Chain-length regulation by WzzE is necessary for, but genetically separable from, cyclic enterobacterial common antigen synthesis

Joseph F. Carr[a], Jennifer S. Rudolf[a]*, Daniel J. Warzecha[a]*, Yohannes H. Rezenom[b], & Angela M. Mitchell[a]#

[a] Department of Biology, Texas A&M University, College Station, Texas, USA

[b] Chemistry Mass Spectrometry Facility, Texas A&M University, College Station, Texas, USA

* JSR and DJW contributed equally to this work. Author order was determined alphabetically.

Address correspondence to Angela M. Mitchell, amitchell@bio.tamu.edu

Running title: Biosynthesis of cyclic enterobacterial common antigen

Abstract

Bacteria of the order *Enterobacterales* (e.g., *Escherichia, Klebsiella, Yersinia*) produce an invariant glycan, enterobacterial common antigen (ECA), that plays a role in maintaining outer membrane (OM) impermeability. ECA is found as the headgroup of a phospholipid and attached to LPS in the OM, and as a periplasmic cyclic form (ECA_{CYC}). WzyE polymerizes ECA repeat units to form the ECA chain with final chain length regulated by WzzE, a class 1 polysaccharide co-polymerase (PCP1). PCP1 function studies have shown interaction of the PCP1 transmembrane helices with the polymerase and a secondary activity found in the periplasmic domain of the PCP1 are required for chain-length regulation. However, WzzE is also necessary for ECA_{CYC} biosynthesis, and it remains unclear why loss of WzzE prevents ECA_{CYC} production but not linear ECA production. Here, we constructed plasmid-based and chromosomal *wzzE* mutants in *E. coli* K-12 with alterations that affect chain-length regulation of other PCP1 and assessed their effects on ECA_{CYC} biogenesis. Our data show loss of chain-length regulation and ECA_{CYC} synthesis from mutations altering transmembrane helix 2 and mutations altering the periplasmic domain of WzzE. We also identified two WzzE variants to the same residue with identical, near wild type linear ECA chain-length regulation but differing effects on ECA_{CYC} production. Specifically, $WzzE^{F104Y}$ produces wild-type levels of ECA_{CYC}, while $WzzE^{F104H}$ produces twofold less ECA_{CYC} than wild type, demonstrating ECA_{CYC} synthesis is genetically separable from chain length regulation. Overall, our results show that chain-length regulation by WzzE is necessary but not sufficient for normal production of ECA_{CYC}.

44 **Importance**

45 The gram-negative cell envelope acts as a permeability barrier excluding many
46 antibiotics. Therefore, understanding envelope synthesis pathways has the potential to identify
47 targets for new antimicrobials. Here, we investigated synthesis requirements for an invariant
48 carbohydrate found throughout *Enterobacterales,* cyclic enterobacterial common antigen
49 (ECA_{CYC}), which is important for maintaining the envelope permeability barrier and relies on a
50 polymerase (WzyE) and co-polymerase (WzzE) for synthesis. We found ECA_{CYC} synthesis
51 depends on WzzE regions necessary for interaction with WzyE and for chain-length regulation.
52 However, we can genetically separate phenotypes of ECA_{CYC} production and chain-length
53 regulation. These data shed light on mechanisms contributing to the envelope permeability
54 barrier and on why the WzyE-WzzE system synthesizes a cyclic carbohydrate while other Wzy-
55 Wzz systems do not.

56

57

58 **Introduction**

59 *Enterobacterales* is an order of gram-negative bacteria containing numerous genera
60 with human pathogens (e.g., *Escherichia, Salmonella, Klebsiella, Yersinia*) (1, 2). These
61 bacteria can result in diseases ranging from respiratory infections, soft tissue infections, urinary
62 tract infections, and endocarditis (3). To survive in their enteric environment, members of this
63 family have a cell envelope and, specifically, an outer membrane which acts as a robust
64 permeability barrier. The outer membrane is an asymmetric bilayer that contains phospholipids
65 in its inner leaflet and predominantly lipopolysaccharides (LPS) in the outer leaflet (4, 5). The
66 LPS creates a strong hydrophilic network of lateral interactions that excludes large compounds
67 and hydrophobic compounds, such as many antibiotics, bile salts, and detergents, from entering
68 the cells (6, 7).

69 In addition to LPS, the outer membrane is decorated with lower-abundant components,
70 including enterobacterial common antigen (ECA) (8). Initially discovered in 1962 by Calvin M.
71 Kunin et al. as a "common antigen" associated with urinary tract infections, ECA is an invariant
72 carbohydrate moiety restricted to the order of *Enterobacterales* and found in all
73 *Enterobacterales* except for endosymbionts with significantly reduced genomes (8-11).
74 Interestingly, ECA_{CYC} plays a role in maintaining the envelope permeability barrier in *E. coli* K-12
75 and is important for the cell's resistance to detergents including bile salts (12). In addition, ECA
76 has been shown to play a role in the pathogenesis of several *Enterobacterales* (13-18). For
77 example, in *Salmonella enterica* serovar Typhimurium, strains lacking ECA exhibit reduced
78 virulence, increased sensitivity to bile salts (14, 15), and establish a low-level persistent
79 infection in mouse models (15).

80 ECA is defined by recurring trisaccharide repeat units (RUs) of GlcNAc (*N*-
81 acetylglucosamine), ManNAcA (*N*-acetyl-D-mannosaminuronic acid), and Fuc4NAc (4-
82 acetamido-4,6-dideoxy-D-galactose) (19). There are three forms of ECA, (i) ECA_{LPS}, which is
83 ECA covalently linked to the LPS core in place of O-antigen; (ii) ECA_{PG}, a cell surface form

84 linked to diacylglycerol-phosphate; and (iii) ECA_{CYC}, a soluble, cyclic form found within the
85 periplasmic space between the inner and outer membranes (10, 19-21). In contrast to the other
86 forms of ECA which have variable number of RU, ECA_{CYC} has a defined size of 4 RUs in
87 *Escherichia coli* K-12 (21). ECA_{PG} and ECA_{LPS} have 1 to 14 RUs and a modal value of 6 to 7 at
88 37 °C (12, 22). Structural analysis by crystallography, nuclear magnetic resonance (NMR), and
89 molecular dynamics suggest ECA_{CYC} can exist in one of two different three-dimensional
90 conformations, a rhomboidal shape and square shape (23, 24).
91 The ECA synthesis genes are located in the *wec* operon. ECA is synthesized through a
92 Wzz/Wzy-dependent pathway analogous to many O-antigen biosynthesis pathways (22, 25-30).
93 After a single trisaccharide ECA RU is assembled and linked to an undecaprenyl-phosphate
94 carrier (8), it is flipped to the outer leaflet of the inner membrane by the flippase, WzxE (26).
95 Then, a complex forms in which translocated undecaprenyl-phosphate-linked ECA subunits are
96 polymerized by WzyE with the number of RU in the final chain regulated by WzzE (31-34).
97 WzzE is an inner membrane protein with a large periplasmic domain that assembles into a
98 homo-octamer (**Figure S1**) (35-38). It belongs to the class 1 polysaccharide co-polymerase
99 (PCP1) family which includes WzzB, FepE, Wzz, and plays a crucial role in controlling
100 carbohydrate chain length during polymerization (39-41). Although PCP1 proteins share low
101 sequence similarity (35), they exhibit a conserved architecture, characterized by a large bell-
102 shaped periplasmic domain and two transmembrane helices (35-38). The C-terminal
103 transmembrane helix (TMH2) of Wzz interacts with Wzy, the polymerase, and this reaction is
104 necessary for O-antigen chain-length regulation (42, 43). However, mutations to Wzz that can
105 disrupt O-antigen chain-length regulation are distributed throughout Wzz's structure and are not
106 confined to TMH2 (44). In *Shigella flexneri*, WzyB and WzzB co-purify together and the second
107 transmembrane region (TMH2) of WzzB and WzzE plays a role in this interaction and in chain
108 length regulation (31). A cryo-EM study in *Pectobacterium atrosepticum* suggested WzyE lies
109 within the central lumen between the transmembrane helices of WzzE (34).

110 The role of WzzE is more complex in ECA synthesis than the role of other PCP1 in O-
111 antigen synthesis. In the absence of WzzE, linear ECA is made with increased variability in
112 number of RU, with both longer and shorter chain lengths being observed (22), whereas PCP1
113 deletion in the context of O-antigen production often leads to a large decrease in chain length
114 with some longer species observed (31, 45-47). In contrast to linear ECA, strains lacking WzzE
115 no longer produce ECA_{CYC}, demonstrating WzzE is necessary for ECA_{CYC} synthesis despite a
116 lack of demonstrated catalytic activity for PCP1 proteins (8, 12). Finally, in *E. coli* K-12, ElyC
117 regulates production of ECA_{PG} in such a way that loss of ElyC causes large increases in ECA_{PG}
118 production (48). This regulation is dependent on the presence of WzzE and likely on levels of
119 ECA_{CYC}, and changes in ECA_{PG} levels with loss of ElyC are much less in the absence of WzzE.
120 Because of this and freeing up of ECA precursors from the synthesis of ECA_{CYC} levels of linear
121 ECA are higher in strains with deletion of *wzzE* than they are in wild type cells (48, 49).
122 Several non-exclusive hypotheses for how PCPs regulate chain length have been
123 proposed (31, 39, 40, 50-54). (i) The inner or outer surface of the PCP may act as a molecular
124 ruler for chain length regulation (51). Specifically, the PCP may bind the extending carbohydrate
125 facilitating catalysis until the polymer length breaks the interaction (55). (ii) The PCP could serve
126 as a scaffold on which multiple polymerase molecules assemble (35). (iii) The elongating
127 carbohydrate could extend in the lumen of the PCP until it reaches the top of the periplasmic
128 domain and is exposed (50). (iv) The PCP could interact with the polymerase to sequester short
129 glycan chains from downstream ligases (i.e., WaaL), increasing the processivity of the
130 polymerization (56). Finally, (v) the loops at the top of the periplasmic bell may move in an
131 alternating up and down arrangement which allows for the polysaccharide chain to extend using
132 a "ratchet"-type mechanism (57).
133 However, the mechanism(s) necessitating WzzE for ECA_{CYC} synthesis are completely
134 unknown. Therefore, we set out to determine what functions of WzzE are required to produce a
135 cyclic carbohydrate in *E. coli* K-12. We used a site-directed mutagenesis approach and

136 biochemical analysis to determine whether the regions important for WzzE-WzyE interaction,
137 ECA chain-length regulation, and any additional mechanisms are necessary for ECA_{CYC}
138 biogenesis. Our findings demonstrate regions important for interaction between WzzE and
139 WzyE and for chain length regulation, are required for ECA_{CYC} biosynthesis. However, our
140 results also demonstrate WzzE chain-length regulation can be genetically separated from
141 synthesis of ECA_{CYC}.
142
143 **Results**
144 **Plasmid-based *wzzE* allows analysis of *wzzE* mutant function.** To identify the functions of
145 WzzE required for ECA_{CYC} synthesis in *E. coli* K-12, we examined the effects of *wzzE* mutations
146 analogous to those known to impact specific functions in other PCP1. We then assessed their
147 influence on the synthesis of both linear ECA and ECA_{CYC}. We first created *wzzE* constructs
148 constitutively expressed by the *tet* promoter on the low copy number pZS21 plasmid (58), which
149 does not contain a Tet repressor (**Figure 1A**). These constructs included wild-type *wzzE*
150 (p*wzzE*), *wzzE* fused with a 3′ glycine-serine linker and triple FLAG tag (p*wzzE*-FLAG) for
151 detection via immunoblot, *wzzE* translationally fused with chloramphenicol acetyltransferase
152 (p*wzzE-cat*), and *wzzE* fused to *cat* and triple FLAG (p*wzzE-cat*-FLAG). Based on the topology
153 of WzzE, we anticipated that the C-terminal fusions would be exposed in the cytoplasm, thereby
154 allowing Cat to function. Thus, this construct allowed us to identify defects in WzzE folding,
155 stability, and octamerization, as measured by their effect on Cat through chloramphenicol
156 resistance. This approach was informed by a translational coupling methodology previously
157 used a tool in *E. coli* (59, 60). This assay is particularly useful because the addition of the bulky
158 Cat (26KDa) protein likely imposes stress on WzzE folding. Additionally, the formation of the
159 functional cat trimer (61) is more likely if WzzE is properly multimerized.
160 After producing the constructs, we tested the function of fusions with each tag by
161 determining whether they complemented a *wzzE* deletion. In the absence of *yhdP*, ECA causes

162 severe sensitivity to vancomycin in a *wzzE* dependent manner (12). Therefore, we expressed
163 the *wzzE* constructs in a Δ*wzzE* Δ*yhdP* strain and tested for vancomycin sensitivity. Our results
164 demonstrated each construct complemented Δ*wzzE* by restoring vancomycin sensitivity in the
165 Δ*wzzE* Δ*yhdP* strain (**Figure 1B**). Additionally, p*wzzE-cat* conferred chloramphenicol resistance
166 to the strain. To determine if the p*wzzE*-FLAG construct could be detected, we immunoblotted
167 for FLAG and for GroEL as a loading control (**Figure 1C**) and determined that the fusion protein
168 could be clearly detected.

169 To determine whether the function of WzzE as a chain-length regulator is complemented
170 by the p*wzzE* construct, we performed immunoblots for linear ECA (**Figure 1D**). Our results
171 revealed a range of ECA bands, with increasing molecular weight indicating molecules with
172 more RUs. The distribution of ECA chain lengths was wider in a Δ*wzzE* mutant (Lane 3) than in
173 wild-type cells (Lane 2) with an increase in longer and shorter ECA bands, indicating loss of
174 chain-length regulation. We also observed more linear ECA with Δ*wzzE* as we would expect
175 from previous findings (48, 62). For simplicity, we will refer to these changes as "loss of chain-
176 length regulation". However, expression of p*wzzE* (Lane 5) but not the empty vector (Lane 4)
177 restored chain-length regulation, albeit with a lower modal chain length than the wild-type strain
178 and much less linear ECA overall. These changes may suggest overactivity of WzzE due to
179 multi-copy expression. The same pattern was observed for the p*wzzE-cat* construct (Lane 7);
180 however, the p*wzzE*-FLAG and p*wzzE-cat*-FLAG constructs resulted in chain-length regulation
181 closer to that of wild type.

182 Finally, to confirm that ECA_{CYC} synthesis is restored by the plasmid-expressed *wzzE*, we
183 expressed the p*wzzE* and empty vector plasmids in a Δ*wzzE* Δ*wecH* strain. WecH has been
184 shown to non-stoichiometrically acetylate ECA (63) and deleting *wecH* allows ECA_{CYC} to be
185 quantitated as a single mass species (12, 63). Unlike the two linear forms of ECA, ECA_{CYC} is
186 uncharged and cannot be detected by immunoblotting. Therefore, we then grew cultures with
187 either a normal (^{14}N) or heavy (^{15}N) nitrogen source, combined labeled and unlabeled samples,

188 purified ECA_{CYC}, and assayed ECA_{CYC} levels by liquid chromatography-mass spectrometry (LC-
189 MS). ECA_{CYC} has an monisotopic mass of 2428.9 Da and the labeling with heavy nitrogen
190 causes a 12 Da increase in this weight due to the presence of 12 nitrogen atoms in ECA_{CYC} (12)
191 (**Figure S2AB**). For this type of assay, each peak consists of multiple minor peaks due to the
192 stochiometric incorporation of naturally occurring isotopes (e.g., ^{14}C) (see **Figure 1EF**, **S2C**).
193 When labeled and unlabeled samples carrying p*wzzE* were combined, we detected peaks at
194 both the unlabeled and labeled m/z (**Figure 1E**). However, when the labeled p*wzzE* sample was
195 combined with an unlabeled sample carrying the empty vector, only the ^{15}N labeled ECA_{CYC}
196 peak was detected (**Figure 1F**). ECA_{CYC} production from the p*wzzE* and p*wzzE-cat* constructs
197 was very similar (**Figure S2C**). These data confirm plasmid-based expression of *wzzE* can
198 facilitate the synthesis of ECA_{CYC}, making this an appropriate system to investigate the role of
199 WzzE in ECA_{CYC} biosynthesis.

200 **Conserved glycines in TMH2 are necessary for WzzE function.** After validating assays for
201 WzzE function, we set off to determine the role of TMH2 in WzzE function. TMH2 has been
202 shown to be important for interaction between WzzB and Wzy in O-antigen synthesis (31).
203 Previous mutagenesis studies of WzzB suggested a conserved GXXXG motif in the TM2 region
204 of WzzB from *Shigella flexneri Y* was important for function (64), likely through effects on helix-
205 helix packing (65-67). A WzzB double mutant G305A/G311A WzzB in *Shigella flexneri Y*
206 (G306A/G312A in *E. coli*) targeting this motif conferred short-chain O antigen (68). Based on
207 these studies, we hypothesized that the GXXXG motif located in TMH2 of WzzE plays a
208 significant role in mediating the helix-helix packing of the transmembrane region and interaction
209 with WzyE.

210 To test this hypothesis, we created TMH2 mutations (**Figure S1B**) in the p*wzzE*
211 constructs described in **Figure 1A**. We first sought to determine whether the levels of WzzE
212 protein produced by p*wzzE*-FLAG or p*wzzE-cat*-FLAG were altered by the mutations (**Figure**

213 **2AB, Figure S3A**). Our results showed a small but significant decrease in protein abundance
214 for our WzzE$^{GG333/4LL}$ in the WzzE-FLAG fusion (**Figure 2AB**), whereas no decrease is apparent
215 in the WzzE-Cat-FLAG fusion (**Figure S3A**). As the levels of the WzzE$^{GG333/4LL}$ variant are 82%
216 of wild type on average and as high as 98% in individual replicates, the decreased level of this
217 protein may not be biologically relevant to its function. To determine whether the mutations
218 might have caused alterations to protein stability or multimerization, we utilized our p*wzzE-cat*
219 construct to assay chloramphenicol resistance conferred by the fusion protein. None of our
220 mutations showed any chloramphenicol sensitivity (**Figure 2C, Figure S3B**), suggesting that
221 they are largely stable and multimerized, despite the small decrease in abundance for
222 WzzE$^{GG333/4LL}$. Next, we set out to determine if the mutations affected WzzE's ability to regulate
223 ECA chain length by performing ECA immunoblots. Three of the changes to TMH2, WzzE^{A323G},
224 WzzE$^{GG333/4AA}$, and WzzE^{G339A} (Lanes 4,5, and 7) did not affect ECA chain length regulation
225 (**Figure 2D**). However, WzzE$^{GG333/4LL}$ (Lane 6) showed a similar or greater disruption of chain
226 length regulation to that observed in the empty vector control (Lane 3).

227 **The TMH2 GXXXG motif affects ECA$_{CYC}$ synthesis.** Given the effect of WzzE$^{GG333/4LL}$ on ECA
228 chain length regulation, we next decided to examine the effect of the TMH2 mutations on
229 ECA$_{CYC}$ biogenesis. Therefore, we utilized LC-MS to detect and quantitate ECA$_{CYC}$ in purified
230 samples as in **Figure 1EF**. We expressed the TMH2 mutants or controls in a Δ*wzzE* Δ*wecH*
231 strain background and labeled the wild-type *wzzE* strain with ^{15}N, while the mutants were grown
232 with ^{14}N. Our quantification demonstrated WzzE^{A323G}, WzzE$^{GG333/4AA}$, and WzzE^{G339A} did not
233 significantly change ECA$_{CYC}$ levels, whereas the WzzE$^{GG333/4LL}$ mutant produced significantly
234 reduced levels of ECA$_{CYC}$, similar to those of the empty vector control (**Figure 3A**). Together
235 with our chain length regulation data, these results demonstrate the chain-length regulating
236 activities of WzzE found in TMH2 are required for ECA$_{CYC}$ synthesis.

237 To visualize the potential effect of the mutants on WzzE and its complex with WzyE, we
238 predicted the structure of the WzyE:WzzE(8) complex for wild-type WzzE and WzzE$^{GG333/4LL}$
239 using AlphaFold 3 (69) (**Figure 3B** & **S4**, **Table S1**) and aligned wild-type WzzE and
240 WzzE$^{GG333/4LL}$ monomers. When we closely examined the transmembrane region of the
241 structure, the model suggested a 2 Å increase in distance between TMH1 and TMH2 in
242 WzzE$^{GG333/4L}$ likely due to steric interference (**Figure 3B**). This suggests a possible hypothesis
243 that the arrangement of the TMHs is responsible for the phenotype of this mutant.

244 **The periplasmic domain of WzzE functions in chain length regulation.** After assessing the
245 importance of WzzE TMH2 for ECA_{CYC} biosynthesis, we wanted to determine whether the
246 chain-length regulating activities found in the periplasmic domain of WzzE were necessary for
247 ECA_{CYC} biosynthesis. There are 21 periplasmic residues conserved between WzzE, FepE, and
248 WzzB in *E. coli*, which are likely to be important for the functions of the PCP1 proteins (35).
249 Furthermore, we hypothesized that other non-conserved residues might also contribute to
250 WzzE's role in chain-length regulation or ECA_{CYC} synthesis functions. Therefore, we made
251 mutations in a conserved cluster of predominantly hydrophobic residues that lie within the core
252 of the α/β domains in strands β1 and β4, as well as helix α2 and the N terminus of helix α6 (see
253 **Figure S1B**). We also targeted non-conserved residues throughout WzzE in positions similar to
254 those that conferred changes to O-antigen chain length in other PCP1s (35).
255 We started characterizing the periplasmic mutants (**Figure S1B**) by investigating WzzE
256 protein levels from the p*wzzE*-FLAG construct. One of the mutants (WzzE$^{\Delta F104}$) had
257 undetectable WzzE-FLAG protein levels, while the other mutants retained wild-type levels of
258 accumulation (**Figure 4AB**). When we tested the levels of WzzE-Cat-FLAG produced from this
259 fusion, we detected a faint band at the molecular weight of the fusion and a second lower
260 molecular weight band for WzzE$^{\Delta F104}$, suggesting that the C-terminal tags were cleaved from the
261 protein (**Figure S3C**). Given the cleavage of the tags, no conclusion can be drawn for the

262 stability of WzzE$^{\Delta F104}$. All other protein levels were similar to wild type. Then, we tested whether
263 the *wzzE* constructs showed changes to protein stability and/or multimerization through
264 chloramphenicol resistance using the p*wzzE-cat* construct. WzzE^{K162A}, WzzE^{L259A}, and
265 WzzE^{F104Y} resulted in partial or full chloramphenicol sensitivity in the WzzE-Cat fusion (**Figure**
266 **S3D**) but showed full chloramphenicol resistance in the WzzE-Cat-FLAG fusion that better
267 complements for linear chain length regulation (**Figure 4C**). Thus, it is possible these mutants
268 could be slightly unstable or have a decreased ability to multimerize. The other mutants
269 conferred resistance to chloramphenicol similar to wild type in both fusions (**Figure 4C, S3D**).
270 WzzE$^{\Delta F104}$ showed wild-type levels of chloramphenicol resistance, likely due to cleavage of the
271 C-terminal tags. We tested an additional 18 mutations in the periplasmic region that caused
272 chloramphenicol sensitivity in the WzzE-Cat background, emphasizing the importance of the
273 periplasmic domain on the structural stability of WzzE (**Figure S5**).
274 After characterizing the effect of the periplasmic mutations on WzzE levels and stability,
275 we investigated their effect on ECA chain-length regulation (**Figure 4D**). The WzzE$^{\Delta F104}$ and
276 WzzE^{K162A} mutants (lanes 4 and 7) prevented regulation of ECA chain length. In contrast,
277 WzzE^{F104Y}, WzzE^{F104H}, WzzE^{K162S}, and WzzE^{K162T} (lanes 5, 6, 8 and 9) demonstrated wild-type
278 ECA chain length regulation. The WzzE^{L259A} and WzzE^{L259I} mutants did not affect ECA chain-
279 length regulation despite the lack of chloramphenicol resistance observed for WzzE^{L259A}. It may
280 be that this protein is partially destabilized but able to function in the absence of the Cat fusion.
281 Overall, these mutants provide an interesting range of phenotypes useful for the investigation of
282 ECA$_{CYC}$ synthesis.

283 **Periplasmic residues of WzzE are important for ECA$_{CYC}$ biogenesis.** We then assessed the
284 impact of the periplasmic mutants on ECA$_{CYC}$ biosynthesis by quantifying ECA$_{CYC}$ levels using
285 LC-MS. Both the WzzE$^{\Delta F104}$ and WzzE^{K162A} mutants exhibited impaired linear chain-length
286 regulation and produced little to no ECA$_{CYC}$ (**Figure 5A**). This lack of ECA$_{CYC}$ production could

287 be due to either (i) insufficient protein accumulation and/or stability, or (ii) the disruption of chain-
288 length regulating activities in these mutants. In contrast, WzzE^{F104Y}, WzzE^{K162S}, WzzE^{K162T},
289 WzzE^{L259A}, and WzzE^{L259I} had similar levels of ECA$_{CYC}$ biosynthesis to wild type. WzzE^{F104H}
290 showed a highly variable level of ECA$_{CYC}$ despite having wild type chain-length regulation and
291 protein levels, suggesting that this mutation might have some sort of variable effect on ECA$_{CYC}$
292 synthesis. These data highlight the importance of the periplasmic domain of WzzE on ECA$_{CYC}$
293 synthesis, likely through its structural role, chain-length regulating activity, and possibly a
294 function specific to ECA$_{CYC}$ synthesis.

295 Given the varying effects of the F104 mutations on ECA$_{CYC}$ synthesis and linear ECA
296 chain length regulation, we decided to visualize the possible effect of these mutations on
297 WzzE's structure using AlphaFold 3 (69) and aligned the wild type and mutant complexes
298 (**Figure 5B** & **S4**, **Table S1**). When examining these structures, we observed that F104 is
299 predicted to make contacts with a neighboring loop. It is possible that alterations in these
300 contacts are responsible for the changes in function for variants at this residue.

301 **ECA$_{CYC}$ biogenesis is genetically separable from chain-length regulation.** Our results,
302 using plasmid-based expression of WzzE, identified several point mutations with differing effects
303 on ECA$_{CYC}$ function. Most interestingly, we identified a change to TMH2 (WzzE$^{GG333/4LL}$) that
304 prevented chain-length regulation and ECA$_{CYC}$ synthesis, despite maintaining protein stability
305 and near wild type protein accumulation. Additionally, several mutations to residue F104
306 exhibited varying effects on WzzE accumulation, chain-length regulation, and ECA$_{CYC}$
307 production. However, plasmid-based expression of wild-type *wzzE* resulted in lower levels of
308 linear ECA and lower chain length than expression of *wzzE* from its native locus (**Figure 1D**),
309 which may have influenced the phenotypes of the mutants assayed. In addition, we observed a
310 high level of variability in ECA$_{CYC}$ levels produced in several of the periplasmic mutants, which
311 we reasoned could be because of the non-native expression of WzzE. Therefore, we used

312 CRISPR-Cas9 to produce selected point mutations to *wzzE* at its native locus and assayed their
313 effect on linear ECA and ECA_{CYC}.
314 We first performed LC-MS on purified ECA_{CYC} to determine the effect of the
315 chromosomal mutants on ECA_{CYC} biosynthesis. The $WzzE^{\Delta F104}$ and $WzzE^{GG333/4LL}$ chromosomal
316 mutants produced very little to no ECA_{CYC}, similar to what is observed with Δ*wzzE* (**Figure 6A**).
317 In contrast, $WzzE^{F104H}$ produced measurable levels of ECA_{CYC} that were about 2-fold lower than
318 the wild type. Finally, $WzzE^{F104Y}$ did not cause a significant change in ECA_{CYC} levels when
319 compared to the wild type. We then assayed linear ECA chain length regulation by immunoblot
320 (**Figure 6B**). Consistent with the plasmid-based system, $WzzE^{\Delta F104}$ and $WzzE^{GG333/4LL}$ were not
321 able to regulate linear ECA chain length and dysregulated chain length was observed.
322 Interestingly, the levels of linear ECA and chain-length dysregulation were even greater than
323 that of Δ*wzzE*, suggesting that these mutants produced a non-functional protein that interacted
324 with ECA synthesis differently than loss of the WzzE protein. $WzzE^{F104H}$ and $WzzE^{F104Y}$ each
325 had identical, intact ECA chain-length regulation, which was close to that of wild-type WzzE but
326 with some increase in chain-length. The identical regulation of linear ECA chain-length by these
327 mutants but differing effects of ECA_{CYC} synthesis demonstrate that levels of ECA_{CYC} production
328 are genetically separable from linear chain-length regulation.
329 We were surprised that the dysregulation of ECA chain length, and even more the levels
330 of ECA, in the $WzzE^{\Delta F104}$ and $WzzE^{GG333/4LL}$ mutants was greater than the Δ*wzzE* control and we
331 wondered whether in a Δ*wzzE* strain the remaining PCP1s could partially regulate ECA chain
332 length through interaction with WzyE. In this scenario, the nonfunctional WzzE mutants could
333 sequester WzyE from the other PCP1s. Therefore, we assayed whether we could cause further
334 disruption of ECA chain-length regulation in a Δ*wzzE* strain by deleting *wzzB* and/or *fepE*.
335 However, our results demonstrated Δ*wzzB*, Δ*fepE*, and Δ*wzzB* Δ*fepE* in a Δ*wzzE* strain has no
336 effect on ECA chain-length regulation (**Figure 6C**). Nevertheless, there was a small increase in
337 the amount of linear ECA observed in the Δ*wzzE* Δ*wzzB* Δ*fepE* strain. Taken together our

338 results demonstrate that mutations to *wzzE* disrupting linear ECA chain length regulation can
339 prevent synthesis of ECA_{CYC}, and that production of normal levels of ECA_{CYC} is genetically
340 separable from regulation of ECA chain length.
341
342 **Discussion**
343 In this study, we constructed and characterized a set of *wzzE* site-directed mutants to
344 identify what functions of WzzE are necessary to produce ECA_{CYC}. Previous mutagenesis
345 studies have suggested regions of *Shigella flexneri* WzzB that are important for proper function
346 (33, 42). However, despite the multitude of hypotheses for PCP1s function, there have been no
347 reports on WzzE's role in producing a cyclic carbohydrate, ECA_{CYC}. Here, we demonstrate that
348 changes to amino acids F104 and K162 in the periplasmic region and to the GXXXG motif in
349 TMH2 can impair ECA chain-length regulation and ECA_{CYC} production, suggesting WzzE's
350 ability to regulate ECA chain length is necessary for ECA_{CYC} synthesis. Moreover, our data
351 demonstrate that regulation of linear ECA chain length and the level of ECA_{CYC} production are
352 genetically separable, as we have identified two mutations to the same WzzE residue (F104)
353 which have differing effects on ECA_{CYC} levels but produce identical regulation of linear ECA
354 chain length. This shows that regulation of chain length by WzzE is necessary but not sufficient
355 for production of normal levels of ECA_{CYC}.
356 Interestingly, despite the low sequence identity amongst PCP1s, they have highly similar
357 three-dimensional structures that are marked by two transmembrane domains and an α-helix/β-
358 sheet domain within the periplasmic space that extends ~100 Å (35-38). Additionally, PCP1s
359 have a conserved GXXXG motif that has been identified to be important for chain length
360 regulation, as a G305A/G311A double mutant in *Shigella flexneri* WzzB that encompassed and
361 extended outside of this motif resulted in short-chain O antigen production (64, 68). Notably, this
362 same study showed that single, double, and triple Gly to Ala mutations in this motif did not affect
363 O-antigen chain length apart from the G305A/G311A mutation (37, 64). Given this, we focused

364 on residues within and around the GXXXG motif, characterized their effect on chain-length
365 regulation and ECA_{CYC} biogenesis, and identified two adjacent, conserved glycine residues in
366 the GXXXG motif of WzzE—analogous to G305/G306 in WzzB—as critical for ECA_{CYC}
367 biosynthesis. We observed no changes to either chain length regulation or ECA_{CYC} production
368 when we introduced an alanine to both sites. However, we did observe dysregulation of ECA
369 chain length and prevention of ECA_{CYC} synthesis by introducing two adjacent leucines.
370 In other proteins, GXXXG motifs have been shown to mediate dimerization and
371 multimerization of transmembrane domains (70, 71). Thus, this motif may be required for proper
372 interaction between TMH1 and TMH2 and the long side chain on the leucine could push TMH1
373 and TMH2 further apart, reducing their interaction. In support of this idea, our structural
374 modeling suggests increased distance between TMH1 and TMH2 in the $WzzE^{GG333/4LL}$ mutant
375 (**Figure 3D**, **S6A**). When we aligned the monomers of WzzE that most closely interacted with
376 WzyE, we found that the angle of WzyE was rotated in the plane perpendicular to the
377 membrane in WzyE-$WzzE^{GG333/4LL}$ model compared to the wild type (**Figure S6BC**). The
378 distance in the structure between the TMH of opposite WzzE monomers was 2.0 to 2.7 Å
379 smaller in the $WzzE^{GG333/4LL}$ model than in the wild-type model. In compensation, WzyE takes on
380 a more condensed structure (1.4 Å narrower) in the $WzzE^{GG333/4LL}$ model. As this type of rotation
381 and structure change could impair WzyE function, this prediction provides a possible hypothesis
382 that $WzzE^{GG3334/LL}$ has impaired interaction with WzyE. Alternatively, the shift of the TMH may
383 affect the less clearly defined activity of WzzE that regulates chain length.
384 In contrast to the TMH2 mutants, many of the mutations that we made in the periplasmic
385 domain of WzzE cause the protein to become unstable as assayed by chloramphenicol
386 resistance conferred by the WzzE-Cat fusion, suggesting the periplasmic domain is critical for
387 WzzE's structural integrity and/or multimerization. Additionally, the periplasmic domain has been
388 proposed to be the "chain-length regulating" portion of WzzE. Studies with other PCP1s have
389 shown that multiple, surface-exposed regions on both the interior and exterior space of the

390 periplasmic domain affect chain-length regulation, resulting in either shorter or longer O-antigen
391 chains (36, 52, 64, 72, 73). More recently, increasing evidence suggests that the inner surface
392 of the WzzE octamer may bind to the elongating carbohydrate to facilitate growth of the chain
393 (34, 40, 56).

394 The internal cavity of WzzE possesses two highly concentrated negatively charged
395 bands of residues separated by a band of neutrally-charged-flexible-loops in between, which
396 may act to bind the growing carbohydrate chain (57). A loop at the top of the periplasmic domain
397 has been suggested to "rachet" up and down to facilitate carbohydrate movement over this
398 binding surface (57). While F104 is found buried inside the bottom of the "bell" of the WzzE
399 periplasmic domain and may have indirect effects on carbohydrate binding, K162 is located
400 near the lower negatively charged band and the membrane, and L259 is located in the loop that
401 has been proposed to drive the ratchet mechanism. Our mutations to L259 did not affect chain-
402 length regulation; however, loss of charge at K162 did prevent chain-length regulation and
403 ECA_{CYC} synthesis, supporting this model of ECA chain-length elongation. As to how the
404 *wzzE*$_{K162A}$ and *wzzE*$_{F104\Delta}$ mutations lead to loss of ECA_{CYC} production, we speculate that the
405 chain-length regulating activity of WzzE—related to the carbohydrate binding and ratcheting
406 mechanism—is necessary for ECA_{CYC} synthesis.

407 In contrast, the *wzzE*$_{F104H}$ mutation reduces ECA_{CYC} synthesis while maintaining chain-
408 length regulation. There is a small difference in linear ECA chain length between this mutant
409 and the wild type; however, the chain length and regulation thereof is identical to that of the
410 *wzzE*$_{F104Y}$ mutant which produces wild-type levels of $ECA_{CYC.}$ These data demonstrate that
411 chain-length regulation by WzzE is not sufficient for normal levels of ECA_{CYC} production and are
412 an indication that WzzE may play a role in ECA_{CYC} synthesis beyond that of chain-length
413 regulation. One possible mechanism would be that WzzE bends ECA molecules of the correct
414 size into an orientation that can be spontaneously cyclized. Another possibility is that WzzE
415 allows the binding of another protein (i.e., a cyclase) that can cyclize ECA chains of the

416 appropriate length. Finally, it is possible that WzzE has specific features that promote the
417 formation of ECA chains with 4 RUs that can later be cyclized to form ECA_{CYC}. Further, genetic
418 and biochemical experiments will be necessary to distinguish between these possibilities and to
419 solidify an additional role for WzzE in ECA_{CYC} biogenesis.

420 Lastly, we would like to draw attention to the observation that the $WzzE^{F104\Delta}$ and
421 $WzzE^{GG333/4LL}$ mutants had greater linear ECA levels than is observed when deleting *wzzE*
422 entirely. We further showed that this is not due to the influence of other PCP1s found in *E. coli*,
423 though there is a slight increase in linear ECA levels in a triple mutant of all the PCP1 genes.
424 Thus, why $WzzE^{F104\Delta}$ and $WzzE^{GG333/4LL}$ have increased linear ECA levels remains a mystery. It
425 is possible that the inactive WzzE octamers sequester some growing ECA chains resulting in
426 the increased variability in chain length. Although this can be investigated in the future, the
427 difference between the Δ*wzzE* strain and these mutants provides evidence that the mutants
428 produced a protein that could not regulate chain length and could not synthesize ECA_{CYC},
429 increasing the strength of our conclusion that chain-length regulation is important for ECA_{CYC}
430 synthesis.

431 Through this work, we have examined only a very small number of the residues in
432 WzzE, so it is quite likely there are mutations that cause effects solely for ECA_{CYC} production.
433 Future work to identify these mutations will provide the need evidence to determine the
434 mechanism of ECA_{CYC} synthesis. Nevertheless, our work provides insight into requirements for
435 carbohydrate cyclization and chain-length regulation. We propose coordinated action between
436 WzyE and WzzE, as well as chain-length regulation, is necessary for ECA_{CYC} biosynthesis and
437 suggest that charged residues in the periplasmic base help to facilitate chain elongation and
438 ECA_{CYC} biosynthesis. We also demonstrated that chain-length regulation can be genetically
439 separated from synthesis of normal levels of ECA_{CYC}, suggesting that WzzE may play a role in
440 ECA_{CYC} biosynthesis beyond that of chain-length regulation. Thus, our work provides insight into
441 the characteristics of WzzE that make it necessary for synthesis of a cyclic carbohydrate.

442

443 **Materials and Method**

444 **Strains and growth conditions**. The strains used in this study are listed in **Table S2**. Cultures
445 were inoculated and grown at 37 °C in LB Lennox unless otherwise noted, and as necessary,
446 cultures were supplemented with 20 mg/L chloramphenicol, 25 mg/L kanamycin, 10 mg/L
447 tetracycline, or a combination thereof. Insertion-deletion alleles containing kanamycin resistance
448 cassettes from the Keio collection (74) were moved to isogenic backgrounds using P1vir
449 transduction (75). Clean deletions were constructed utilizing the Flp recombinase-FRT system
450 as has described (76). Chromosomal mutagenesis in WzzE was constructed using a CRISPR
451 Cas-9 system harboring a pCRISPR::protospacer plasmid in W3110 containing a heat shock
452 inducible λcI857 lysogen as has been described (77, 78). The pCas9 (Addgene #42876) and
453 pCRISPR (Addgene #42975) plasmids were a kind gift from Professor Luciano Marraffini
454 (Rockefeller University). Oligos used to form the protospacer and repair templates used to make
455 the desired mutations are listed in **Table S3**. Prior to chromosomal mutagenesis, *metE*::Tn*10*
456 was linked to *wzzE* (79) allowing the mutations to be moved to MG1655 through P1vir
457 transduction. Strains carrying *metE*::Tn*10* were supplemented with 1μm cobalamin (vitamin
458 B12) when grown in minimal media (80).

459 **Plasmids, cloning, and site directed mutagenesis.** All primers used in this study are listed in
460 **Table S4**. To sub-clone the p*wzzE* expression plasmid, chromosomal *wzzE* with an ochre stop
461 codon was PCR amplified using primers pZS21-wzzE_fwd and wzzE-pZS21_rev. The low copy
462 number pZS21 plasmid (58) was amplified using primers pZS21_fwd and pZS21_rev. The
463 plasmid was assembled using HiFi Assembly Master Mix (New England Biolabs) as per the
464 manufacturer's instructions. To construct p*wzzE-cat*, *wzzE* without a stop codon was amplified
465 using primers pZS21-wzzE_fwd and wzzE-cat_rev. The chloramphenicol acetyltransferase
466 gene, *cat*, was amplified from pBAD33 (81) using primers wzzE-cat_fwd and cat-pZS21_rev.

467 These fragments were assembled into pZS21 amplified as before using HiFi Assembly. To
468 construct p*wzzE*-FLAG, p*wzzE* was amplified using primers p*wzzE*_x3F_Fwd and
469 p*wzzE*_x3F_Rvr, which contained sequence homology to a glycine serine linked FLAG tag that
470 incorporated a stop codon. The glycine serine linked FLAG tag was a gene block fragment
471 (**Table S3**) assembled into the p*wzzE* plasmid through HiFi Assembly. To clone the p*wzzE-cat-*
472 FLAG plasmid, *cat* was amplified using the WzzE:Cmr Fp and WzzE:Cmr Rp and the p*wzzE-*
473 FLAG plasmid was amplified using the Cmr: GS-x3F FP and wzzE-cat_rev. The fragments were
474 assembled using HiFi Assembly.

475 Site-direct mutations were constructed using Q5 PCR mutagenesis (New England
476 Biolabs). Briefly, non-overlapping primers containing the mutation of interest were used to
477 amplify the indicated pZS21-sanA plasmids (pZS21-*wzzE*, pZS21-*wzzE-cat*, and pZS21-*wzzE-*
478 FLAG). The PCR products were then digested with DpnI (New England Biolabs),
479 phosphorylated with T4 polynucleotide kinase (New England Biolabs) and self-ligated using T4
480 DNA ligase (New England Biolabs). Ligations were transformed into Turbo cloning cells (New
481 England Biolabs) and the mutations were confirmed by sequencing (Plasmidsaurs). The primers
482 used for each site-directed mutant are described in **Table S4**. Modeling of the predicted mutant
483 structure was performed using AlphaFold 3 (69) with eight monomers of WzzE and one of
484 WzyE. Relevant confidence scores are given in **Table S1** and **Figure S2**.

485 **Antibiotic Sensitivity.** Efficiency of plating (EOP) assays were performed by plating 10-fold
486 dilutions of overnight cultures on LB plates containing the indicated antibiotics using a 48-pin
487 tool. Plates were incubated overnight at 30 °C before imaging.

488 **Immunoblot Analysis.** Immunoblots for WzzE-FLAG and ECA were conducted as previously
489 described with minor modification (12, 82). For analysis of WzzE-FLAG protein levels, samples
490 with equal OD_{600} from overnight cultures of the indicated strains were lysed by boiling for 5
491 minutes in sample loading buffer and were loaded on 12% SDS-PAGE gels. Samples for ECA

492 analysis were prepared and immunoblotted as previously described. The antibodies used for
493 blotting were as follows: M2 αFLAG (Millipore Sigma, 1:50,000 dilution), αGroEL (Millipore
494 Sigma, 1: 60,000 dilution), αECA (a gift from Professor Renato Morona (University of Adelaide),
495 1:30,000), rabbit α-mouse (Prometheus, 1:100,000 dilution), goat α-rabbit (Prometheus,
496 1:100,000 dilution). Levels of WzzE-FLAG were quantitated by densitometry using ImageJ as
497 has been previously described (12) on immunoblots with the lowest exposure at which WzzE
498 bands could be detected.

499 **Quantification of Cyclic ECA**. Quantification of ECA_{CYC} was performed as previously
500 described with minor modifications (12). Briefly, cells were grown in M63 glucose minimum
501 media without nitrogen supplemented with 0.2% $(^{14}NH_4)_2SO_4$ or $(^{15}NH_4)_2SO_4$ and ECA was
502 purified as previously described. Sample analysis was performed using a Thermo Scientific Q
503 Exactive Focus and liquid chromatography unit (UltiMate 3000 RS). Purified samples were
504 dissolved in 0.1% formic acid and separated by injecting 20 µL of sample into the Synergi
505 Hydro-RP (4.6 x 250 mm; 4.0 µm) column (Phenomenax). The mobile phase consisted of 0.1%
506 Formic acid (eluent A) and methanol (eluent B). The flow rate was set at 800 µL/min with the
507 following gradient: 0–5.0 min 5% B, 5.0-12.0 min 5-60% B, 12-14 min 60-90% B, 14-17 min 90%
508 B, and returned back to 5% B in 0.1 min and held for 5 min at 5% B. The Q Exactive Focus ESI
509 source was operated in full MS (100-1000 m/z). The mass resolution was tuned to 70000
510 FWHM at m/z 200. The spray voltage was set to 3.3 kV in negative mode, and the sheath gas
511 and auxiliary gas flow rates were set to 60 and 20 arbitrary units, respectively. The transfer
512 capillary and auxiliary gas heater temperatures were held at 380 and 300 °C, respectively. The
513 S-Lens RF level was set at 70 V. Exactive Series 2.11 /Xcalibur 4.2.47 software was used for
514 data acquisition and processing. For relative quantification, the ratio of the areas of the ^{14}N to
515 ^{15}N ECA_{CYC} peaks was calculated.
516

517 **Acknowledgements**

518 Thank you to the members of the Mitchell Lab for helpful and productive discussions and for
519 critical reading of our manuscript. We also thank Drs. Jolene Ramsey, Deborah Siegele,
520 Jennifer Herman, and Ryland Young (Texas A&M University), as well as members of the
521 Ramsey lab, for their insightful comments. The ECA antibody was a kind gift from Professor
522 Renato Morana (University of Adelaide). The pCas9 and pCRISPR plasmids were a kind gift
523 from Professor Luciano Marraffini (Rockefeller University) through the Addgene Repository. This
524 work was supported by the National Institute of Allergy and Infectious Disease under award
525 number R01-AI155915 (to A.M.M.).

526

527 **References**

528 1. Doern CD. 2024. Chapter 2 - Classification of medically important bacteria, p 9-21. *In*
529 Tang Y-W, Hindiyeh MY, Liu D, Sails A, Spearman P, Zhang J-R (ed), Molecular Medical
530 Microbiology (Third Edition) doi:https://doi.org/10.1016/B978-0-12-818619-0.00029-0.
531 Academic Press.

532 2. Adeolu M, Alnajar S, Naushad S, S. Gupta R. 2016. Genome-based phylogeny and
533 taxonomy of the '*Enterobacteriales*': proposal for *Enterobacterales* ord. nov. divided into
534 the families *Enterobacteriaceae*, *Erwiniaceae* fam. nov., *Pectobacteriaceae* fam. nov.,
535 *Yersiniaceae* fam. nov., *Hafniaceae* fam. nov., *Morganellaceae* fam. nov., and
536 *Budviciaceae* fam. nov. International Journal of Systematic and Evolutionary
537 Microbiology 66:5575-5599.

538 3. Ramirez D, Giron M. 2024. Enterobacter Infections, StatPearls. StatPearls Publishing
539 Copyright © 2024, StatPearls Publishing LLC., Treasure Island (FL).

540 4. Maher C, Hassan KA. 2023. The Gram-negative permeability barrier: tipping the balance
541 of the in and the out. mBio 14:e01205-23.

542 5. Silhavy TJ, Kahne D, Walker S. 2010. The bacterial cell envelope. Cold Spring Harb
543 Perspect Biol 2:a000414.
544 6. Farfour E, Dortet L, Guillard T, Chatelain N, Poisson A, Mizrahi A, Fournier D, Bonnin
545 RA, Degand N, Morand P, Janvier F, Fihman V, Corvec S, Broutin L, Le Brun C, Yin N,
546 Héry-Arnaud G, Grillon A, Bille E, ..., On Behalf Of The Gmc Study G. 2022.
547 Antimicrobial Resistance in *Enterobacterales* Recovered from Urinary Tract Infections in
548 France. Pathogens 11:356.
549 7. Aminov RI. 2010. A brief history of the antibiotic era: lessons learned and challenges for
550 the future. Front Microbiol 1:134.
551 8. Rai AK, Mitchell AM. 2020. Enterobacterial Common Antigen: Synthesis and Function of
552 an Enigmatic Molecule. mBio 11:mBio.01914-20.
553 9. Jenkins C, Rentenaar RJ, Landraud L, Brisse S. 2017. 180 - *Enterobacteriaceae*, p
554 1565-1578.e2. *In* Cohen J, Powderly WG, Opal SM (ed), Infectious Diseases (Fourth
555 Edition) doi:https://doi.org/10.1016/B978-0-7020-6285-8.00180-5. Elsevier.
556 10. Kunin CM. 1963. Separtion, chacterization, and Biological Significance of a Common
557 Antigen in *Enterobacteriaceae*. J Exp Med 118:565-86.
558 11. Kunin CM, Beard MV, Halmagyi NE. 1962. Evidence for a Common Hapten Associated
559 with Endotoxin Fractions of *E. coli* and other *Enterobacteriaceae*. Proc Soc Exp Biol Med
560 111:160-166.
561 12. Mitchell AM, Srikumar T, Silhavy TJ. 2018. Cyclic Enterobacterial Common Antigen
562 Maintains the Outer Membrane Permeability Barrier of *Escherichia coli* in a Manner
563 Controlled by YhdP. mBio 9:01321-18.
564 13. Aguiniga LM, Yaggie RE, Schaeffer AJ, Klumpp DJ. 2016. Lipopolysaccharide Domains
565 Modulate Urovirulence. Infection and Immunity 84:3131-3140.
566 14. Kalynych S, Morona R, Cygler M. 2014. Progress in understanding the assembly
567 process of bacterial O-antigen. FEMS Microbiol Rev 38:1048-65.

568 15. Gilbreath JJ, Colvocoresses Dodds J, Rick PD, Soloski MJ, Merrell DS, Metcalf ES.
569 2012. Enterobacterial common antigen mutants of *Salmonella enterica* serovar
570 Typhimurium establish a persistent infection and provide protection against subsequent
571 lethal challenge. Infect Immun 80:441-50.
572 16. Armbruster CE, Forsyth-DeOrnellas V, Johnson AO, Smith SN, Zhao L, Wu W, Mobley
573 HLT. 2017. Genome-wide transposon mutagenesis of *Proteus mirabilis*: Essential genes,
574 fitness factors for catheter-associated urinary tract infection, and the impact of
575 polymicrobial infection on fitness requirements. PLOS Pathogens 13:e1006434.
576 17. Holmes CL, Wilcox AE, Forsyth V, Smith SN, Moricz BS, Unverdorben LV, Mason S, Wu
577 W, Zhao L, Mobley HLT, Bachman MA. 2023. *Klebsiella pneumoniae* causes bacteremia
578 using factors that mediate tissue-specific fitness and resistance to oxidative stress.
579 PLOS Pathogens 19:e1011233.
580 18. Guérin F, Lallement C, Goudergues B, Isnard C, Sanguinetti M, Cacaci M, Torelli R,
581 Cattoir V, Giard J-C. 2020. Landscape of in vivo Fitness-Associated Genes of
582 *Enterobacter cloacae* Complex. Frontiers in Microbiology 11:1609.
583 19. Mannel D, Mayer H. 1978. Isolation and Chemical Characterization of the
584 Enterobacterial Common Antigen. European Journal of Biochemistry 86:361-370.
585 20. Goździewicz TK, Łukasiewicz J, Ługowski C. 2015. [The structure and significance of
586 enterobacterial common antigen (ECA)]. Postepy Hig Med Dosw (Online) 69:1003-12.
587 21. Kajimura J, Rahman A, Rick PD. 2005. Assembly of cyclic enterobacterial common
588 antigen in *Escherichia coli* K-12. J Bacteriol 187:6917-6927.
589 22. Barr K, Klena J, Rick P. 1999. The Modality of Enterobacterial Common Antigen
590 Polysaccharide Chain Lengths Is Regulated by o349 of the wec Gene Cluster of
591 *Escherichia coli K-12*. Journal of bacteriology 181:6564-8.

592 23. Färnbäck M, Eriksson L, Senchenkova Sy, Zych K, Knirel YA, Sidorczyk Z, Widmalm G.
593 2003. Crystal Structure of a Cyclic Enterobacterial Common Antigen. Angewandte
594 Chemie International Edition 42:2543-2546.
595 24. Erbel PJ, Seidel R, Macintosh SE, Gentile LN, Amor JC, Kahn RA, Prestegard JH,
596 McIntosh LP, Gardner KH. 2004. Cyclic enterobacterial common antigen: potential
597 contaminant of bacterially expressed protein preparations. J Biomol NMR 29:199-204.
598 25. Eade CR, Wallen TW, Gates CE, Oliverio CL, Scarbrough BA, Reid AJ, Jorgenson MA,
599 Young KD, Troutman JM. 2021. Making the enterobacterial common antigen glycan and
600 measuring its substrate sequestration. ACS Chem Biol 16:691-700.
601 26. Rick PD, Mayer H, Neumeyer BA, Wolski S, Bitter-Suermann D. 1985. Biosynthesis of
602 enterobacterial common antigen. J Bacteriol 162:494-503.
603 27. Barr K, Rick PD. 1987. Biosynthesis of enterobacterial common antigen in *Escherichia*
604 *coli.* In vitro synthesis of lipid-linked intermediates. J Biol Chem 262:7142-50.
605 28. Barr K, Nunes-Edwards P, Rick PD. 1989. In vitro synthesis of a lipid-linked trisaccharide
606 involved in synthesis of enterobacterial common antigen. J Bacteriol 171:1326-32.
607 29. Bertani B, Ruiz N. 2018. Function and Biogenesis of Lipopolysaccharides. EcoSal Plus
608 8:ESP-0001-2018.
609 30. Erbel PJ, Barr K, Gao N, Gerwig GJ, Rick PD, Gardner KH. 2003. Identification and
610 biosynthesis of cyclic enterobacterial common antigen in Escherichia coli. J Bacteriol
611 185:1995-2004.
612 31. Leo V, Tran E, Morona R. 2020. Polysaccharide co-polymerase WzzB/WzzE chimeras
613 reveal transmembrane 2 region of WzzB is important for interaction with WzyB. J
614 Bacteriol 203:00598-20.
615 32. Maczuga N, Tran ENH, Qin J, Morona R. 2022. Interdependence of *Shigella flexneri* O
616 Antigen and Enterobacterial Common Antigen Biosynthetic Pathways. J Bacteriol
617 204:e0054621.

618 33. Nath P, Morona R. 2015. Detection of Wzy/Wzz interaction in *Shigella flexneri*.
619 Microbiology (Reading) 161:1797-1805.
620 34. Weckener M, Woodward LS, Clarke BR, Liu H, Ward PN, Le Bas A, Bhella D, Whitfield
621 C, Naismith JH. 2023. The lipid linked oligosaccharide polymerase Wzy and its
622 regulating co-polymerase, Wzz, from enterobacterial common antigen biosynthesis form
623 a complex. Open Biology 13:220373.
624 35. Tocilj A, Munger C, Proteau A, Morona R, Purins L, Ajamian E, Wagner J, Papadopoulos
625 M, Van Den Bosch L, Rubinstein JL, Féthière J, Matte A, Cygler M. 2008. Bacterial
626 polysaccharide co-polymerases share a common framework for control of polymer
627 length. Nat Struct Mol Biol 15:130-8.
628 36. Kalynych S, Yao D, Magee J, Cygler M. 2012. Structural characterization of closely
629 related O-antigen lipopolysaccharide (LPS) chain length regulators. J Biol Chem
630 287:15696-705.
631 37. Collins RF, Kargas V, Clarke BR, Siebert CA, Clare DK, Bond PJ, Whitfield C, Ford RC.
632 2017. Full-length, Oligomeric Structure of Wzz Determined by Cryoelectron Microscopy
633 Reveals Insights into Membrane-Bound States. Structure 25:806-815.e3.
634 38. Kalynych S, Cherney M, Bostina M, Rouiller I, Cygler M. 2015. Quaternary structure of
635 WzzB and WzzE polysaccharide copolymerases. Protein Science 24:58-69.
636 39. Wiseman B, Nitharwal RG, Widmalm G, Högbom M. 2021. Structure of a full-length
637 bacterial polysaccharide co-polymerase. Nat Commun 12:369.
638 40. Wiseman B, Widmalm G, Högbom M. 2023. Alternating L4 loop architecture of the
639 bacterial polysaccharide co-polymerase WzzE. Commun Biol 6:802.
640 41. Morona R, Purins L, Tocilj A, Matte A, Cygler M. 2009. Sequence–structure relationships
641 in polysaccharide co-polymerase (PCP) proteins. Trends in biochemical sciences 34:78-
642 84.

643 42. Ascari A, Tran ENH, Eijkelkamp BA, Morona R. 2022. Identification of the *Shigella*
644 *flexneri* Wzy Domain Modulating Wzz(pHS-2) Interaction and Detection of the
645 Wzy/Wzz/Oag Complex. J Bacteriol 204:e00224-22.
646 43. Leo V, Teh MY, Tran ENH, Morona R. 2021. Identification of a Region in *Shigella flexneri*
647 WzyB Disrupting the Interaction with Wzz(pHS2). J Bacteriol 203:e0041321.
648 44. Whitfield C, Williams DM, Kelly SD. 2020. Lipopolysaccharide O-antigens-bacterial
649 glycans made to measure. J Biol Chem 295:10593-10609.
650 45. Ascari A, Tran ENH, Eijkelkamp BA, Morona R. 2022. Identification of the *Shigella*
651 *flexneri* Wzy Domain Modulating Wzz_{pHS-2} Interaction and Detection of the Wzy/Wzz/Oag
652 Complex. Journal of Bacteriology 0:e00224-22.
653 46. Huszczynski SM, Coumoundouros C, Pham P, Lam JS, Khursigara CM. 2019. Unique
654 regions of the polysaccharide copolymerase Wzz_2 from *Pseudomonas aeruginosa* are
655 essential for O-specific antigen chain length control. Journal of Bacteriology 201:e00165-
656 19.
657 47. Kalynych S, Ruan X, Valvano MA, Cygler M. 2011. Structure-guided investigation of
658 lipopolysaccharide O-antigen chain length regulators reveals regions critical for modal
659 length control. Journal of Bacteriology 193:3710-3721.
660 48. Rai AK, Carr JF, Bautista DE, Wang W, Mitchell AM. 2021. ElyC and cyclic
661 enterobacterial common antigen regulate synthesis of phosphoglyceride-linked
662 enterobacterial common antigen. mBio 12:e02846-21.
663 49. Mitchell AM, Srikumar T, Silhavy TJ. 2018. Cyclic enterobacterial common antigen
664 maintains the outer membrane permeability barrier of *Escherichia coli* in a manner
665 controlled by Yhdp. mBio 9:e01321-18.
666 50. Islam ST, Lam JS. 2014. Synthesis of bacterial polysaccharides via the Wzx/Wzy-
667 dependent pathway. Can J Microbiol 60:697-716.

668 51. Kintz EN, Goldberg JB. 2011. Site-directed mutagenesis reveals key residue for O
669 antigen chain length regulation and protein stability in *Pseudomonas aeruginosa* Wzz2.
670 J Biol Chem 286:44277-44284.
671 52. Papadopoulos M, Morona R. 2010. Mutagenesis and chemical cross-linking suggest that
672 Wzz dimer stability and oligomerization affect lipopolysaccharide O-antigen modal chain
673 length control. J Bacteriol 192:3385-93.
674 53. Morona R, van den Bosch L, Manning PA. 1995. Molecular, genetic, and topological
675 characterization of O-antigen chain length regulation in *Shigella flexneri*. J Bacteriol
676 177:1059-68.
677 54. Weckener M, Woodward LS, Clarke BR, Liu H, Ward PN, Le Bas A, Bhella D, Whitfield
678 C, Naismith JH. 2023. The lipid linked oligosaccharide polymerase Wzy and its
679 regulating co-polymerase, Wzz, from enterobacterial common antigen biosynthesis form
680 a complex. Open Biol 13:220373.
681 55. Kalynych S, Valvano MA, Cygler M. 2012. Polysaccharide co-polymerases: the
682 enigmatic conductors of the O-antigen assembly orchestra. Protein Eng Des Sel 25:797-
683 802.
684 56. Hong Y, Hu D, Verderosa Anthony D, Qin J, Totsika M, Reeves Peter R. 2023. Repeat-
685 Unit Elongations To Produce Bacterial Complex Long Polysaccharide Chains, an O-
686 Antigen Perspective. EcoSal Plus 11:eesp-0020-2022.
687 57. Wiseman B, Widmalm G, Högbom M. 2023. Alternating L4 loop architecture of the
688 bacterial polysaccharide co-polymerase WzzE. Communications Biology 6:802.
689 58. Lutz R, Bujard H. 1997. Independent and tight regulation of transcriptional units in
690 *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic
691 Acids Res 25:1203-10.

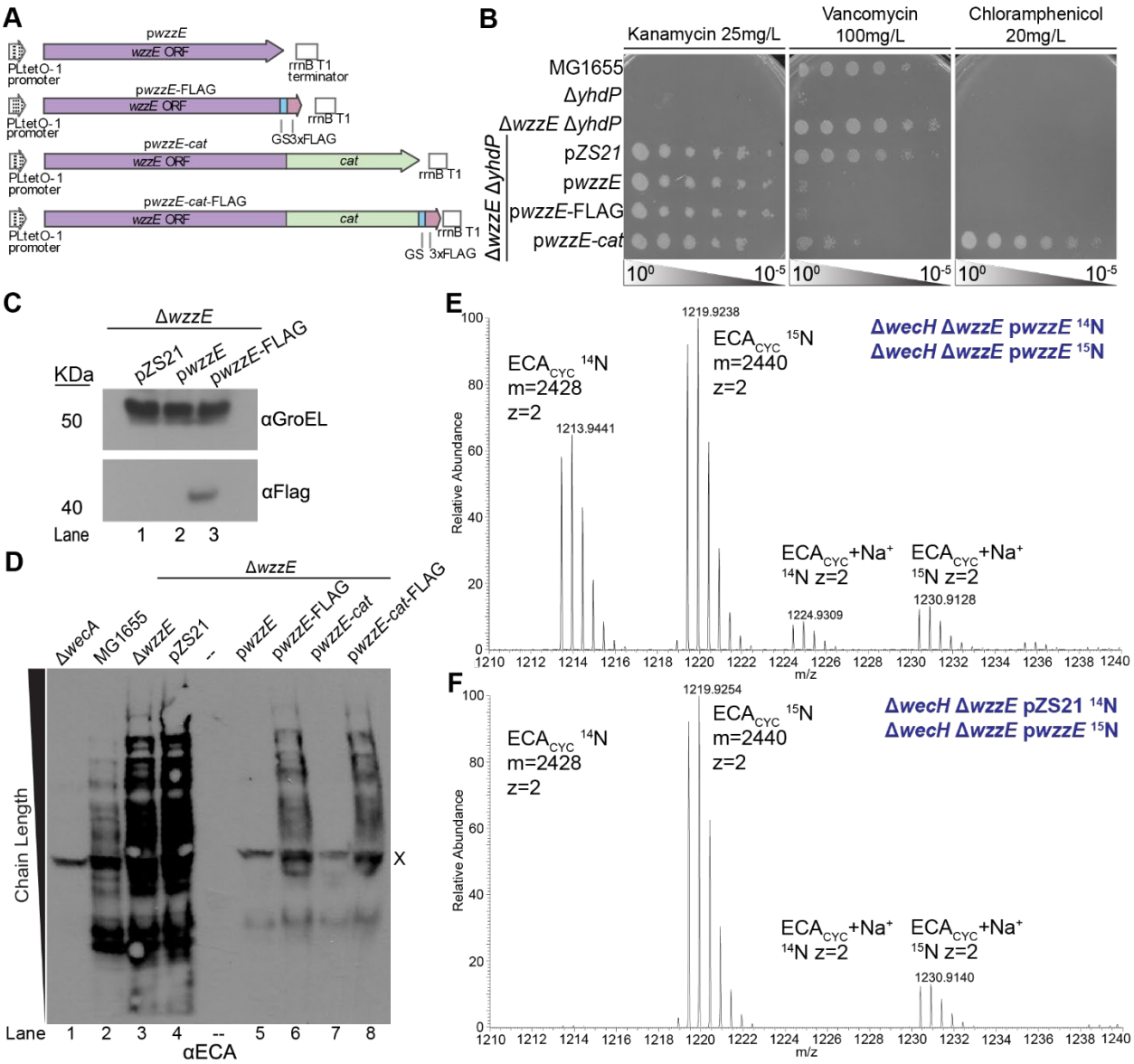
692 59. Mendez-Perez D, Gunasekaran S, Orler VJ, Pfleger BF. 2012. A translation-coupling
693 DNA cassette for monitoring protein translation in *Escherichia coli*. Metabolic
694 Engineering 14:298-305.
695 60. Sieber V. 2003. Selection for Soluble Proteins via Fusion with Chloramphenicol
696 Acetyltransferase, p 45-55. *In* Arnold FH, Georgiou G (ed), Directed Enzyme Evolution:
697 Screening and Selection Methods doi:10.1385/1-59259-396-8:45. Humana Press,
698 Totowa, NJ.
699 61. Andreeva AE, Borissova BE, Mironova R, Glykos NM, Kotsifaki D, Ivanov I, Krysteva M,
700 Kokkinidis M. 2000. Crystallization of type I chloramphenicol acetyltransferase: an
701 approach based on the concept of ionic strength reducers. Acta Crystallogr D Biol
702 Crystallogr 56:101-3.
703 62. Mitchell AM, Silhavy TJ. 2019. Envelope stress responses: balancing damage repair and
704 toxicity. Nature Reviews Microbiology 17:417-428.
705 63. Kajimura J, Rahman A, Hsu J, Evans MR, Gardner KH, Rick PD. 2006. O acetylation of
706 the enterobacterial common antigen polysaccharide is catalyzed by the product of the
707 *yiaH* Gene of *Escherichia coli* K-12. Journal of Bacteriology 188:7542-7550.
708 64. Daniels C, Morona R. 1999. Analysis of *Shigella flexneri* wzz (Rol) function by
709 mutagenesis and cross-linking: wzz is able to oligomerize. Mol Microbiol 34:181-94.
710 65. Teese MG, Langosch D. 2015. Role of GxxxG Motifs in Transmembrane Domain
711 Interactions. Biochemistry 54:5125-35.
712 66. Prakash A, Janosi L, Doxastakis M. 2011. GxxxG motifs, phenylalanine, and cholesterol
713 guide the self-association of transmembrane domains of ErbB2 receptors. Biophys J
714 101:1949-58.
715 67. Mueller BK, Subramaniam S, Senes A. 2014. A frequent, GxxxG-mediated,
716 transmembrane association motif is optimized for the formation of interhelical Cα–H
717 hydrogen bonds. Proceedings of the National Academy of Sciences 111:E888-E895.

718 68. Papadopoulos M, Tran ENH, Murray GL, Morona R. 2016. Conserved transmembrane
719 glycine residues in the *Shigella flexneri* polysaccharide co-polymerase protein WzzB
720 influence protein-protein interactions. Microbiology (Reading) 162:921-929.
721 69. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L,
722 Ballard AJ, Bambrick J, Bodenstein SW, Evans DA, Hung C-C, O'Neill M, Reiman D,
723 Tunyasuvunakool K, Wu Z, Žemgulytė A, Arvaniti E, Beattie C, Bertolli O, Bridgland A,
724 Cherepanov A, Congreve M, Cowen-Rivers AI, Cowie A, Figurnov M, Fuchs FB,
725 Gladman H, Jain R, Khan YA, Low CMR, Perlin K, Potapenko A, Savy P, Singh S,
726 Stecula A, Thillaisundaram A, Tong C, Yakneen S, Zhong ED, Zielinski M, Žídek A, Bapst
727 V, Kohli P, Jaderberg M, Hassabis D, Jumper JM. 2024. Accurate structure prediction of
728 biomolecular interactions with AlphaFold 3. Nature 630:493-500.
729 70. Parrish HL, Glassman CR, Keenen MM, Deshpande NR, Bronnimann MP, Kuhns MS.
730 2015. A Transmembrane Domain GGxxG Motif in CD4 Contributes to Its Lck-
731 Independent Function but Does Not Mediate CD4 Dimerization. PLOS ONE
732 10:e0132333.
733 71. Faingold O, Cohen T, Shai Y. 2012. A GxxxG-like Motif within HIV-1 Fusion Peptide Is
734 Critical to Its Immunosuppressant Activity, Structure, and Interaction with the
735 Transmembrane Domain of the T-cell Receptor*. Journal of Biological Chemistry
736 287:33503-33511.
737 72. Hong M, Payne SM. 1997. Effect of mutations in *Shigella flexneri* chromosomal and
738 plasmid-encoded lipopolysaccharide genes on invasion and serum resistance. Mol
739 Microbiol 24:779-91.
740 73. Tran ENH, Morona R. 2013. Residues located inside the *Escherichia coli* FepE protein
741 oligomer are essential for lipopolysaccharide O-antigen modal chain length regulation.
742 Microbiology (Reading) 159:701-714.

743 74. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M,
744 Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene
745 knockout mutants: the Keio collection. Mol Syst Biol 2:2006.0008.
746 75. Silhavy TJ, Berman ML, Enquist LW. 1984. Experiments with gene fusions. (No Title).
747 76. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in
748 *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of
749 Sciences 97:6640-6645.
750 77. Pyne ME, Moo-Young M, Chung DA, Chou CP. 2015. Coupling the CRISPR/Cas9
751 System with Lambda Red Recombineering Enables Simplified Chromosomal Gene
752 Replacement in *Escherichia coli*. Appl Environ Microbiol 81:5103-14.
753 78. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. 2013. RNA-guided editing of bacterial
754 genomes using CRISPR-Cas systems. Nature biotechnology 31:233-239.
755 79. Nichols BP, Shafiq O, Meiners V. 1998. Sequence analysis of Tn10 insertion sites in a
756 collection of *Escherichia coli* strains used for genetic mapping and strain construction. J
757 Bacteriol 180:6408-11.
758 80. Banerjee RV, Johnston NL, Sobeski JK, Datta P, Matthews RG. 1989. Cloning and
759 sequence analysis of the *Escherichia coli* metH gene encoding cobalamin-dependent
760 methionine synthase and isolation of a tryptic fragment containing the cobalamin-binding
761 domain. J Biol Chem 264:13888-95.
762 81. Guzman LM, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and
763 high-level expression by vectors containing the arabinose P_{BAD} promoter. J Bacteriol
764 177:4121-30.
765 82. Morris KN, Mitchell AM. 2023. Phosphatidylglycerol Is the Lipid Donor for Synthesis of
766 Phospholipid-Linked Enterobacterial Common Antigen. Journal of Bacteriology
767 205:e00403-22.
768

769 **Figures and Legends**

770



771

772 **Figure 1: Plasmid-based expression of *wzzE* partially complements a chromosomal *wzzE***

773 **deletion. (A)** Diagrams of constructs for expression of wild-type WzzE, WzzE with a C-terminal

774 glycine-serine linker and triple FLAG tag, WzzE translationally fused with chloramphenicol

775 acetyltransferase (Cat), and WzzE fused with Cat and triple FLAG are shown. The indicated

776 constructs were expressed from the low copy number pZS21 plasmid (58). **(B)** The vancomycin

777 sensitivity of a Δ*yhdP* strain can be suppressed by loss of WzzE (12). When the Δ*wzzE* Δ*yhdP*

778 strain carried the pZS21 empty plasmid, it remained resistance to vancomycin as assayed by
779 EOP but became sensitive when *wzzE, wzzE*-FLAG, or *wzzE-cat* was expressed, indicating all
780 three constructs complemented Δ*wzzE*. In addition, translational fusion of *cat* and *wzzE*
781 conferred chloramphenicol resistance, suggesting this construct could readout WzzE stability
782 and multimerization based on chloramphenicol resistance. **(C)** WzzE-FLAG was detectable by
783 immunoblotting. GroEL serves as a loading control. **(D)** ECA chain-length regulation was
784 assayed through immunoblotting using an ECA antibody. Plasmid-base expression of WzzE
785 restored chain-length regulation of linear ECA, although lower levels of ECA and shorter chain-
786 length were observed than in the wild-type strain for the p*wzzE* and p*wzzE* constructs, while the
787 p*wzzE*-FLAG and p*wzzE-cat*-FLAG constructs complemented better. The Δ*wecA* strain
788 produces no ECA and serves as negative control. X: nonspecific band. **(E-F)** The strain carrying
789 a plasmid with wild-type *wzzE* was grown with ^{15}N to label ECA and then combined with an
790 unlabeled strain carrying either wild-type *wzzE* (E) or the empty vector (F). ECA_{CYC} was purified
791 and then detected by LC-MS. MS traces are shown for the indicated samples. The mass,
792 charge, and identity of the ECA_{CYC}-associated peaks are shown. Expression of *wzzE* from
793 p*wzzE* restored ECA_{CYC} synthesis in the Δ*wzzE* strain. The strains contained Δ*wecH* to prevent
794 non-stochiometric acetylation of ECA. All images are representative of at least three
795 independent experiments.
796

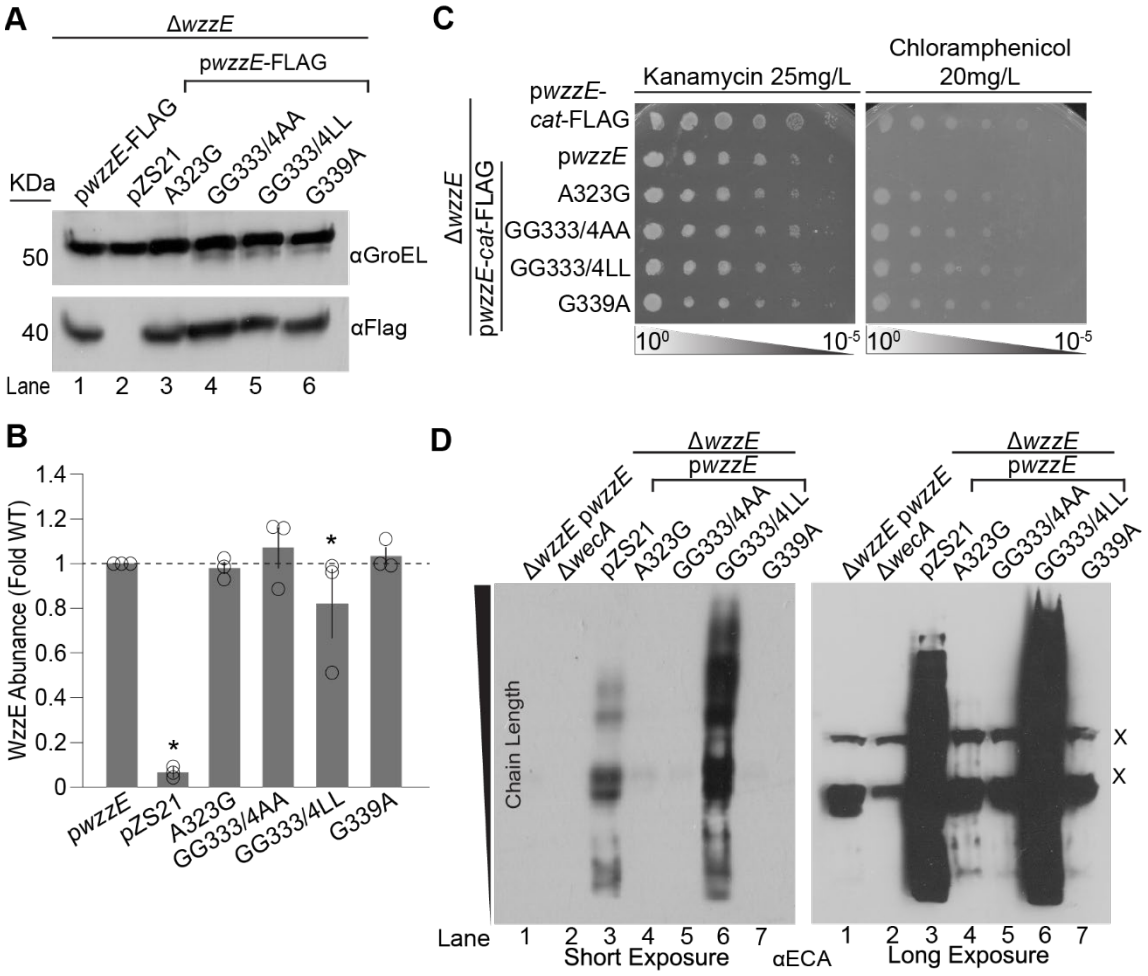


Figure 2: Alteration of WzzE TMH2 can prevent ECA chain length regulation without affecting WzzE stability. (A) The effect of the indicated *wzzE* site-directed mutations on WzzE-FLAG protein levels was assessed using immunoblotting. GroEL serves as a loading control. **(B)** Densitometry was used for quantification of the immunoblots in (A). Levels of WzzE-FLAG compared to GroEL are shown normalized to the wild-type *wzzE* complemented strain. All variants show similar protein levels to wild-type WzzE. Data are shown as the average of three biological replicates ± SEM and individual datapoints. * $p < 0.05$ by Mann-Whitney T-test. **(C)** Strains carrying the indicated plasmids were assayed for their chloramphenicol resistance by EOP as a proxy for WzzE stability and multimerization ability. All *wzzE* mutants retained full chloramphenicol resistance, suggesting WzzE stability and multimerization. **(D)** Chain-length regulation of the linear ECA forms by the indicated WzzE variants was assayed by

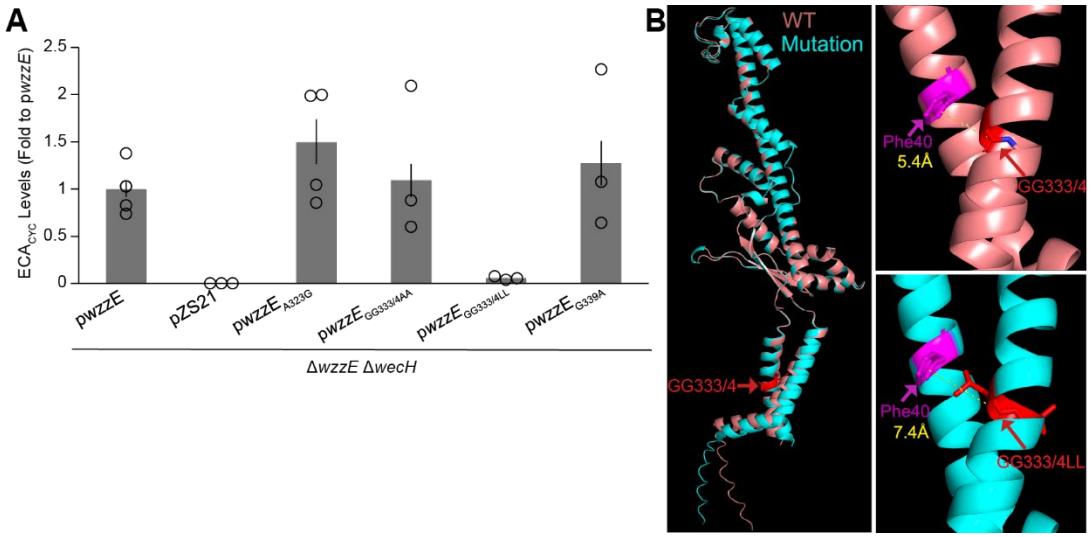
809 immunoblotting. Short and long exposures are shown. “X” indicates non-specific bands.

810 WzzE$^{GG333/4LL}$ showed loss of chain-length regulation while the other variants retained wild-type

811 function. The Δ*wecA* strain serves as negative control. All images are representative of at least

812 three independent experiments.

813



814

815 **Figure 3. Conserved glycines in WzzE TMH 2 are required for ECA$_{CYC}$ synthesis. (A)** The
816 strain carrying a plasmid with wild-type *wzzE* was grown with ^{15}N to label ECA and then
817 combined with an unlabeled strain carrying the indicated *wzzE* mutant or control. ECA$_{CYC}$ was
818 purified and the relative amount of unlabeled to labeled ECA$_{CYC}$ was determined by LC-MS.
819 WzzE$^{GG333/4LL}$ showed a loss of the ability to produce ECA$_{CYC}$. All plasmids were expressed in a
820 Δ*wzzE* Δ*wecH* strain background to prevent non-stochiometric acetylation of ECA. Data are
821 shown as the average of three biological replicates ± SEM and as individual datapoints. * $p <$
822 0.05 by Mann-Whitney T-test. **(B)** The predicted structure of a WzzE octamer with WzyE was
823 calculated using AlphaFold 3 (69). The structures of wild-type WzzE and WzzE$^{GG333/4LL}$
824 monomers were aligned with the mutated residues shown in red. The insets show the predicted
825 distance between the two WzzE TMH in the wild type WzzE and WzzE$^{GG333/4LL}$.

826

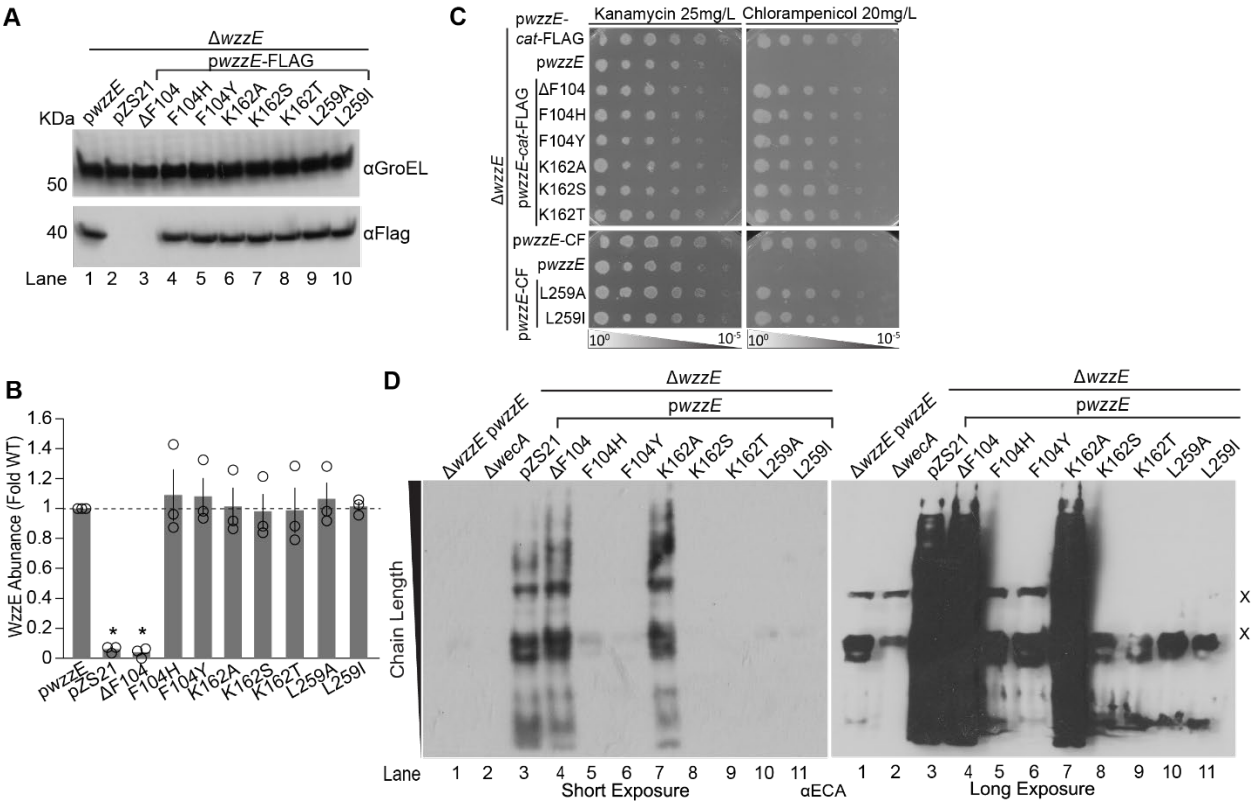


Figure 4: The periplasmic region of WzzE contributes to WzzE stability and chain length regulation. (A) Immunoblotting was used to determine protein levels for the indicated WzzE-FLAG variants. GroEL serves as a loading control. **(B)** WzzE-FLAG levels from (A) were quantified by densitometry. WzzE-FLAG levels compared to GroEL are shown normalized to the wild-type *wzzE* complemented strain. All variants are similar to wild type WzzE except for WzzE$^{\Delta F104}$ which shows significantly reduced protein levels. Data are shown as the average of three biological replicates ± SEM and individual datapoints. * $p < 0.05$ by Mann-Whitney T-test. **(C)** The indicated strains were assayed for WzzE-Cat-FLAG stability as measured by chloramphenicol resistance. All *wzzE* point mutations showed chloramphenicol resistance indicating stability and multimerization, with the exception of WzzE$^{\Delta F104}$ where the tags are cleaved (**Figure S3C**). **(D)** Linear ECA chain-length regulation by WzzE variants was assayed by immunoblotting with ECA antibody. Long and short exposures are shown. The periplasmic point mutants have variable effects on chain length regulation as well as ECA_{CYC} biogenesis.

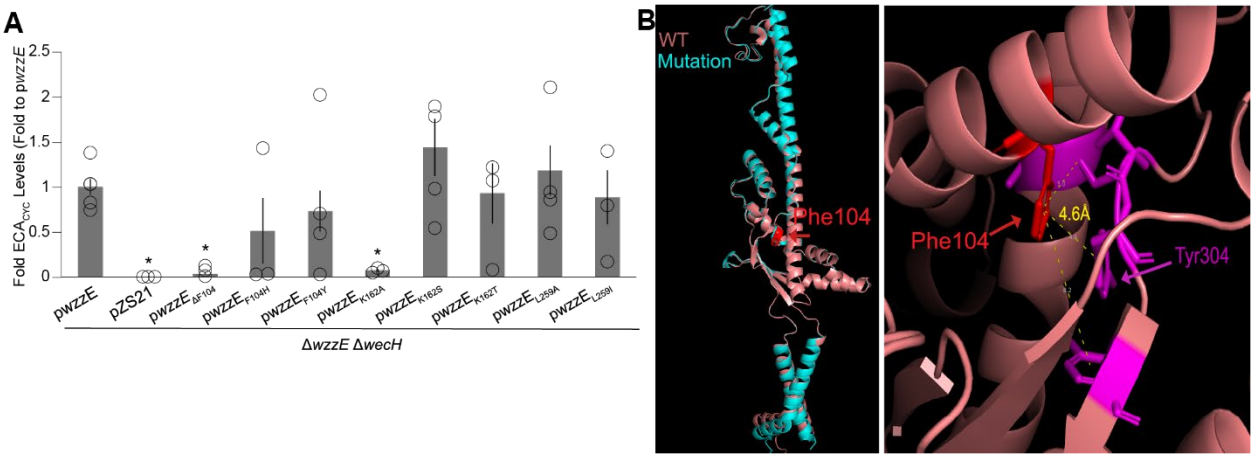
841 The WzzE$^{\Delta F104}$ and WzzE^{K162A} variants showed loss of ECA chain-length regulation while the

842 other variants had wild-type chain-length regulation. The Δ*wecA* strain serves as negative

843 control. All images are representative of at least three independent experiments.

844

845



846

847 **Figure 5: The periplasmic domain of WzzE contributes to ECA$_{CYC}$ synthesis. (A)** The strain
848 carrying wild-type *wzzE* was grown with ^{15}N to label ECA and then combined with an unlabeled
849 strain carrying the indicated *wzzE* mutant or control strain. After ECA$_{CYC}$ purification, the relative
850 amount of unlabeled to labeled ECA$_{CYC}$ was determined by LC-MS. WzzE$^{\Delta F104}$ and WzzE^{K162A}
851 had little to no ECA$_{CYC}$ production while WzzE^{F104H} showed an variable phenotype. All strains
852 included Δ*wecH* to prevent non-stochiometric acetylation of ECA. Data are shown as the
853 average of three biological replicates ± SEM and as individual datapoints. * $p < 0.05$ by Mann-
854 Whitney T-test. **(B)** Structures predicted as in **Figure 3** for wild-type WzzE and WzzE$^{F104\Delta}$ were
855 aligned. The mutated residue is shown in red on a WzzE monomer (B). The inset shows the
856 interaction between alpha helix 2 and an adjoining loop in wild-type WzzE.

857

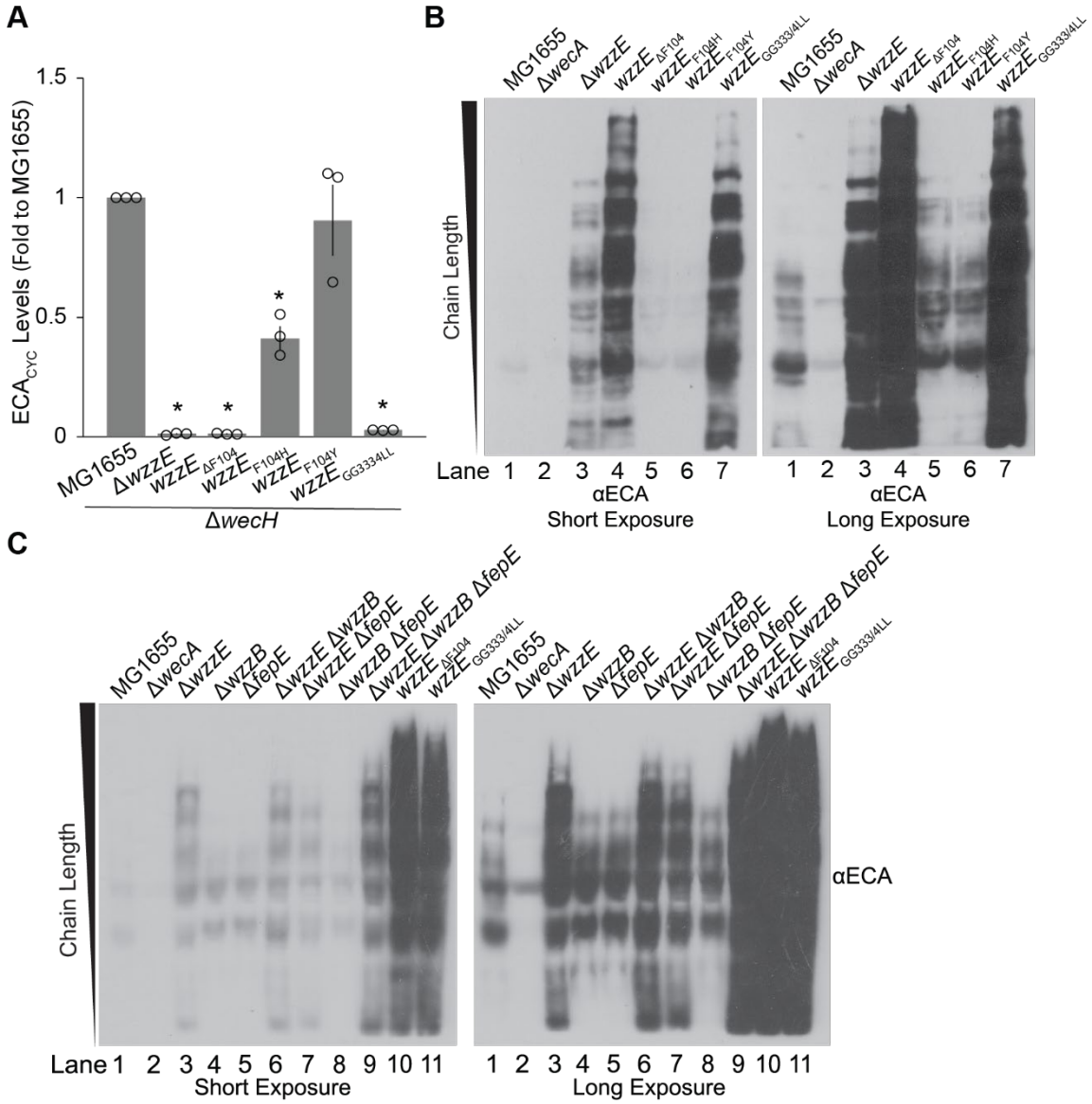


Figure 6: Levels of ECA$_{CYC}$ production and chain-length regulation are genetically separable. CRISPR-Cas9 was used to make the indicated mutations to *wzzE* in its native locus. All strains contain *metE*::Tn*10*. **(A)** The strain wild type for *wzzE* was grown with ^{15}N to label ECA and then combined with an unlabeled strain with the indicated *wzzE* mutant or control strain. After ECA$_{CYC}$ purification, the ratio of unlabeled to labeled ECA$_{CYC}$ was determined by LC-MS. All variants other than WzzE^{F104Y} produced significantly reduced ECA$_{CYC}$ levels. **(B)** Immunoblotting for ECA was conducted on the indicated chromosomal mutants. A short and long exposure are shown. The WzzE^{ΔF104} and WzzE$^{GG333/4LL}$ variants resulted in ECA chain

867 length dysregulation, while the other mutations more closely resembled wild-type WzzE chain
868 length regulation. The Δ*wecA* strain serves as negative control. **(C)** Immunoblotting was used to
869 assay ECA chain-length regulation in the indicated strains. Deletion of *wzzB* and/or *fepE* in
870 combination with Δ*wzzE* did not cause the increased levels of linear ECA observed with the
871 loss-of-function point mutants; however, the Δ*wzzE* Δ*wzzB* Δ*fepE* strain did have a small
872 increase in linear ECA levels compared to Δ*wzzE* alone.
873