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

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

[a] Department of Biology, Texas A&M University, College Station, Texas, USA (jfcarr@tamu.edu)

[b] Department of Biology, Texas A&M University, College Station, Texas, USA (jsrudolf@tamu.edu)

[c] Department of Biology, Texas A&M University, College Station, Texas, USA (dwarzecha03@tamu.edu)

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

[e] Department of Biology, Texas A&M University, College Station, Texas, USA (amitchell@bio.tamu.edu)

* 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

Keywords: gram negative envelope, enterobacterial common antigen, O antigen, Wzz, glycan biosynthesis, chain length regulation

Funding Statement: This work was supported by the National Institute of Allergy and Infectious Disease under award number R01-AI155915 (to A.M.M.).

Conflict Statement: The authors declare that they have no conflicts of interest.

Abstract

Enterobacterial common antigen (ECA) is a conserved glycan that supports outer membrane impermeability and intrinsic antibiotic resistance in Enterobacterales. ECA exists in outer membrane diacylglycerol-phosphate- and lipopolysaccharide-linked forms, and a cyclic periplasmic form (ECA_{CYC}). Intriguingly, ECA_{CYC} both affects the outer-membrane permeability barrier and functions in regulation of diacylglycerol-phosphate-linked ECA levels. While the length of linear ECA polymers generated by WzyE is regulated by the co-polymerase WzzE, WzzE is also required for ECA_{CYC} biogenesis and no ECA_{CYC} is synthesized in its absence. To uncover WzzE functions necessary for ECA_{CYC} biosynthesis, we generated plasmid-borne and chromosomal *wzzE* mutants in *Escherichia coli* K-12 and quantified their effects on linear chain-length regulation and on ECA_{CYC} synthesis. Mutations affecting linear ECA chain-length regulation in either transmembrane helix 2 or the periplasmic domain abolished ECA_{CYC} synthesis. Notably, two variants at residue F104 exhibited identical linear ECA chain-length regulation which was similar to wild type, but resulted in sharply different ECA_{CYC} production: chromosomal $WzzE_{F104Y}$ produced near wild-type levels of ECA_{CYC}, whereas $WzzE_{F104H}$ produced approximately two-fold less ECA_{CYC}. This difference in ECA_{CYC} production vs. linear chain-length regulation demonstrates ECA_{CYC} synthesis is genetically separable from this regulation. Thus, WzzE-mediated chain-length regulation is necessary for ECA_{CYC} biogenesis, but not sufficient.

46 **Introduction**

47 *Enterobacterales* is an order of gram-negative bacteria that includes numerous genera
48 with human pathogens (e.g., *Escherichia, Salmonella, Klebsiella, Yersinia*) (Doern, 2024,
49 Adeolu et al., 2016). These organisms cause a variety of diseases including respiratory, soft
50 tissue, and urinary tract infections, as well as endocarditis (Ramirez and Giron, 2024). To
51 survive in their enteric environment, members of this family take advantage of their outer
52 membrane (OM) which acts as a robust permeability barrier composed of an asymmetric bilayer
53 with phospholipids in the inner leaflet and predominantly lipopolysaccharides (LPS) in the outer
54 leaflet (Maher and Hassan, 2023, Silhavy et al., 2010). LPS molecules form an extensive
55 hydrophilic network of lateral interactions that excludes many large and hydrophobic
56 compounds—such as many antibiotics, bile salts, and detergents—from entering the cell
57 (Farfour et al., 2022, Aminov, 2010).

58 Beyond LPS, the outer membrane is decorated with other components critical for
59 fitness, including enterobacterial common antigen (ECA) (Bennett and Mitchell, 2025, Rai and
60 Mitchell, 2020). Identified in 1962 by Calvin M. Kunin et al. as a "common antigen" associated
61 with urinary tract infections, ECA is an invariant carbohydrate restricted to *Enterobacterales* and
62 absent only in endosymbionts with significantly reduced genomes (Rai and Mitchell, 2020,
63 Jenkins et al., 2017, Kunin, 1963, Kunin et al., 1962). ECA has been shown to play a role in the
64 pathogenesis in several members within the order of *Enterobacterales* (Aguiniga et al., 2016,
65 Kalynych et al., 2014, Gilbreath et al., 2012, Armbruster et al., 2017, Holmes et al., 2023,
66 Guérin et al., 2020). For example, *Salmonella enterica* serovar Typhimurium LT2 strains lacking
67 ECA display reduced virulence, heightened sensitivity to bile salts, and establish a low-level
68 persistent infection in mouse models (Kalynych et al., 2014, Gilbreath et al., 2012).

69 ECA is defined by recurring trisaccharide repeat units (RUs) of GlcNAc (*N*-acetyl-D-
70 glucosamine), ManNAcA (*N*-acetyl-D-mannosaminuronic acid), and Fuc4NAc (4-acetamido-4,6-
71 dideoxy-D-galactose) (Mannel and Mayer, 1978). There are three forms of ECA, (i) ECA_{LPS},

72 which replaces O-antigen (Oag) on the LPS core; (ii) ECA_{PG}, a cell-surface form linked to
73 diacylglycerol-phosphate; and (iii) ECA_{CYC}, a soluble, cyclic form found in the periplasm
74 (Goździewicz et al., 2015, Kunin, 1963, Mannel and Mayer, 1978, Kajimura et al., 2005). ECA_{PG}
75 and ECA_{LPS} vary in the number of RUs with a range of 1 to 14 and has a modal value of 5 to 7
76 at 37°C (Mitchell et al., 2018, Barr et al., 1999), whereas ECA_{CYC} has a defined number of
77 exactly 4 RUs in *Escherichia coli* K-12 (Kajimura et al., 2005). Structural analysis studies
78 including crystallography, nuclear magnetic resonance (NMR), and molecular dynamics indicate
79 that ECA_{CYC} adopts either a rhomboidal or square conformations (Färnbäck et al., 2003, Erbel
80 et al., 2004). We have previously found that ECA_{CYC} contributes to OM integrity in *E.coli* K-12
81 and mediates resistance to detergents and bile salts (Mitchell et al., 2018).
82 ECA biosynthesis genes reside mainly in the *wec* operon and employs a pathway
83 analogous to the Wzy/Wzz-dependent pathway of Oag synthesis (Eade et al., 2021, Rick et al.,
84 1985, Barr and Rick, 1987, Barr et al., 1989, Bertani and Ruiz, 2018, Barr et al., 1999, Erbel et
85 al., 2003). After a single undecaprenyl-phosphate-linked ECA RU is synthesized, WzxE flips the
86 intermediate to the outer leaflet of the inner membrane (Rai and Mitchell, 2020). WzyE and
87 WzzE form a complex where the translocated undecaprenyl-phosphate-linked ECA subunits are
88 polymerized by WzyE and the number of RU in the resulting chain is regulated by WzzE (Leo et
89 al., 2021b, Maczuga et al., 2022b, Nath and Morona, 2015, Weckener et al., 2023). WzzE is an
90 octameric inner-membrane protein with a large periplasmic domain (**Figure S1**) (Tocilj et al.,
91 2008, Kalynych et al., 2012b, Collins et al., 2017, Kalynych et al., 2015). A cryo-EM structure
92 from *Pectobacterium atrosepticum* suggest that WzyE resides within the central lumen formed
93 by WzzE's transmembrane helices (Weckener et al., 2023).
94 WzzE is a member of the class 1 polysaccharide co-polymerase (PCP1) family of chain
95 length regulators, which includes WzzB, FepE, Wzz among others (Wiseman et al., 2021,
96 Morona et al., 2009, Wiseman et al., 2023). Despite low sequence identity, PCP1 proteins share
97 a conserved architecture consisting of two transmembrane helices and a bell-shaped

98 periplasmic domain. (Tocilj et al., 2008, Kalynych et al., 2012b, Collins et al., 2017, Kalynych et
99 al., 2015). In *Shigella flexneri*, WzzB C-terminal transmembrane helix two (TMH2) is known to
100 interact with WzyB, the polymerase, and is necessary for Oag chain-length regulation (Ascari et
101 al., 2022, Leo et al., 2021a). WzzB and WzyB co-purify, and the TMH2 of WzzE can maintain
102 this interaction and contribute to Oag chain length regulation (Leo et al., 2021b). It has also
103 been shown that mutations disrupting Oag chain-length regulation are found throughout Wzz
104 and are not restricted to TMH2 (Whitfield et al., 2020).

105 WzzE plays a more complex role in ECA synthesis than other PCP1s do in Oag
106 production. Loss of WzzE generates linear ECA with increased heterogeneity—producing ECA
107 with both shorter and longer RUs—rather than the pronounced shortening typical of PCP1
108 deletions in Oag pathways (Leo et al., 2021b, Ascari et al., 2022, Huszczynski et al., 2019,
109 Kalynych et al., 2011). In contrast to linear ECA, strains lacking WzzE fail to produce ECA_{CYC},
110 indicating that WzzE is required for ECA_{CYC} synthesis despite a lack of known catalytic activity
111 (Rai and Mitchell, 2020, Mitchell et al., 2018, Kajimura et al., 2005). In addition, WzzE plays a
112 role in regulating levels of ECA_{PG} in *E. coli* K-12. ElyC regulates production of ECA_{PG} in such a
113 way that loss of ElyC causes large increases in ECA_{PG} production (Rai et al., 2021). This
114 regulation is dependent on WzzE and likely on ECA_{CYC}, and changes in ECA_{PG} levels with loss
115 of ElyC are much less in the absence of WzzE (Mitchell et al., 2018, Rai et al., 2021). Finally,
116 ECA takes on aberrant activity in the absence of the anterograde phospholipid transporter,
117 YhdP, causing increased sensitivity to vancomycin and the detergent SDS, and this phenotype
118 is reversed when *wzzE* is absent (Mitchell et al., 2018).

119 Several non-exclusive hypotheses for how PCPs regulate chain length have been
120 proposed (Islam and Lam, 2014, Leo et al., 2021b, Wiseman et al., 2021, Wiseman et al., 2023,
121 Kintz and Goldberg, 2011, Papadopoulos and Morona, 2010, Morona et al., 1995, Weckener et
122 al., 2023). (i) The inner or outer surface of the PCP may act as a molecular ruler for chain length
123 regulation (Kintz and Goldberg, 2011). Specifically, the PCP may bind the extending

124 carbohydrate, facilitating catalysis until the polymer length breaks the interaction (Kalynych et
125 al., 2012a). (ii) The PCP could serve as a scaffold on which multiple polymerase molecules
126 assemble (Tocilj et al., 2008). (iii) The elongating carbohydrate could extend in the lumen of the
127 PCP until it reaches the top of the periplasmic domain and is exposed (Islam and Lam, 2014).
128 (iv) The PCP could interact with the polymerase to sequester short glycan chains from
129 downstream ligases (i.e., WaaL), increasing the processivity of polymerization (Hong et al.,
130 2023). Finally, (v) the loops at the top of the periplasmic bell may move in an alternating up and
131 down arrangement which allows for the polysaccharide chain to extend using a "ratchet"-type
132 mechanism (Wiseman et al., 2023). However, the mechanism(s) used by WzzE for ECA_{CYC}
133 synthesis is unknown. Here, we used site-directed mutagenesis and biochemical analysis to
134 test whether WzzE requires only polymerase-interaction/chain-length regulation to produce
135 ECA_{CYC}, or if an additional function of WzzE contributes to ECA_{CYC} biogenesis. Our findings
136 demonstrate that polymerase interaction and chain-length regulation are necessary for ECA_{CYC}
137 biosynthesis. Moreover, we found that ECA_{CYC} synthesis genetically separated from linear ECA
138 regulation, suggesting an additional function of WzzE in ECA_{CYC} synthesis.
139

140 **Results**

141 **Plasmid-based *wzzE* allows analysis of *wzzE* mutant function.** In order to define the
142 functions of WzzE required for ECA_{CYC} synthesis in *E. coli* K-12, we examined the effects of
143 *wzzE* mutations analogous to known residues affecting PCP1 Oag regulation. We first
144 generated a set of *wzzE* constructs constitutively expressed by the *tet* promoter on the low copy
145 number pZS21 plasmid (Lutz and Bujard, 1997), which does not contain a Tet repressor (**Figure**
146 **1A**). These included wild-type *wzzE* (p*wzzE*) and C-terminal fusions bearing a glycine-serine
147 linker and triple FLAG tag (p*wzzE*-FLAG), a translation fusion to chloramphenicol
148 acetyltransferase (p*wzzE-cat*), and a p*wzzE-cat* triple FLAG tag fusion (p*wzzE-cat*-FLAG).
149 Based on the topology of WzzE, we anticipated that the C-terminal fusions would be exposed in

150 the cytoplasm, thereby allowing Cat to function. Thus, these constructs allowed us to potentially
151 identify defects in WzzE's folding or stability as measured by chloramphenicol resistance. This
152 approach was informed by a translational coupling methodology previously used a tool in *E. coli*
153 (Mendez-Perez et al., 2012, Sieber, 2003), in which an unstable or degraded protein construct
154 was shown to not allow chloramphenicol resistance. Additionally, we hypothesized that the
155 formation of the functional cat trimer (Andreeva et al., 2000) is more likely if WzzE is properly
156 multimerized, allowing Cat monomers to come into contact (with up to two Cat trimers
157 potentially formed per octamer), giving insight into potential disruptions in octamerization. This
158 assay is particularly useful because the addition of the bulky Cat (26KDa) protein likely imposes
159 stress on WzzE folding. However, the chloramphenicol resistance results remain suggestive
160 due to the potential limitations of this assay, such as the multimerization of Cat affecting
161 octamer formation (i.e., a loss of chloramphenicol resistance is more indicative of a loss of
162 stability and/or multimerization than chloramphenicol resistance is indicative of multimerization).
163 We next assessed whether these constructs complemented Δ*wzzE.* In the absence of
164 *yhdP*, ECA causes sensitivity to vancomycin in a *wzzE* dependent manner (Mitchell et al.,
165 2018). Therefore, we expressed the *wzzE* constructs in a Δ*wzzE* Δ*yhdP* strain and tested for
166 vancomycin sensitivity. Our results demonstrated each construct complemented Δ*wzzE* by
167 restoring vancomycin sensitivity (**Figure 1B**). Additionally, p*wzzE-cat* conferred chloramphenicol
168 resistance to the strain indicating that the Cat fusion was functional. To determine if the p*wzzE*-
169 FLAG construct could be detected, we immunoblotted for FLAG using GroEL as a loading
170 control (**Figure 1C**) and determined that the fusion protein could be clearly detected.
171 To determine whether linear ECA regulation was restored in the *wzzE* plasmid
172 constructs, we first performed immunoblots for linear ECA (**Figure 1D**) using a polyclonal ECA
173 antibody (Leo et al., 2021b). Our results revealed a range of ECA bands consistent with
174 previous publications (Mitchell et al., 2018, Rai et al., 2021, Morris and Mitchell, 2023), with
175 increasing molecular weight indicating molecules with more RUs. The distribution of ECA chain

176 lengths was wider in a Δ*wzzE* mutant (Lane 3) than in wild-type cells (Lane 2) with an increase
177 in longer and shorter ECA bands, indicating loss of chain-length regulation. We also observed
178 more linear ECA with Δ*wzzE* as we would expect from previous findings (Rai et al., 2021,
179 Mitchell et al., 2018). For simplicity, we will refer to these changes as "loss of chain-length
180 regulation". Expression of p*wzzE* (Lane 5) but not the empty vector (Lane 4) restored chain-
181 length regulation, albeit with a lower modal chain length than the wild-type strain and much less
182 linear ECA overall. These changes may suggest overactivity of WzzE due to multi-copy
183 expression. The same pattern was observed for the p*wzzE-cat* construct (Lane 7); however, the
184 p*wzzE*-FLAG and p*wzzE-cat*-FLAG constructs resulted in chain-length regulation closer to that
185 of wild type.

186 Finally, to test whether the plasmid-based *wzzE* constructs rescue ECA_{CYC} synthesis, we
187 expressed the p*wzzE* and empty vector plasmids in a Δ*wzzE* Δ*wecH* strain. WecH has been
188 shown to non-stoichiometrically acetylate ECA (Kajimura et al., 2006) and deleting *wecH* allows
189 ECA_{CYC} to be quantitated as a single mass species (Kajimura et al., 2006, Mitchell et al., 2018).
190 Unlike the two linear forms of ECA, ECA_{CYC} is uncharged and cannot be detected by
191 immunoblotting. Therefore, we grew cultures with either a normal (^{14}N) or heavy (^{15}N) nitrogen
192 source, combined labeled and unlabeled samples, purified ECA_{CYC}, and assayed ECA_{CYC} levels
193 by liquid chromatography-mass spectrometry (LC-MS). ECA_{CYC} has a defined size of 4 RU in *E.*
194 *coli* K-12 (Kajimura et al., 2005) and this molecule has an monoisotopic mass of 2428.9 Da in
195 the absence of acetylation. Labeling with heavy nitrogen causes a 12 Da increase in this weight
196 due to the presence of 12 nitrogen atoms in ECA_{CYC} (one in each sugar monomer) (Mitchell et
197 al., 2018) (**Figure S2AB**). For this type of assay, each peak consists of multiple minor peaks
198 due to the stochiometric incorporation of naturally occurring isotopes (e.g., ^{14}C) (see **Figure**
199 **1EF**, **S2C**). When labeled and unlabeled samples carrying p*wzzE* were combined, we detected
200 peaks at both the unlabeled and labeled m/z (**Figure 1E**). However, when the labeled p*wzzE*
201 sample was combined with an unlabeled sample carrying the empty vector, only the ^{15}N labeled

202 ECA_{CYC} peak was detected (**Figure 1F**). ECA_{CYC} production from the p*wzzE* and p*wzzE-cat*
203 constructs was very similar (**Figure S2C**). These data confirm plasmid-based expression of
204 *wzzE* can facilitate the synthesis of ECA_{CYC}, making this an appropriate system to investigate
205 the role of WzzE in ECA_{CYC} biosynthesis.

206 **Conserved glycines in TMH2 are necessary for WzzE function.** After validating our WzzE
207 functional assays, we next examined the contribution of TMH2 on WzzE function. TMH2 has
208 been shown to mediate interactions between WzzB and Wzy in O-antigen synthesis (Leo et al.,
209 2021b), and prior mutagenesis studies of WzzB identified a conserved GXXXG motif in the
210 TMH2 region of WzzB from *Shigella flexneri Y* that was important for function (Daniels and
211 Morona, 1999), likely due to effects on helix-helix packing (Teese and Langosch, 2015, Prakash
212 et al., 2011, Mueller et al., 2014). A corresponding double mutant in *Shigella flexneri Y*
213 (G305A/G311A WzzB; G306A/G312A in *E. coli*) disrupts Oag chain length regulation
214 (Papadopoulos et al., 2016); therefore, we hypothesized that the analogous mutations to
215 GXXXG motif located in WzzE would play a significant role in mediating the helix-helix packing
216 of the transmembrane region and its interaction with WzyE.
217 To test this, we generated TMH2 mutations (**Figure S1B**) in the p*wzzE* constructs
218 described in **Figure 1A**. We first sought to determine whether the levels of WzzE protein
219 produced by p*wzzE*-FLAG or p*wzzE-cat*-FLAG were altered by the mutations (**Figure 2AB,**
220 **Figure S3A**). Our results showed a small but significant decrease in protein abundance for our
221 $WzzE^{GG333/4LL}$ in the WzzE-FLAG fusion (**Figure 2AB**), whereas no decrease is apparent in the
222 WzzE-Cat-FLAG fusion (**Figure S3A**). As the levels of the $WzzE^{GG333/4LL}$ variant are 82% of wild
223 type on average and as high as 98% in individual replicates, the biological relevance of this
224 decrease is unclear. To determine whether the mutations might have caused alterations to
225 protein stability or multimerization, we utilized our p*wzzE-cat* construct to assay
226 chloramphenicol resistance conferred by the fusion protein. None of our mutations showed any

227 chloramphenicol sensitivity (**Figure 2C, Figure S3B**), suggesting that they are largely stable
228 and multimerized, despite the small decrease in abundance for WzzE$^{GG333/4LL}$. Next, we set out
229 to determine if the mutations affected WzzE's ability to regulate ECA chain length by performing
230 ECA immunoblots. Three of the TMH2 mutations, WzzE^{A323G}, WzzE$^{GG333/4AA}$, and WzzE^{G339A}
231 (Lanes 4,5, and 7) retained wild-type ECA chain length regulation (**Figure 2DE**). In contrast,
232 WzzE$^{GG333/4LL}$ (Lane 6) markedly disrupted chain-length regulation, resembling or exceeding the
233 defect observed in our empty-vector control (Lane 3). These results identify the TMH2 GXXXG
234 region as a key determinant of WzzE-mediated ECA chain-length regulation

235 **The TMH2 GXXXG motif affects ECA$_{CYC}$ synthesis.** Given the strong defect of WzzE$^{GG333/4LL}$
236 in regulating ECA chain length, we assessed whether the TMH2 mutations impair ECA$_{CYC}$
237 biogenesis. We utilized LC-MS to detect and quantitate ECA$_{CYC}$ in purified samples as in **Figure**
238 **1EF**. We expressed the TMH2 mutants or controls in a Δ*wzzE* Δ*wecH* strain background and
239 labeled the wild type p*wzzE* strain with ^{15}N, while the mutants were grown with ^{14}N. The
240 quantification demonstrated WzzE^{A323G}, WzzE$^{GG333/4AA}$, and WzzE^{G339A} did not significantly
241 change ECA$_{CYC}$ levels, whereas the WzzE$^{GG333/4LL}$ mutant produced significantly reduced levels
242 of ECA$_{CYC}$, similar to those of the empty vector control (**Figure 3A**). These data demonstrate
243 that the TMH2-dependent chain-length regulation is required for ECA$_{CYC}$ synthesis.
244 To visualize how these mutations might alter WzzE function, we modeled the
245 WzyE:WzzE(8) complex for wild-type WzzE and WzzE$^{GG333/4LL}$ using AlphaFold 3 (Abramson et
246 al., 2024) (**Figure 3B** & **S4**, **Table S1**). Alignment of monomers revealed a ~2 Å expansion
247 between TMH1 and TMH2 in WzzE$^{GG333/4LL}$ consistent with steric disruption of helix-helix
248 packing (**Figure 3B**). This structural shift provides a potential mechanistic basis for the loss of
249 both chain-length regulation and ECA$_{CYC}$ biosynthesis in our GXXXG mutant.

250 **The periplasmic domain of WzzE functions in chain length regulation.** After assessing the
251 importance of WzzE TMH2 for ECA$_{CYC}$ biosynthesis, we next sought to determine whether

252 chain-length regulating activities within the periplasmic domain of WzzE are also required for
253 ECA_{CYC} biosynthesis. Twenty-one periplasmic residues are conserved among WzzE, FepE, and
254 WzzB in *E. coli*, suggesting that they are important for PCP1 function (Tocilj et al., 2008). We
255 also hypothesized that non-conserved residues could contribute to WzzE-mediated chain-length
256 regulation or ECA_{CYC} synthesis. Therefore, we made mutations in a conserved cluster of
257 predominantly hydrophobic residues located in the core of the α/β domains—spanning strands
258 β1 and β4, helix α2 and the N terminus of helix α6 (see **Figure S1B**). Additionally, we targeted
259 non-conserved residues throughout WzzE in positions similar to those that conferred changes
260 to Oag chain length in other PCP1s (Tocilj et al., 2008).

261 We first characterized the periplasmic mutants (**Figure S1B**) by examining WzzE protein
262 levels using the p*wzzE*-FLAG construct. $WzzE^{\Delta F104}$ (an in-frame, single codon deletion) had
263 undetectable WzzE-FLAG protein levels, whereas all other mutants retained wild-type protein
264 accumulation (**Figure 4AB**). When we examined the levels of WzzE-Cat-FLAG for $WzzE^{\Delta F104}$,
265 we observed a faint band at the expected molecular weight of the fusion and a second, lower-
266 molecular-weight band, indicating cleavage of the C-terminal tag (**Figure S3C**). Given the
267 cleavage of the tags, no conclusion could be drawn for the stability of $WzzE^{\Delta F104}$. All other
268 protein levels were similar to those of wild type.

269 We next evaluated whether the *wzzE* mutations might affect protein stability and/or
270 multimerization by measuring chloramphenicol resistance conferred by the p*wzzE*-cat fusion.
271 $WzzE^{K162A}$, $WzzE^{L259A}$, and $WzzE^{F104Y}$ caused partial or full chloramphenicol sensitivity in the
272 p*wzzE*-Cat fusion (**Figure S3D**) but showed full chloramphenicol resistance in the p*wzzE*-Cat-
273 FLAG fusion, which better complements linear chain length regulation (**Figure 4C**). These
274 differences suggest these variants may exhibit mild instability or reduced multimerization.
275 All other mutants conferred chloramphenicol resistance similar to wild type in both fusions
276 (**Figure 4C, S3D**). $WzzE^{\Delta F104}$ showed wild-type levels of chloramphenicol resistance, likely
277 reflecting cleavage of the C-terminal tag. We also tested an additional 18 periplasmic-domain

278 mutations that caused chloramphenicol sensitivity in the WzzE-Cat background, emphasizing
279 the importance of this domain for WzzE structural stability and multimerization (**Figure S5**).
280 After characterizing the effect of the periplasmic mutations on WzzE levels and stability,
281 we next examined how these periplasmic mutations affect ECA chain-length regulation (**Figure**
282 **4DE**). WzzE$^{\Delta F104}$ and WzzE^{K162A} (lanes 4 and 7) disrupted chain-length regulation when
283 expressed from p*wzzE* (**Figure 4D**). However, when expressed from p*wzzE*-cat-Flag,
284 WzzE$^{\Delta F104}$, exhibited partial impairment of chain-length regulation, while WzzE^{K162A} showed wild-
285 type chain-length regulation (**Figure 4E**). These differences parallel the chloramphenicol
286 resistance phenotypes observed for the corresponding fusions (**Figure 4C**, **S3D**). In contrast,
287 WzzE^{F104Y}, WzzE^{F104H}, WzzE^{K162S}, WzzE^{K162T}, WzzE^{L259A}, and WzzE^{L259I} (lanes 5, 6, 8-11)
288 demonstrated wild-type chain length regulation in both the p*wzzE* and p*wzzE-cat-FLAG*
289 constructs. Together, these mutants provide a useful range of phenotypes for dissecting the
290 requirements for ECA$_{CYC}$ synthesis.

291 **Periplasmic residues of WzzE are important for ECA$_{CYC}$ biogenesis.** We next assessed the
292 impact of our periplasmic mutants on ECA$_{CYC}$ biosynthesis by quantifying ECA$_{CYC}$ levels using
293 LC-MS. Both the WzzE$^{\Delta F104}$ and WzzE^{K162A} mutants exhibited impaired linear chain-length
294 regulation and produced little to no ECA$_{CYC}$ (**Figure 5A**). This loss of ECA$_{CYC}$ production could
295 results from either (i) insufficient protein accumulation and/or stability or (ii) disruption of chain-
296 length-regulating activities in these mutants. In contrast, WzzE^{F104Y}, WzzE^{K162S}, WzzE^{K162T},
297 WzzE^{L259A}, and WzzE^{L259I} produced ECA$_{CYC}$ at levels comparable to wild type. WzzE^{F104H}
298 showed highly variable ECA$_{CYC}$ levels despite maintaining wild type chain-length regulation and
299 protein levels, suggesting that this mutation may have a specific ECA$_{CYC}$ synthesis effect. These
300 data highlight the importance of WzzE's periplasmic domain on ECA$_{CYC}$ synthesis, likely through
301 its structural role, its chain-length-regulating activity, and possibly a function specific to ECA$_{CYC}$
302 synthesis. Given the differing effects of the F104 mutations on ECA$_{CYC}$ synthesis and linear ECA

303 chain-length regulation, we examined their potential structural consequences using AlphaFold 3
304 (Abramson et al., 2024) and aligned wild type and WzzE$^{\Delta F104}$ complexes (**Figure 5B** & **S4**,
305 **Table S1**). These models predict that F104 forms contacts with a neighboring loop, and
306 alterations in these interactions may underlie the functional differences observed for
307 substitutions at this residue.

308 **ECA$_{CYC}$ biogenesis is genetically separable from chain-length regulation.** Our plasmid-
309 based expression analyses of *wzzE* identified several point mutations with distinct effects on
310 ECA$_{CYC}$ production. Notably, substitution of the TMH2 residues (WzzE$^{GG333/4LL}$) abolished chain-
311 length regulation and ECA$_{CYC}$ synthesis, while maintaining protein stability and near-wild-type
312 protein accumulation. Additionally, several mutations to residue F104 also produced variable
313 effects on WzzE accumulation, chain-length regulation, and ECA$_{CYC}$ production. However,
314 plasmid-based expression of wild-type *wzzE* yielded lower levels of linear ECA and reduced
315 modal chain-length compared expression of *wzzE* from its native locus (**Figure 1D**), a
316 difference likely contributing to some of the phenotypes observed in the plasmid system. We
317 also noted substantial variability in ECA$_{CYC}$ levels among periplasmic mutants, which we
318 attributed to non-native gene expression. To circumvent these limitations, we used CRISPR-
319 Cas9 to introduce select point mutations into the native *wzzE* locus and assayed their effect on
320 linear ECA and ECA$_{CYC}$.

321 LC-MS analysis of purified ECA$_{CYC}$ revealed that chromosomal WzzE$^{\Delta F104}$ and
322 WzzE$^{GG333/4LL}$ mutants produced little to no ECA$_{CYC}$, similar to what is observed with Δ*wzzE*
323 (**Figure 6A**). In contrast, WzzE^{F104H} produced measurable levels of ECA$_{CYC}$ that were about 2-
324 fold lower than the wild type, whereas WzzE^{F104Y} showed no significant change in ECA$_{CYC}$ levels
325 relative to the wild type. We then assayed linear ECA chain length regulation by immunoblot
326 (**Figure 6B**). Consistent with the plasmid-based system, WzzE$^{\Delta F104}$ and WzzE$^{GG333/4LL}$ failed to
327 regulate linear ECA. The magnitude of chain-length dysregulation in these mutants exceeded

328 that of Δ*wzzE*, potentially suggesting an aberrant interaction of non-functional WzzE and ECA
329 synthesis that differs from the absence of WzzE. WzzE^{F104H} and WzzE^{F104Y} each displayed
330 intact and nearly wild-type chain-length regulation (**Figure 6B**), with a possible slight increase in
331 modal chain length that was not present in all replicates. However, the linear ECA regulation of
332 WzzE^{F104H} and WzzE^{F104Y} was consistently identical between these two mutants.
333 In our ECA immunoblot protocol, we take steps to minimize ECA degradation (see Experimental
334 Procedures), as ECA can be delipidated by mild acid hydrolysis (BASU et al., 1987). One of
335 these steps is to avoid long incubations, such as those used to proteinase K (ProK) treat LPS
336 samples (May and Silhavy, 2018). However, protein in our samples cause ECA to spread over a
337 larger area of the gel and potentially could mask small changes in chain length regulation. To
338 improve detection of small changes in chain length while minimizing ECA degradation, we
339 pretreated our immunoblot samples with various concentrations of ProK and incubation times at
340 56 °C found in the literature (Leo et al., 2020, Jiang et al., 2020, Maczuga et al., 2022a, Liu et
341 al., 2020, Peters et al., 1985), as well as treatments optimized based on our initial results. The
342 resulting blots and comparative signal quality are shown in (**Figure S5**). We selected a one-
343 hour treatment with a high concentration of ProK to move forward with as this treatment resulted
344 in loss of protein signal while minimizing ECA degradation. After treatment with ProK in these
345 optimized conditions, both WzzE^{F104H} and WzzE^{F104Y} showed ECA chain length that was
346 extremely similar to wild type (**Figure 6C**). Overall, the identical regulation of linear ECA chain-
347 length by WzzE^{F104H} and WzzE^{F104Y} but differing effects on ECA$_{CYC}$ synthesis demonstrate that
348 levels of ECA$_{CYC}$ production are genetically separable from linear chain-length regulation.
349 We were surprised that the dysregulation of ECA chain length, and even more the levels
350 of ECA, in the WzzE$^{\Delta F104}$ and WzzE$^{GG333/4LL}$ mutants was still greater than the Δ*wzzE* control,
351 and so we wondered whether, in a Δ*wzzE* strain, the remaining PCP1s could partially regulate
352 ECA chain length through interaction with WzyE. In this model, nonfunctional WzzE mutants
353 might sequester WzyE from the other PCP1s, exacerbating dysregulation. Therefore, we

354 assayed whether we could cause further disruption of ECA chain-length regulation in a Δ*wzzE*
355 strain by deleting *wzzB* and/or *fepE*. Deletion of Δ*wzzB*, Δ*fepE*, and Δ*wzzB* Δ*fepE*, or both had
356 no detectable effect on chain regulation in a Δ*wzzE* strain (**Figure 6D**), though the triple mutant
357 showed a modest increase in linear ECA levels. Taken together, these findings demonstrate that
358 mutations in *wzzE* that disrupt linear ECA chain-length regulation can also abolish ECA_{CYC}
359 synthesis, and that normal ECA_{CYC} production is genetically separable from regulation of linear
360 ECA chain-length.
361
362 **Discussion**
363 In this study, we constructed and characterized a set of *wzzE* site-directed mutants to
364 define which functions of WzzE are required for ECA_{CYC} biogenesis. Although previous
365 mutagenesis studies have identified regions of *Shigella flexneri* WzzB important for activity
366 (Nath and Morona, 2015, Ascari et al., 2022), no prior work has examined how WzzE
367 contributes to the production of cyclic ECA. Here, we show that substitutions at F104 within the
368 periplasmic domain and to the GXXXG motif in TMH2 can compromise both ECA chain-length
369 regulation and ECA_{CYC} synthesis. Moreover, our data demonstrate that regulation of linear ECA
370 chain length and the level of ECA_{CYC} production are genetically separable, as we have identified
371 two mutations at F104 with divergent effects on ECA_{CYC} levels, but produce identical regulation
372 of linear ECA chain length. These data show regulation of chain length by WzzE is necessary
373 but not sufficient for production of normal levels of ECA_{CYC}.
374 Despite low sequence conservation among PCP1s, their overall structures are highly
375 similar, consisting of two transmembrane domains and an α-helix/β-sheet domain within the
376 periplasmic space that extends ~100 Å above the membrane (Tocilj et al., 2008, Kalynych et al.,
377 2012b, Collins et al., 2017, Kalynych et al., 2015). PCP1s also harbor a conserved GXXXG
378 motif that has been shown to contribute to chain-length regulation. In WzzB, a G305A/G311A
379 double mutant in *Shigella flexneri* WzzB that encompassed and extended outside of this motif

380 resulted in short-chain O-antigen (Oag) (Daniels and Morona, 1999, Papadopoulos et al., 2016),
381 whereas single, double, and triple glycine to alanine mutations restricted to the GXXXG motif
382 did not affect chain-length (Collins et al., 2017, Daniels and Morona, 1999). Motivated by these
383 findings, we focused on residues within and around the GXXXG motif, characterized their effect
384 on chain-length regulation and ECA_{CYC} biogenesis, and identified two adjacent, conserved
385 glycine residues in the GXXXG motif of WzzE—analogous to G305/G306 in WzzB—as critical
386 for ECA chain-length regulation and ECA_{CYC} biosynthesis. Interestingly Ala substitutions at
387 these positions had no detectable effect on chain-length regulation or ECA_{CYC} production,
388 whereas introduction of adjacent Leu residues led to pronounced dysregulation of ECA chain-
389 length and significantly reduced levels of ECA_{CYC}.

390 GXXXG motifs often mediate dimerization and multimerization of transmembrane
391 domains (Parrish et al., 2015, Faingold et al., 2012), suggesting that the motif may facilitate
392 positioning of TMH1 and TMH2 of WzzE. In support of this idea, our structural modeling
393 suggests increased distance between TMH1 and TMH2 in the $WzzE^{GG333/4LL}$ mutant (**Figure 3D**,
394 **S6A**). Alignments of WzzE monomers predicted to closely interact with WzyE, shows that WzyE
395 rotates in the plane perpendicular to the membrane in WzyE-$WzzE^{GG333/4LL}$ model compared to
396 the wild type (**Figure S6BC**). The alignment also shows the distance in the structure between
397 the TMH of opposite WzzE monomers was 2.0 to 2.7 Å smaller in the $WzzE^{GG333/4LL}$ model than
398 in the wild-type model. In compensation, WzyE takes on a more condensed structure (1.4 Å
399 narrower) in the $WzzE^{GG333/4LL}$ model. As this type of rotation and structure change could impair
400 WzyE function, this prediction provides a possible hypothesis that $WzzE^{GG3334/LL}$ has impaired
401 interaction with WzyE. Alternatively, disruption of TMH packing may interfere with an intrinsic
402 and lesser defined activity of WzzE.

403 In contrast to mutations in TMH2, many substitutions introduced into the periplasmic
404 domain rendered WzzE unstable, as evidenced by loss of chloramphenicol resistance conferred
405 by the WzzE-Cat fusion. This observation underscores the importance of the periplasmic

406 domain for WzzEs structural integrity and/or multimerization. Prior work on other PCP1s
407 suggests that multiple surface-exposed residues within the periplasmic domain affect chain-
408 length regulation, producing either shorter or longer Oag chains (Kalynych et al., 2012b,
409 Papadopoulos and Morona, 2010, Daniels and Morona, 1999, Hong and Payne, 1997, Tran and
410 Morona, 2013). More recently, increasing evidence suggests that the inner surface of the WzzE
411 octamer may bind to the elongating carbohydrate to facilitate growth of the chain (Wiseman et
412 al., 2023, Hong et al., 2023, Weckener et al., 2023).

413 WzzE's internal cavity possesses two highly concentrated, negatively charged bands,
414 separated by a flexible neutrally charged loop that may bind the growing carbohydrate chain
415 (Wiseman et al., 2023). A loop at the apex of the periplasmic domain has been suggested to
416 "ratchet" up and down to facilitate carbohydrate movement over this binding surface (Wiseman
417 et al., 2023). F104 is buried near the base of the periplasmic "bell" and may affect this binding
418 surface indirectly. In contrast, K162 is positioned adjacent to the lower negatively charged band,
419 and L259 residues within the proposed ratchet loop. Mutations to L259 did not affect chain-
420 length regulation; however, loss of charge at K162 did prevent chain-length regulation and
421 ECA_{CYC} synthesis, perhaps supporting this model of ECA chain-length elongation. As to how the
422 *wzzE*$_{K162A}$ and *wzzE* $_{\Delta F104}$ mutations lead to loss of ECA_{CYC} production, we speculate that the
423 chain-length regulating activity of WzzE—related to the carbohydrate binding and ratcheting
424 mechanism—is necessary for ECA_{CYC} synthesis.

425 In contrast, the *wzzE*$_{F104H}$ mutation reduces ECA_{CYC} levels while maintaining
426 comparatively wild-type chain-length regulation. Moreover, its chain-length regulation is
427 indistinguishable from *wzzE*$_{F104Y}$, which produces wild-type ECA_{CYC} levels. These results
428 demonstrate that chain-length regulation alone is insufficient for normal ECA_{CYC} synthesis and
429 suggest that WzzE plays an additional role specific to ECA_{CYC} synthesis beyond that of chain-
430 length regulation. One possible mechanism would be that WzzE bends ECA molecules of the
431 correct size into an orientation that can be spontaneously cyclized. Another possibility is that

432 WzzE allows the binding of another protein (i.e., a cyclase) that can cyclize ECA chains of the
433 appropriate length. Finally, it is possible that WzzE has specific features that promote the
434 formation of ECA chains with 4 RUs that can later be cyclized to form ECA_{CYC}. Further, genetic
435 and biochemical experiments will be necessary to distinguish between these possibilities and to
436 solidify an additional role for WzzE in ECA_{CYC} biogenesis.

437 Unexpectedly, we consistently observed that $WzzE^{\Delta F104}$ and $WzzE^{GG333/4LL}$ mutants had
438 greater linear ECA levels than what was observed in our *wzzE* deletion. We confirmed that this
439 phenotype does not arise from compensation by other PCP1s in *E.coli*, although we did notice a
440 modest increase in linear ECA in strains lacking all other PCP1 genes. Therefore, why $WzzE^{\Delta F104}$
441 and $WzzE^{GG333/4LL}$ have elevated linear ECA levels remains unclear. One explanation is that
442 inactive WzzE octamers partially sequester growing ECA chains, increasing variability in chain
443 length. Regardless, we believe that the phenotypic distinction between Δ*wzzE* strain and the
444 mutants strengthens the conclusion that these mutants produce inactive proteins which
445 accumulate but are incapable of regulating chain-length or supporting ECA_{CYC} biosynthesis.

446 Overall, our analysis interrogates only a small fraction of WzzE residues, and additional
447 mutations that selectively impair ECA_{CYC} biosynthesis likely exist. Identification of such mutants
448 will clarify the molecular mechanism of ECA_{CYC} biosynthesis. Nonetheless, our findings reveal
449 fundamental requirements for ECA_{CYC} and chain-length regulation. We propose coordinated
450 activity between WzyE and WzzE, together with periplasmic charge elements within WzzE that
451 facilitate chain elongation, is required for ECA_{CYC} biogenesis. Importantly, we demonstrate that
452 chain-length regulation can be genetically uncoupled from ECA_{CYC} synthesis, indicating that
453 WzzE possesses a second function beyond chain-length regulation that is important for ECA_{CYC}
454 biosynthesis. These insights help to define key features of WzzE that enable synthesis of a
455 cyclic carbohydrate.

456

457 **Experimental Procedures**

458 *Strains and growth conditions*. The strains used in this study are listed in **Table S2**. Cultures
459 were inoculated and grown at 37 °C in LB Lennox unless otherwise noted, and, when
460 necessary, were supplemented with 20 mg/L chloramphenicol, 25 mg L^{-1} kanamycin, 10 mg L^{-1}
461 tetracycline, or a combination thereof. Insertion-deletion alleles containing kanamycin resistance
462 cassettes from the Keio collection (Baba et al., 2006) were moved to isogenic backgrounds
463 using P1vir transduction (Silhavy et al., 1984). Clean deletions were constructed using the Flp
464 recombinase-FRT system as described (Datsenko and Wanner, 2000). Chromosomal
465 mutagenesis in *wzzE* was performed using a CRISPR-Cas9 system harboring a
466 pCRISPR::protospacer plasmid in W3110 containing a heat shock-inducible λcI857 lysogen, as
467 previously described (Pyne et al., 2015, Jiang et al., 2013). The pCas9 (Addgene #42876) and
468 pCRISPR (Addgene #42975) plasmids were a kind gift from Professor Luciano Marraffini
469 (Rockefeller University). Oligonucleotides used to generate the protospacer and repair
470 templates used for mutagenesis are listed in **Table S3**. Prior to chromosomal mutagenesis,
471 *metE*::Tn*10* was linked to *wzzE* (Nichols et al., 1998), allowing the mutations to be transferred
472 into MG1655 via P1vir transduction. Strains carrying *metE*::Tn*10* were supplemented with 1μm
473 cobalamin (vitamin B_{12}) when grown in minimal media (Banerjee et al., 1989).

474 *Plasmids, cloning, and site directed mutagenesis***.** All primers used in this study are listed in
475 **Table S4**. To subclone the p*wzzE* expression plasmid, chromosomal *wzzE* with an ochre stop
476 codon was PCR-amplified using primers pZS21-wzzE_fwd and wzzE-pZS21_rev. The low copy
477 number plasmid pZS21 (Lutz and Bujard, 1997) was amplified using primers pZS21_fwd and
478 pZS21_rev. The plasmid was assembled using HiFi Assembly Master Mix (New England
479 Biolabs) according to the manufacturer's instructions. To construct p*wzzE-cat*, *wzzE* lacking a
480 stop codon was amplified using primers pZS21-wzzE_fwd and wzzE-cat_rev. The
481 chloramphenicol acetyltransferase gene (*cat*) was amplified from pBAD33 (Guzman et al., 1995)
482 using primers wzzE-cat_fwd and cat-pZS21_rev. These fragments were assembled into pZS21,

483 amplified as described above, using HiFi Assembly. To construct p*wzzE*-FLAG, p*wzzE* was
484 amplified using primers p*wzzE*_x3F_Fwd and p*wzzE*_x3F_Rvr, which contained sequence
485 homology to a glycine-serine-linked FLAG tag that incorporating a stop codon. The glycine-
486 serine-linked FLAG tag was supplied as a gene-block fragment (**Table S3**) and assembled into
487 the p*wzzE* plasmid via HiFi Assembly. To generate the p*wzzE-cat*-FLAG plasmid, *cat* was
488 amplified using primers WzzE:Cmr Fp and WzzE:Cmr Rp and the p*wzzE*-FLAG plasmid
489 backbone was amplified using Cmr: GS-x3F_FP and wzzE-cat_rev. The fragments were
490 assembled using HiFi Assembly.

491 Site-direct mutations were generated using Q5 PCR mutagenesis (New England
492 Biolabs). Briefly, non-overlapping primers containing the mutation of interest were used to
493 amplify the indicated pZS21-sanA plasmids (pZS21-*wzzE*, pZS21-*wzzE-cat*, and pZS21-*wzzE*-
494 FLAG). PCR products were digested with DpnI (New England Biolabs), phosphorylated using
495 T4 polynucleotide kinase (New England Biolabs) and self-ligated with T4 DNA ligase (New
496 England Biolabs). Ligations were transformed into Turbo cloning cells (New England Biolabs),
497 and the mutations were confirmed by sequencing (Plasmidsaurs). Primers used all site-directed
498 mutants are listed in **Table S4**. Modeling of the predicted mutant structures was performed
499 using AlphaFold 3 (Abramson et al., 2024) with eight monomers of WzzE and one of WzyE.
500 Relevant confidence scores are given in **Table S1** and **Figure S2**.

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

504 *Immunoblot Analysis***.** Immunoblots for WzzE-FLAG and ECA were conducted as previously
505 described with minor modification (Mitchell et al., 2018, Morris and Mitchell, 2023). For analysis
506 of WzzE-FLAG protein levels, samples were normalized to equal OD_{600} from overnight cultures,
507 lysed by boiling for 5 minutes in sample loading buffer, and loaded on 12% SDS-PAGE gels.

508 Levels of WzzE-FLAG were quantitated by densitometry using ImageJ as has been previously
509 described (Mitchell et al., 2018) on immunoblots with the lowest exposure at which WzzE bands
510 could be detected. Samples for ECA analysis were prepared and immunoblotted as previously
511 described with minor modification (Mitchell et al., 2018, Morris and Mitchell, 2023). To avoid
512 ECA degradation, ultrapure water, commercial lysis reagents, commercial running buffer, and
513 precast gels were used for ECA immunoblots. In addition, all western equipment was washed
514 before and after ECA immunoblots with Extran 300 Detergent (Millipore Sigma) and rinsed 10
515 times with ultrapure water. To determine whether proteins affect the apparent chain length of
516 ECA, we pre-treated samples with ProK. Samples were subjected to different pre-treatment
517 conditions, varying both in the concentration of ProK and time of treatment (**Figure S5A**), based
518 on protocols used in previous (Peters et al., 1985, Liu et al., 2020, Leo et al., 2020, Maczuga et
519 al., 2022a) papers. All treatments were performed at 56 °C. Following ProK treatment, samples
520 were boiled again for 10min, and loaded on 15% SDS-PAGE gels. The outcomes of each
521 treatment condition and their respective effects on the resolution and clarity of the immunoblot, as
522 well as effects on ECA levels are shown in (**Figure S5B**). Antibodies used for blotting were as
523 follows: M2 αFLAG (Millipore Sigma, 1:50,000 dilution), αGroEL (Millipore Sigma, 1: 60,000
524 dilution), αECA (a gift from Professor Renato Morona, University of Adelaide, 1:30,000), rabbit
525 α-mouse (Prometheus, 1:100,000 dilution), goat α-rabbit (Prometheus, 1:100,000 dilution), or
526 αBamA (a gift from Professor Thomas Silhavy, Princeton University, 1:50,000 dilution).

527 *Quantification of Cyclic ECA*. Quantification of ECA_{CYC} was performed as previously described,
528 with minor modifications (Mitchell et al., 2018). Briefly, cells were grown in M63 glucose
529 minimum media lacking nitrogen and supplemented with either 0.2% $(^{14}NH_4)_2SO_4$ or
530 $(^{15}NH_4)_2SO_4$. ECA_{CYC} was purified as previously described. Sample analysis was performed
531 using a Thermo Scientific Q Exactive Focus and liquid chromatography unit (UltiMate 3000 RS).
532 Purified samples were dissolved in 0.1% formic acid and separated by injecting 20 µL of sample

533 into the Synergi Hydro-RP (4.6 x 250 mm; 4.0 µm) column (Phenomenax). The mobile phase
534 consisted of 0.1% Formic acid (eluent A) and methanol (eluent B). The flow rate was set at
535 800 µL/min with the following gradient: 0–5.0 min 5% B, 5.0-12.0 min 5-60% B, 12-14 min 60-
536 90% B, 14-17 min 90% B, and returned to 5% B in 0.1 min and held for 5 min at 5% B. The Q
537 Exactive Focus ESI source was operated in full MS (100-1000 m/z). The mass resolution was
538 tuned to 70000 FWHM at m/z 200. The spray voltage was set to 3.3 kV in negative mode, and
539 the sheath gas and auxiliary gas flow rates were set to 60 and 20 arbitrary units, respectively.
540 The transfer capillary and auxiliary gas heater temperatures were held at 380 and 300 °C,
541 respectively. The S-Lens RF level was set at 70 V. Exactive Series 2.11 /Xcalibur 4.2.47
542 software was used for data acquisition and processing. For relative quantification, the ratio of
543 the areas of the ^{14}N to ^{15}N ECA_{CYC} peaks was calculated.
544

545 **Acknowledgements**

546 Thank you to the members of the Mitchell Lab for helpful and productive discussions and for
547 critical reading of our manuscript. We also thank Drs. Jolene Ramsey, Deborah Siegele,
548 Jennifer Herman, and Ryland Young (Texas A&M University), as well as members of the
549 Ramsey lab, for their insightful comments. The ECA antibody was a kind gift from Professor
550 Renato Morana (University of Adelaide). The pCas9 and pCRISPR plasmids were a kind gift
551 from Professor Luciano Marraffini (Rockefeller University) through the Addgene Repository. This
552 work was supported by the National Institute of Allergy and Infectious Disease under award
553 number R01-AI155915 (to A.M.M.).
554

555 **References**

556 ABRAMSON, J., ADLER, J., DUNGER, J., EVANS, R., GREEN, T., PRITZEL, A.,
557 RONNEBERGER, O., WILLMORE, L., BALLARD, A. J., BAMBRICK, J., BODENSTEIN,
558 S. W., EVANS, D. A., HUNG, C.-C., O'NEILL, M., REIMAN, D., TUNYASUVUNAKOOL,

K., WU, Z., ŽEMGULYTė, A., ARVANITI, E., ... & JUMPER, J. M. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature,* 630**,** 493-500.

ADEOLU, M., ALNAJAR, S., NAUSHAD, S. & S. GUPTA, R. 2016. Genome-based phylogeny and taxonomy of the '*Enterobacteriales*': proposal for *Enterobacterales* ord. nov. divided into the families *Enterobacteriaceae*, *Erwiniaceae* fam. nov., *Pectobacteriaceae* fam. nov., *Yersiniaceae* fam. nov., *Hafniaceae* fam. nov., *Morganellaceae* fam. nov., and *Budviciaceae* fam. nov. *International Journal of Systematic and Evolutionary Microbiology,* 66**,** 5575-5599.

AGUINIGA, L. M., YAGGIE, R. E., SCHAEFFER, A. J. & KLUMPP, D. J. 2016. Lipopolysaccharide Domains Modulate Urovirulence. *Infection and Immunity,* 84**,** 3131-3140.

AMINOV, R. I. 2010. A brief history of the antibiotic era: lessons learned and challenges for the future. *Front Microbiol,* 1**,** 134.

ANDREEVA, A. E., BORISSOVA, B. E., MIRONOVA, R., GLYKOS, N. M., KOTSIFAKI, D., IVANOV, I., KRYSTEVA, M. & KOKKINIDIS, M. 2000. Crystallization of type I chloramphenicol acetyltransferase: an approach based on the concept of ionic strength reducers. *Acta Crystallogr D Biol Crystallogr,* 56**,** 101-3.

ARMBRUSTER, C. E., FORSYTH-DEORNELLAS, V., JOHNSON, A. O., SMITH, S. N., ZHAO, L., WU, W. & MOBLEY, H. L. T. 2017. Genome-wide transposon mutagenesis of *Proteus mirabilis*: Essential genes, fitness factors for catheter-associated urinary tract infection, and the impact of polymicrobial infection on fitness requirements. *PLOS Pathogens,* 13**,** e1006434.

ASCARI, A., TRAN, E. N. H., EIJKELKAMP, B. A. & MORONA, R. 2022. Identification of the *Shigella flexneri* Wzy Domain Modulating Wzz(pHS-2) Interaction and Detection of the Wzy/Wzz/Oag Complex. *J Bacteriol,* 204**,** e00224-22.

584 BABA, T., ARA, T., HASEGAWA, M., TAKAI, Y., OKUMURA, Y., BABA, M., DATSENKO, K. A.,
585 TOMITA, M., WANNER, B. L. & MORI, H. 2006. Construction of *Escherichia coli* K-12 in-
586 frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol,* 2**,** 2006.0008.
587 BANERJEE, R. V., JOHNSTON, N. L., SOBESKI, J. K., DATTA, P. & MATTHEWS, R. G. 1989.
588 Cloning and sequence analysis of the *Escherichia coli* metH gene encoding cobalamin-
589 dependent methionine synthase and isolation of a tryptic fragment containing the
590 cobalamin-binding domain. *J Biol Chem,* 264**,** 13888-95.
591 BARR, K., KLENA, J. & RICK, P. 1999. The Modality of Enterobacterial Common Antigen
592 Polysaccharide Chain Lengths Is Regulated by o349 of the wec Gene Cluster of
593 *Escherichia coli K-12*. *Journal of bacteriology,* 181**,** 6564-8.
594 BARR, K., NUNES-EDWARDS, P. & RICK, P. D. 1989. In vitro synthesis of a lipid-linked
595 trisaccharide involved in synthesis of enterobacterial common antigen. *J Bacteriol,* 171**,**
596 1326-32.
597 BARR, K. & RICK, P. D. 1987. Biosynthesis of enterobacterial common antigen in *Escherichia*
598 *coli.* In vitro synthesis of lipid-linked intermediates. *J Biol Chem,* 262**,** 7142-50.
599 BASU, S., KUHN, H.-M., NESZMELYI, A., HIMMELSPACH, K. & MAYER, H. 1987. Chemical
600 characterization of enterobacterial common antigen isolated from *Plesiomonas*
601 *shigelloides* ATCC 14029. *European Journal of Biochemistry,* 162**,** 75-81.
602 BENNETT, H. C. & MITCHELL, A. M. 2025. Recent advances in understanding of
603 enterobacterial common antigen synthesis and regulation. *Open Biology,* 15**,** 250055.
604 BERTANI, B. & RUIZ, N. 2018. Function and Biogenesis of Lipopolysaccharides. *EcoSal Plus,*
605 8**,** ESP-0001-2018.
606 COLLINS, R. F., KARGAS, V., CLARKE, B. R., SIEBERT, C. A., CLARE, D. K., BOND, P. J.,
607 WHITFIELD, C. & FORD, R. C. 2017. Full-length, Oligomeric Structure of Wzz
608 Determined by Cryoelectron Microscopy Reveals Insights into Membrane-Bound States.
609 *Structure,* 25**,** 806-815.e3.

610 DANIELS, C. & MORONA, R. 1999. Analysis of *Shigella flexneri* wzz (Rol) function by
611 mutagenesis and cross-linking: wzz is able to oligomerize. *Mol Microbiol,* 34**,** 181-94.
612 DATSENKO, K. A. & WANNER, B. L. 2000. One-step inactivation of chromosomal genes in
613 *Escherichia coli K-12* using PCR products. *Proceedings of the National Academy of*
614 *Sciences,* 97**,** 6640-6645.
615 DOERN, C. D. 2024. Chapter 2 - Classification of medically important bacteria. *In:* TANG, Y.-W.,
616 HINDIYEH, M. Y., LIU, D., SAILS, A., SPEARMAN, P. & ZHANG, J.-R. (eds.) *Molecular*
617 *Medical Microbiology (Third Edition).* Academic Press.
618 EADE, C. R., WALLEN, T. W., GATES, C. E., OLIVERIO, C. L., SCARBROUGH, B. A., REID, A.
619 J., JORGENSON, M. A., YOUNG, K. D. & TROUTMAN, J. M. 2021. Making the
620 enterobacterial common antigen glycan and measuring its substrate sequestration. *ACS*
621 *Chem Biol,* 16**,** 691-700.
622 ERBEL, P. J., BARR, K., GAO, N., GERWIG, G. J., RICK, P. D. & GARDNER, K. H. 2003.
623 Identification and biosynthesis of cyclic enterobacterial common antigen in *Escherichia*
624 *coli*. *J Bacteriol,* 185**,** 1995-2004.
625 ERBEL, P. J., SEIDEL, R., MACINTOSH, S. E., GENTILE, L. N., AMOR, J. C., KAHN, R. A.,
626 PRESTEGARD, J. H., MCINTOSH, L. P. & GARDNER, K. H. 2004. Cyclic
627 enterobacterial common antigen: potential contaminant of bacterially expressed protein
628 preparations. *J Biomol NMR,* 29**,** 199-204.
629 FAINGOLD, O., COHEN, T. & SHAI, Y. 2012. A GxxxG-like Motif within HIV-1 Fusion Peptide Is
630 Critical to Its Immunosuppressant Activity, Structure, and Interaction with the
631 Transmembrane Domain of the T-cell Receptor*. *Journal of Biological Chemistry,* 287**,**
632 33503-33511.
633 FARFOUR, E., DORTET, L., GUILLARD, T., CHATELAIN, N., POISSON, A., MIZRAHI, A.,
634 FOURNIER, D., BONNIN, R. A., DEGAND, N., MORAND, P., JANVIER, F., FIHMAN, V.,
635 CORVEC, S., BROUTIN, L., LE BRUN, C., YIN, N., HéRY-ARNAUD, G., GRILLON, A.,

636 BILLE, E., ... & ON BEHALF OF THE GMC STUDY, G. 2022. Antimicrobial Resistance in
637 *Enterobacterales* Recovered from Urinary Tract Infections in France. *Pathogens,* 11**,**
638 356.
639 FäRNBäCK, M., ERIKSSON, L., SENCHENKOVA, S. Y., ZYCH, K., KNIREL, Y. A.,
640 SIDORCZYK, Z. & WIDMALM, G. 2003. Crystal Structure of a Cyclic Enterobacterial
641 Common Antigen. *Angewandte Chemie International Edition,* 42**,** 2543-2546.
642 GILBREATH, J. J., COLVOCORESSES DODDS, J., RICK, P. D., SOLOSKI, M. J., MERRELL,
643 D. S. & METCALF, E. S. 2012. Enterobacterial common antigen mutants of *Salmonella*
644 *enterica serovar Typhimurium* establish a persistent infection and provide protection
645 against subsequent lethal challenge. *Infect Immun,* 80**,** 441-50.
646 GOźDZIEWICZ, T. K., ŁUKASIEWICZ, J. & ŁUGOWSKI, C. 2015. [The structure and
647 significance of enterobacterial common antigen (ECA)]. *Postepy Hig Med Dosw (Online),*
648 69**,** 1003-12.
649 GUéRIN, F., LALLEMENT, C., GOUDERGUES, B., ISNARD, C., SANGUINETTI, M., CACACI,
650 M., TORELLI, R., CATTOIR, V. & GIARD, J.-C. 2020. Landscape of in vivo Fitness-
651 Associated Genes of *Enterobacter cloacae* Complex. *Frontiers in Microbiology,* 11**,** 1609.
652 GUZMAN, L. M., BELIN, D., CARSON, M. J. & BECKWITH, J. 1995. Tight regulation,
653 modulation, and high-level expression by vectors containing the arabinose P_{BAD}
654 promoter. *J Bacteriol,* 177**,** 4121-30.
655 HOLMES, C. L., WILCOX, A. E., FORSYTH, V., SMITH, S. N., MORICZ, B. S.,
656 UNVERDORBEN, L. V., MASON, S., WU, W., ZHAO, L., MOBLEY, H. L. T. &
657 BACHMAN, M. A. 2023. *Klebsiella pneumoniae* causes bacteremia using factors that
658 mediate tissue-specific fitness and resistance to oxidative stress. *PLOS Pathogens,* 19**,**
659 e1011233.

660 HONG, M. & PAYNE, S. M. 1997. Effect of mutations in *Shigella flexneri* chromosomal and
661 plasmid-encoded lipopolysaccharide genes on invasion and serum resistance. *Mol*
662 *Microbiol,* 24**,** 779-91.
663 HONG, Y., HU, D., VERDEROSA ANTHONY, D., QIN, J., TOTSIKA, M. & REEVES PETER, R.
664 2023. Repeat-Unit Elongations To Produce Bacterial Complex Long Polysaccharide
665 Chains, an O-Antigen Perspective. *EcoSal Plus,* 11**,** eesp-0020-2022.
666 HUSZCZYNSKI, S. M., COUMOUNDOUROS, C., PHAM, P., LAM, J. S. & KHURSIGARA, C. M.
667 2019. Unique regions of the polysaccharide copolymerase Wzz_2 from *Pseudomonas*
668 *aeruginosa* are essential for O-specific antigen chain length control. *Journal of*
669 *Bacteriology,* 201**,** e00165-19.
670 ISLAM, S. T. & LAM, J. S. 2014. Synthesis of bacterial polysaccharides via the Wzx/Wzy-
671 dependent pathway. *Can J Microbiol,* 60**,** 697-716.
672 JENKINS, C., RENTENAAR, R. J., LANDRAUD, L. & BRISSE, S. 2017. 180 -
673 *Enterobacteriaceae*. *In:* COHEN, J., POWDERLY, W. G. & OPAL, S. M. (eds.) *Infectious*
674 *Diseases (Fourth Edition).* Elsevier.
675 JIANG, W., BIKARD, D., COX, D., ZHANG, F. & MARRAFFINI, L. A. 2013. RNA-guided editing
676 of bacterial genomes using CRISPR-Cas systems. *Nature biotechnology,* 31**,** 233-239.
677 JIANG, X. E., TAN, W. B., SHRIVASTAVA, R., SEOW, D. C. S., CHEN, S. L., GUAN, X. L. &
678 CHNG, S.-S. 2020. Mutations in enterobacterial common antigen biosynthesis restore
679 outer membrane barrier function in Escherichia coli tol-pal mutants. *Molecular*
680 *Microbiology,* 114**,** 991-1005.
681 KAJIMURA, J., RAHMAN, A., HSU, J., EVANS, M. R., GARDNER, K. H. & RICK, P. D. 2006. O
682 acetylation of the enterobacterial common antigen polysaccharide is catalyzed by the
683 product of the *yiaH* Gene of *Escherichia coli* K-12. *Journal of Bacteriology,* 188**,** 7542-
684 7550.

685 KAJIMURA, J., RAHMAN, A. & RICK, P. D. 2005. Assembly of cyclic enterobacterial common
686 antigen in *Escherichia coli* K-12. *J Bacteriol,* 187**,** 6917-6927.
687 KALYNYCH, S., CHERNEY, M., BOSTINA, M., ROUILLER, I. & CYGLER, M. 2015. Quaternary
688 structure of WzzB and WzzE polysaccharide copolymerases. *Protein Science,* 24**,** 58-
689 69.
690 KALYNYCH, S., MORONA, R. & CYGLER, M. 2014. Progress in understanding the assembly
691 process of bacterial O-antigen. *FEMS Microbiol Rev,* 38**,** 1048-65.
692 KALYNYCH, S., RUAN, X., VALVANO, M. A. & CYGLER, M. 2011. Structure-guided
693 investigation of lipopolysaccharide O-antigen chain length regulators reveals regions
694 critical for modal length control. *Journal of Bacteriology,* 193**,** 3710-3721.
695 KALYNYCH, S., VALVANO, M. A. & CYGLER, M. 2012a. Polysaccharide co-polymerases: the
696 enigmatic conductors of the O-antigen assembly orchestra. *Protein Eng Des Sel,* 25**,**
697 797-802.
698 KALYNYCH, S., YAO, D., MAGEE, J. & CYGLER, M. 2012b. Structural characterization of
699 closely related O-antigen lipopolysaccharide (LPS) chain length regulators. *J Biol Chem,*
700 287**,** 15696-705.
701 KINTZ, E. N. & GOLDBERG, J. B. 2011. Site-directed mutagenesis reveals key residue for O
702 antigen chain length regulation and protein stability in *Pseudomonas aeruginosa* Wzz2.
703 *J Biol Chem,* 286**,** 44277-44284.
704 KUNIN, C. M. 1963. Separtion, chacterization, and Biological Significance of a Common Antigen
705 in *Enterobacteriaceae*. *J Exp Med,* 118**,** 565-86.
706 KUNIN, C. M., BEARD, M. V. & HALMAGYI, N. E. 1962. Evidence for a Common Hapten
707 Associated with Endotoxin Fractions of *E. coli* and other *Enterobacteriaceae*. *Proc Soc*
708 *Exp Biol Med,* 111**,** 160-166.

709 LEO, V., TEH, M. Y., TRAN, E. N. H. & MORONA, R. 2021a. Identification of a Region in
710 Shigella flexneri WzyB Disrupting the Interaction with Wzz(pHS2). *J Bacteriol,* 203**,**
711 e0041321.
712 LEO, V., TRAN, E. & MORONA, R. 2020. Polysaccharide co-polymerase WzzB/WzzE chimeras
713 reveal transmembrane 2 region of WzzB is important for interaction with WzyB. *J*
714 *Bacteriol,* 203.
715 LEO, V., TRAN, E. & MORONA, R. 2021b. Polysaccharide co-polymerase WzzB/WzzE
716 chimeras reveal transmembrane 2 region of WzzB is important for interaction with WzyB.
717 *J Bacteriol,* 203**,** 00598-20.
718 LIU, Q., SHEN, X., BIAN, X. & KONG, Q. 2020. Effect of deletion of gene cluster involved in
719 synthesis of Enterobacterial common antigen on virulence and immunogenicity of live
720 attenuated Salmonella vaccine when delivering heterologous *Streptococcus*
721 *pneumoniae* antigen PspA. *BMC Microbiology,* 20**,** 150.
722 LUTZ, R. & BUJARD, H. 1997. Independent and tight regulation of transcriptional units in
723 *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic*
724 *Acids Res,* 25**,** 1203-10.
725 MACZUGA, N., TRAN, E. N. H. & MORONA, R. 2022a. Subcellular localization of the
726 enterobacterial common antigen GT-E-like glycosyltransferase, WecG. *Mol Microbiol,*
727 118**,** 403-416.
728 MACZUGA, N., TRAN, E. N. H., QIN, J. & MORONA, R. 2022b. Interdependence of *Shigella*
729 *flexneri* O Antigen and Enterobacterial Common Antigen Biosynthetic Pathways. *J*
730 *Bacteriol,* 204**,** e0054621.
731 MAHER, C. & HASSAN, K. A. 2023. The Gram-negative permeability barrier: tipping the
732 balance of the in and the out. *mBio,* 14**,** e01205-23.
733 MANNEL, D. & MAYER, H. 1978. Isolation and Chemical Characterization of the Enterobacterial
734 Common Antigen. *European Journal of Biochemistry,* 86**,** 361-370.

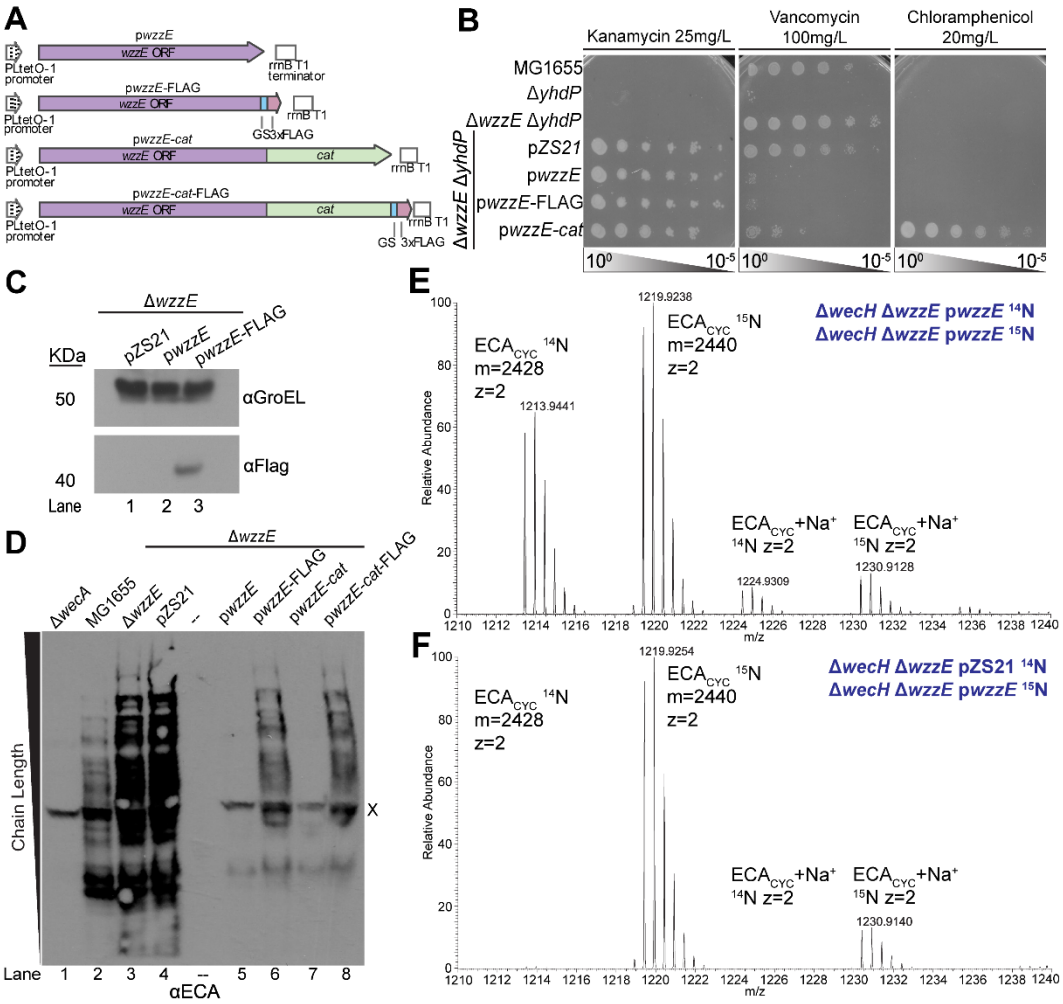
735 MAY, K. L. & SILHAVY, T. J. 2018. The *Escherichia coli* phospholipase PldA regulates outer
736 membrane homeostasis via lipid signaling. *mBio,* 9**,** e00379-18.
737 MENDEZ-PEREZ, D., GUNASEKARAN, S., ORLER, V. J. & PFLEGER, B. F. 2012. A
738 translation-coupling DNA cassette for monitoring protein translation in *Escherichia coli*.
739 *Metabolic Engineering,* 14**,** 298-305.
740 MITCHELL, A. M., SRIKUMAR, T. & SILHAVY, T. J. 2018. Cyclic Enterobacterial Common
741 Antigen Maintains the Outer Membrane Permeability Barrier of *Escherichia coli* in a
742 Manner Controlled by YhdP. *mBio,* 9**,** 01321-18.
743 MORONA, R., PURINS, L., TOCILJ, A., MATTE, A. & CYGLER, M. 2009. Sequence–structure
744 relationships in polysaccharide co-polymerase (PCP) proteins. *Trends in biochemical*
745 *sciences,* 34**,** 78-84.
746 MORONA, R., VAN DEN BOSCH, L. & MANNING, P. A. 1995. Molecular, genetic, and
747 topological characterization of O-antigen chain length regulation in *Shigella flexneri*. *J*
748 *Bacteriol,* 177**,** 1059-68.
749 MORRIS, K. N. & MITCHELL, A. M. 2023. Phosphatidylglycerol Is the Lipid Donor for Synthesis
750 of Phospholipid-Linked Enterobacterial Common Antigen. *Journal of Bacteriology,* 205**,**
751 e00403-22.
752 MUELLER, B. K., SUBRAMANIAM, S. & SENES, A. 2014. A frequent, GxxxG-mediated,
753 transmembrane association motif is optimized for the formation of interhelical Cα–H
754 hydrogen bonds. *Proceedings of the National Academy of Sciences,* 111**,** E888-E895.
755 NATH, P. & MORONA, R. 2015. Detection of Wzy/Wzz interaction in *Shigella flexneri*.
756 *Microbiology (Reading),* 161**,** 1797-1805.
757 NICHOLS, B. P., SHAFIQ, O. & MEINERS, V. 1998. Sequence analysis of Tn10 insertion sites
758 in a collection of Escherichia coli strains used for genetic mapping and strain
759 construction. *J Bacteriol,* 180**,** 6408-11.

760 PAPADOPOULOS, M. & MORONA, R. 2010. Mutagenesis and chemical cross-linking suggest
761 that Wzz dimer stability and oligomerization affect lipopolysaccharide O-antigen modal
762 chain length control. *J Bacteriol,* 192**,** 3385-93.
763 PAPADOPOULOS, M., TRAN, E. N. H., MURRAY, G. L. & MORONA, R. 2016. Conserved
764 transmembrane glycine residues in the *Shigella flexneri* polysaccharide co-polymerase
765 protein WzzB influence protein-protein interactions. *Microbiology (Reading),* 162**,** 921-
766 929.
767 PARRISH, H. L., GLASSMAN, C. R., KEENEN, M. M., DESHPANDE, N. R., BRONNIMANN, M.
768 P. & KUHNS, M. S. 2015. A Transmembrane Domain GGxxG Motif in CD4 Contributes to
769 Its Lck-Independent Function but Does Not Mediate CD4 Dimerization. *PLOS ONE,* 10**,**
770 e0132333.
771 PETERS, H., JüRS, M., JANN, B., JANN, K., TIMMIS, K. N. & BITTER-SUERMANN, D. 1985.
772 Monoclonal antibodies to enterobacterial common antigen and to Escherichia coli
773 lipopolysaccharide outer core: demonstration of an antigenic determinant shared by
774 enterobacterial common antigen and E. coli K5 capsular polysaccharide. *Infect Immun,*
775 50**,** 459-66.
776 PRAKASH, A., JANOSI, L. & DOXASTAKIS, M. 2011. GxxxG motifs, phenylalanine, and
777 cholesterol guide the self-association of transmembrane domains of ErbB2 receptors.
778 *Biophys J,* 101**,** 1949-58.
779 PYNE, M. E., MOO-YOUNG, M., CHUNG, D. A. & CHOU, C. P. 2015. Coupling the
780 CRISPR/Cas9 System with Lambda Red Recombineering Enables Simplified
781 Chromosomal Gene Replacement in *Escherichia coli*. *Appl Environ Microbiol,* 81**,** 5103-
782 14.
783 RAI, A. K., CARR, J. F., BAUTISTA, D. E., WANG, W. & MITCHELL, A. M. 2021. ElyC and cyclic
784 enterobacterial common antigen regulate synthesis of phosphoglyceride-linked
785 enterobacterial common antigen. *mBio,* 12**,** e02846-21.

786 RAI, A. K. & MITCHELL, A. M. 2020. Enterobacterial Common Antigen: Synthesis and Function
787 of an Enigmatic Molecule. *mBio,* 11**,** mBio.01914-20.
788 RAMIREZ, D. & GIRON, M. 2024. Enterobacter Infections. *StatPearls.* Treasure Island (FL):
789 StatPearls Publishing Copyright © 2024, StatPearls Publishing LLC.
790 RICK, P. D., MAYER, H., NEUMEYER, B. A., WOLSKI, S. & BITTER-SUERMANN, D. 1985.
791 Biosynthesis of enterobacterial common antigen. *J Bacteriol,* 162**,** 494-503.
792 SIEBER, V. 2003. Selection for Soluble Proteins via Fusion with Chloramphenicol
793 Acetyltransferase. *In:* ARNOLD, F. H. & GEORGIOU, G. (eds.) *Directed Enzyme*
794 *Evolution: Screening and Selection Methods.* Totowa, NJ: Humana Press.
795 SILHAVY, T. J., BERMAN, M. L. & ENQUIST, L. W. 1984. Experiments with gene fusions. *(No*
796 *Title)*.
797 SILHAVY, T. J., KAHNE, D. & WALKER, S. 2010. The bacterial cell envelope. *Cold Spring Harb*
798 *Perspect Biol,* 2**,** a000414.
799 TEESE, M. G. & LANGOSCH, D. 2015. Role of GxxxG Motifs in Transmembrane Domain
800 Interactions. *Biochemistry,* 54**,** 5125-35.
801 TOCILJ, A., MUNGER, C., PROTEAU, A., MORONA, R., PURINS, L., AJAMIAN, E., WAGNER,
802 J., PAPADOPOULOS, M., VAN DEN BOSCH, L., RUBINSTEIN, J. L., FéTHIèRE, J.,
803 MATTE, A. & CYGLER, M. 2008. Bacterial polysaccharide co-polymerases share a
804 common framework for control of polymer length. *Nat Struct Mol Biol,* 15**,** 130-8.
805 TRAN, E. N. H. & MORONA, R. 2013. Residues located inside the *Escherichia coli* FepE
806 protein oligomer are essential for lipopolysaccharide O-antigen modal chain length
807 regulation. *Microbiology (Reading),* 159**,** 701-714.
808 WECKENER, M., WOODWARD, L. S., CLARKE, B. R., LIU, H., WARD, P. N., LE BAS, A.,
809 BHELLA, D., WHITFIELD, C. & NAISMITH, J. H. 2023. The lipid linked oligosaccharide
810 polymerase Wzy and its regulating co-polymerase, Wzz, from enterobacterial common
811 antigen biosynthesis form a complex. *Open Biology,* 13**,** 220373.

812 WHITFIELD, C., WILLIAMS, D. M. & KELLY, S. D. 2020. Lipopolysaccharide O-antigens-
813 bacterial glycans made to measure. *J Biol Chem,* 295**,** 10593-10609.
814 WISEMAN, B., NITHARWAL, R. G., WIDMALM, G. & HöGBOM, M. 2021. Structure of a full-
815 length bacterial polysaccharide co-polymerase. *Nat Commun,* 12**,** 369.
816 WISEMAN, B., WIDMALM, G. & HöGBOM, M. 2023. Alternating L4 loop architecture of the
817 bacterial polysaccharide co-polymerase WzzE. *Commun Biol,* 6**,** 802.
818
819

820 **Figures and Legends**



821

822 **Figure 1: Plasmid-based expression of *wzzE* partially complements a chromosomal *wzzE***
823 **deletion. (A)** Diagrams of constructs for expression of wild-type WzzE, WzzE with a C-terminal
824 glycine-serine linker and triple FLAG tag, WzzE translationally fused with chloramphenicol
825 acetyltransferase (Cat), and WzzE fused with Cat and triple FLAG are shown. The indicated
826 constructs were expressed from the low copy number pZS21 plasmid (Lutz and Bujard, 1997).
827 **(B)** The vancomycin sensitivity of a Δ*yhdP* strain can be suppressed by loss of WzzE (Mitchell
828 et al., 2018). When the Δ*wzzE* Δ*yhdP* strain carried the pZS21 empty plasmid, it remained
829 resistance to vancomycin as assayed by EOP but became sensitive when *wzzE, wzzE*-FLAG,
830 or *wzzE-cat* was expressed, indicating all three constructs complemented Δ*wzzE*. In addition,

831 translational fusion of *cat* and *wzzE* conferred chloramphenicol resistance, suggesting this
832 construct could readout WzzE stability and multimerization based on chloramphenicol
833 resistance. **(C)** WzzE-FLAG was detectable by immunoblotting. GroEL serves as a loading
834 control. **(D)** ECA chain-length regulation was assayed through immunoblotting using an ECA
835 antibody. Plasmid-base expression of WzzE restored chain-length regulation of linear ECA,
836 although lower levels of ECA and shorter chain-length were observed than in the wild-type strain
837 for the p*wzzE* and p*wzzE* constructs, while the p*wzzE*-FLAG and p*wzzE-cat*-FLAG constructs
838 complemented better. The Δ*wecA* strain produces no ECA and serves as negative control. X:
839 nonspecific band. **(E-F)** The strain carrying a plasmid with wild-type *wzzE* was grown with ^{15}N to
840 label ECA and then combined with an unlabeled strain carrying either wild-type *wzzE* (E) or the
841 empty vector (F). ECA_{CYC} was purified and then detected by LC-MS. MS traces are shown for
842 the indicated samples. The mass, charge, and identity of the ECA_{CYC}-associated peaks are
843 shown. Expression of *wzzE* from p*wzzE* restored ECA_{CYC} synthesis in the Δ*wzzE* strain. The
844 strains contained Δ*wecH* to prevent non-stochiometric acetylation of ECA. All images are
845 representative of at least three independent experiments.
846

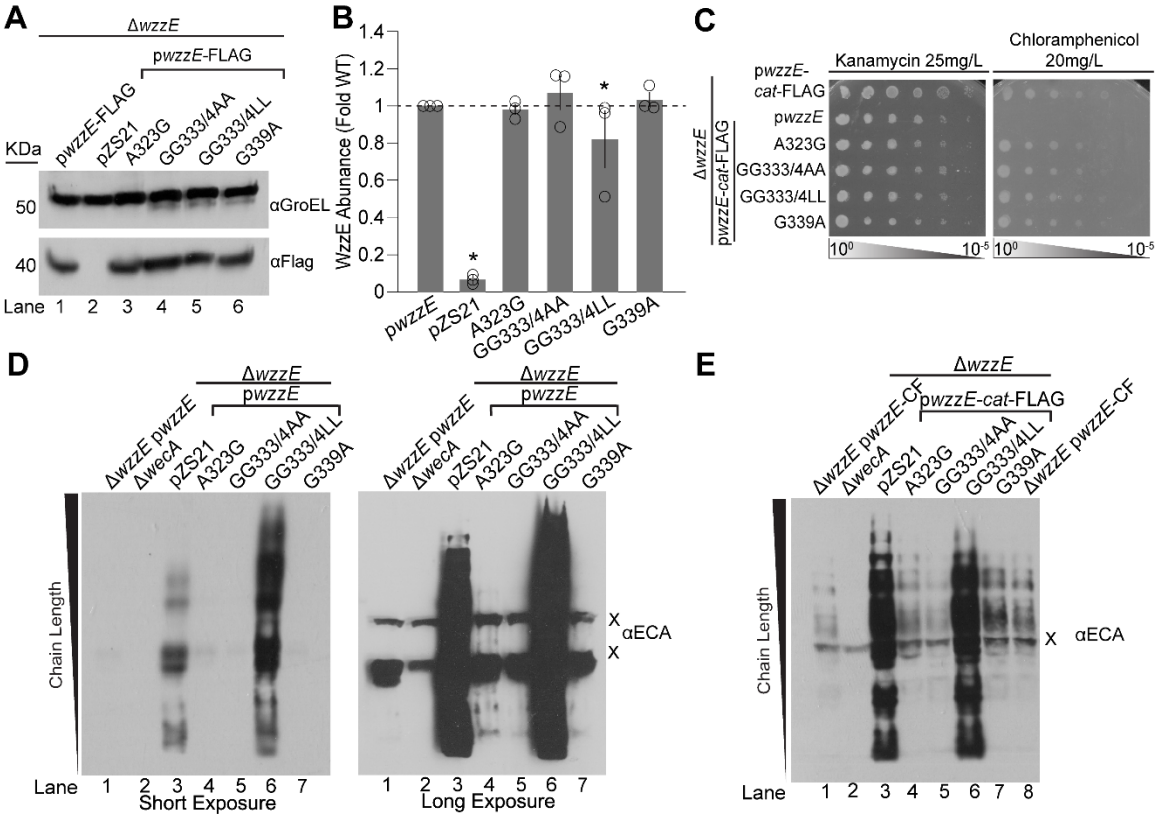


Figure 2: Alteration of WzzE TMH2 can prevent ECA chain length regulation without affecting WzzE stability. (A) The effect of the indicated *wzzE* site-directed mutations on WzzE-FLAG protein levels was assessed using immunoblotting. GroEL serves as a loading control. **(B)** Densitometry was used for 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, except for a small, but significant decrease in the WzzE$^{GG333/4LL}$ variant. Data are shown as the average of three biological replicates ± SEM and individual datapoints. * $p<0.05$ by Mann-Whitney 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 immunoblotting. Short and long exposures are shown. “X” indicates non-specific bands. WzzE$^{GG333/4LL}$ showed loss of

861 chain-length regulation while the other variants retained wild-type function. The Δ*wecA* strain
862 serves as negative control. **(E)** Chain-length regulation of the WzzE-Cat-FLAG variants was
863 assayed by immunoblot. Similar phenotypes were observed to those found in the WzzE
864 variants. “X” indicates non-specific band. All images are representative of at least three
865 independent experiments.
866

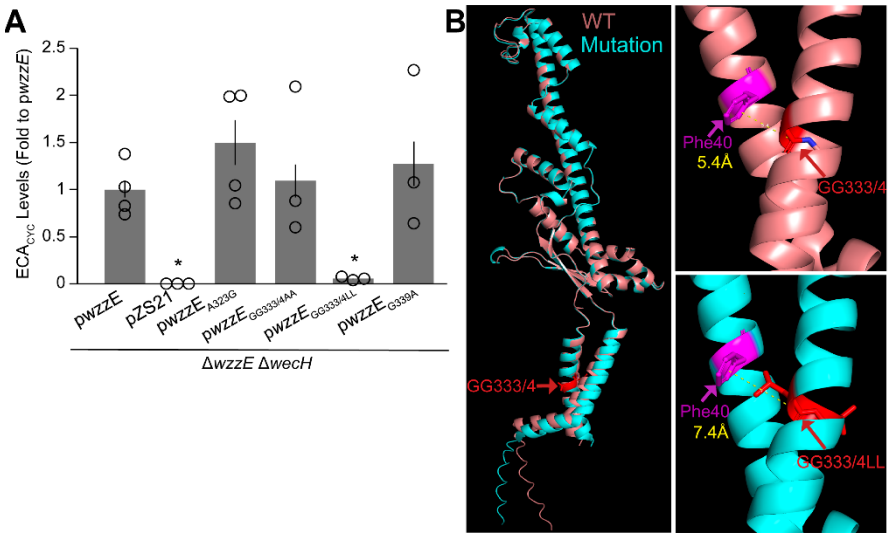


Figure 3. Conserved glycines in WzzE TMH 2 are required for ECA$_{CYC}$ synthesis. (A) The strain carrying a plasmid with wild-type *wzzE* was grown with ^{15}N to label ECA and then combined with an unlabeled strain carrying the indicated *wzzE* mutant or control. ECA$_{CYC}$ was purified and the relative amount of unlabeled to labeled ECA$_{CYC}$ was determined by LC-MS. WzzE$^{GG333/4LL}$ showed a loss of the ability to produce ECA$_{CYC}$. All plasmids were expressed in a Δ*wzzE* Δ*wecH* strain background to prevent non-stochiometric acetylation of ECA. Data are shown as the average of three biological replicates ± SEM and as individual datapoints. * $p<0.05$ by Mann-Whitney test. **(B)** The predicted structure of a WzzE octamer with WzyE was calculated using AlphaFold 3 (Abramson et al., 2024). The structures of wild-type WzzE and WzzE$^{GG333/4LL}$ monomers were aligned with the mutated residues shown in red. The insets show the predicted distance between the two WzzE TMH in the wild type WzzE and WzzE$^{GG333/4LL}$.

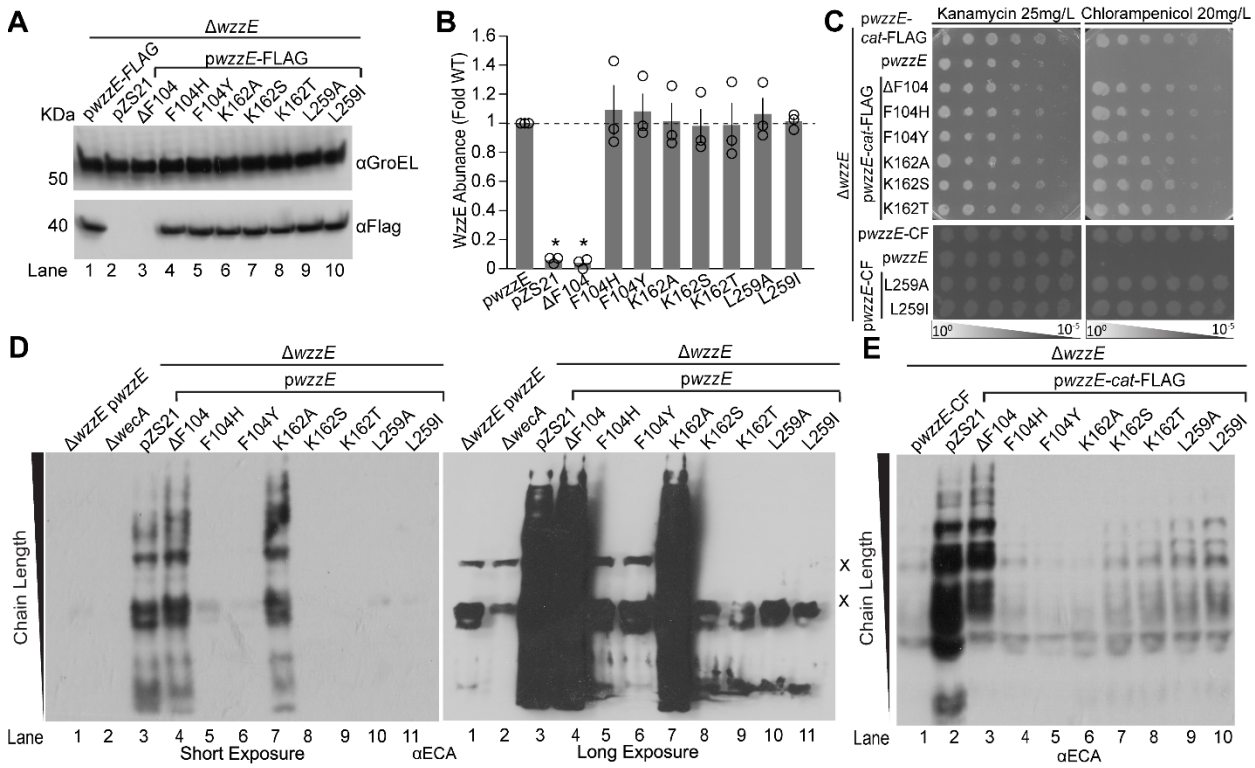


Figure 4: The periplasmic region of WzzE contributes to WzzE stability and chain length regulation. (A) Immunoblotting was used to determine protein levels for the indicated WzzE-FLAG variants. GroEL serves as a loading control. **(B)** WzzE-FLAG levels from (A) were quantified by densitometry. WzzE-FLAG 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 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 tag is 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. The $WzzE^{\Delta F104}$ and $WzzE^{K162A}$ variants showed loss of ECA chain-length regulation while the other

895 variants had wild-type chain-length regulation. The Δ*wecA* strain serves as negative control. **(E)**
896 The effect of WzzE-Cat-FLAG variants on ECA chain-length regulation was assayed by
897 immunoblot. The variants show similar phenotypes as in (D) except for $WzzE^{K162A}$, which shows
898 normal chain-length regulation and $WzzE^{\Delta F104}$ which shows partial chain-length dysregulation.
899 All images are representative of at least three independent experiments.
900

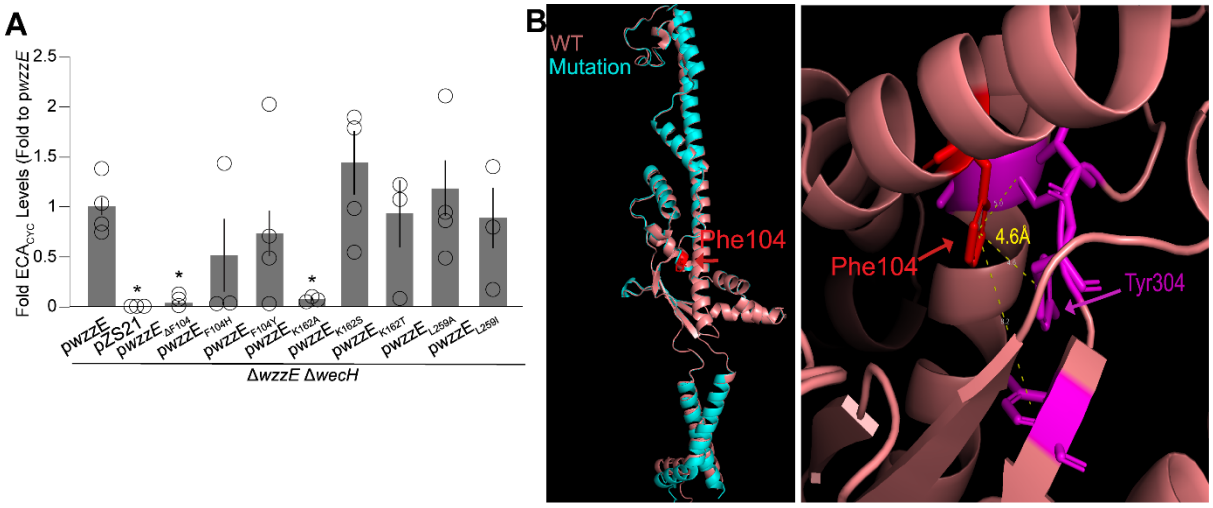


Figure 5: The periplasmic domain of WzzE contributes to ECA_{CYC} synthesis. (A) The strain carrying wild-type *wzzE* was grown with ^{15}N to label ECA and then combined with an unlabeled strain carrying the indicated *wzzE* mutant or control strain. After ECA_{CYC} purification, the relative amount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

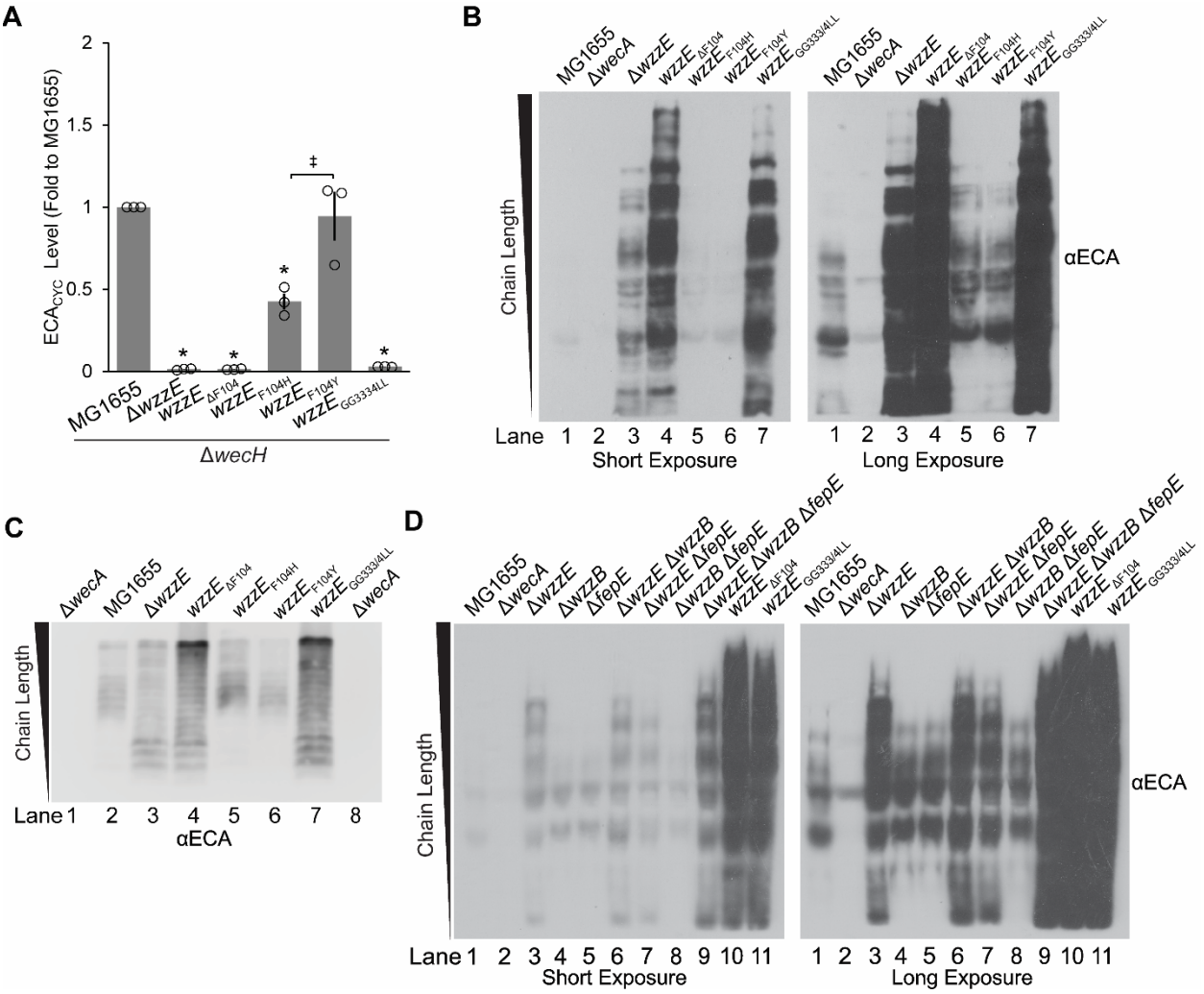


Figure 6: Levels of ECA$_{CYC}$ production and chain-length regulation are genetically separable. CRISPR-Cas9 was used to make the indicated mutations to *wzzE* in its native locus. All strains contain *metE*::Tn*10*. **(A)** The strain wild type for *wzzE* was grown with ^{15}N to label ECA and then 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. Data are shown as the average of three biological replicates ± SEM and as individual datapoints. * $p<0.05$ vs. wild type by Mann-Whitney test; ‡ $p<0.05$ between the indicated strains by Mann-Whitney test. **(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

924 resulted in linear ECA dysregulation, while the other mutations more closely resembled wild-
925 type linear ECA regulation. The Δ*wecA* strain serves as negative control. **(C)**. ECA
926 Immunoblotting was performed on the indicated chromosomal mutants after ProK treatment to
927 more clearly assess chain length. WzzE mutants ΔF104 and GG333/4LL resulted in
928 dysregulation of linear ECA, whereas F104Y and F104H strongly resembled wild-type linear
929 ECA regulation. **(D)** Immunoblotting was used to assay ECA chain-length regulation in the
930 indicated strains. Deletion of *wzzB* and/or *fepE* in combination with Δ*wzzE* did not cause the
931 increased levels of linear ECA observed with the loss-of-function point mutants; however, the
932 Δ*wzzE* Δ*wzzB* Δ*fepE* strain did have a small increase in linear ECA levels compared to Δ*wzzE*
933 alone. All images are representative of at least three independent experiments.