

1 **Engineered shell proteins confer improved encapsulated pathway behavior in** 2 **a bacterial microcompartment**

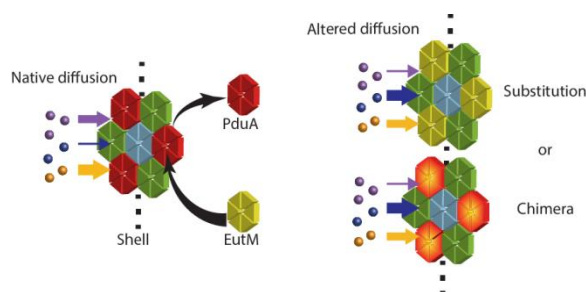
3 Marilyn F. Slininger Lee[†], Christopher M. Jakobson[†], Danielle Tullman-Ercek^{‡*}

4 [†] Department of Chemical and Biomolecular Engineering, University of California Berkeley,
5 United States

6 [‡] Department of Chemical and Biological Engineering, Northwestern University, 2145 Sheridan
7 Rd, Technological Institute E136, Evanston, IL 60208-3120, USA; Chemistry of Life Processes
8 Institute, Northwestern University, 2170 Campus Drive, Evanston, IL 60208-3120

9 *Corresponding author

10 **TOC Abstract Graphic**



11
12 **Abstract:** Bacterial microcompartments are a class of proteinaceous organelles comprising a
13 characteristic protein shell enclosing a set of enzymes. Compartmentalization can prevent escape of
14 volatile or toxic intermediates, prevent off-pathway reactions, and create private cofactor pools.
15 Encapsulation in synthetic microcompartment organelles will enhance the function of heterologous
16 pathways, but to do so, it is critical to understand how to control diffusion in and out of the
17 microcompartment organelle. To this end, we explored how small differences in the shell protein structure
18 result in changes in the diffusion of metabolites through the shell. We found that the ethanolamine
19 utilization (Eut) protein EutM properly incorporates into the 1,2-propanediol utilization (Pdu)
20 microcompartment, altering native metabolite accumulation and the resulting growth on 1,2-propanediol
21 as the sole carbon source. Further, we identified a single pore-lining residue mutation that confers the
22 same phenotype as substitution of the full EutM protein, indicating that small molecule diffusion through
23 the shell is the cause of growth enhancement. Finally, we show that the hydropathy index and charge of
24 pore amino acids are important indicators to predict how pore mutations will affect growth on 1,2-

25 propanediol, likely by controlling diffusion of one or more metabolites. This study highlights the success
26 of two strategies to engineer microcompartment control over metabolite transport: altering the existing
27 shell protein pore via mutation of the pore-lining residues, and generating chimeras using shell proteins
28 with the desired pores.

29 **Keywords:** Bacterial microcompartment, Protein shell assembly, Protein pore diffusion, Metabolosome,
30 *Salmonella enterica*, 1,2-propanediol

31 Bacterial microcompartments (MCPs) are a prokaryotic form of cellular organization consisting
32 of a protein shell used to enhance pathways by encapsulation. MCP-associated pathways include those
33 found in hosts across many bacterial phyla for metabolism of niche carbon sources, such as 1,2-
34 propanediol (1,2-PD), ethanolamine, and ethanol (involving Pdu, Eut, and Eut MCPs, respectively), and
35 those found in cyanobacteria and chemoautotrophs for carbon fixation by RuBisCo in the carboxysome.¹⁻⁴
36 Additional naturally-encapsulated pathways have been recently discovered and a summary of this topic
37 may be found in a number of review articles.⁵⁻⁸ Some bacterial species have several MCP systems,
38 encoding different organelles tailored to metabolize specific substrates.⁹⁻¹² For example, two well-
39 characterized model systems are found in *Salmonella enterica*, which can produce both the Eut and Pdu
40 MCPs from two separate operons. These enable growth on ethanolamine and 1,2-PD, respectively, which
41 are both found in the human gut. The Pdu MCP encapsulates enzymes converting 1,2-PD through a
42 propionaldehyde intermediate into either 1-propanol or propionate (Figure 1a), while the Eut MCP
43 similarly converts ethanolamine through an acetaldehyde intermediate into ethanol and acetyl phosphate.
44 The benefits of encapsulating a pathway include preventing the escape or toxicity of intermediates as well
45 as cycling of cofactors.¹³⁻¹⁷

46 The *S. enterica* Pdu MCP assembly is an irregular polyhedron 150-200 nm in diameter and is
47 enclosed by a shell composed of a few thousand copies of shell proteins of several types.¹⁸ There are eight
48 Pdu shell protein paralogs that are present in the final assembly in distinct and reproducible ratios.¹⁹ The
49 *S. enterica* Eut MCP is similarly composed of five shell proteins. MCP shell proteins form trimers,

50 pentamers, or hexamers that come together to assemble the hexagonal or pentagonal tiles of the
51 polyhedron. A pore is formed at the symmetric center of each tile.^{18–21}

52 The central pores of MCP shell proteins are implicated in the transport of small molecule
53 substrates and products across the shell.^{22, 23} For instance, Pdu MCP shell protein PduA is abundant in the
54 shell and is suggested to be involved in the selective transport of 1,2-PD.²³ Specifically, occlusion of its
55 pore by PduA^{S40GSG}, PduA^{S40L}, PduA^{S40C}, and PduA^{S40Q} mutations lead to growth defects that are rescued
56 by high 1,2-PD concentrations, while the PduA^{S40A} mutation increased the escape of propionaldehyde.²³
57 Similarly, the pores of the carboxysome are hypothesized to control flux of bicarbonate and CO₂ across
58 the shell.²⁴ All characterized carboxysome pore structures are positively charged at the narrowest point,
59 which could enhance the flux of negatively charged bicarbonate relative to the escape of CO₂. Since
60 different pore structures are postulated to lead to altered diffusivity of metabolites, engineering the pores
61 could be the key to tailoring MCP properties for encapsulating non-native systems.

62 An alternative strategy for engineering MCP properties for non-native pathways is to use novel
63 combinations of natural shell proteins. Shell proteins from different MCP systems have the potential to
64 interact to form chimeric assemblies if expressed together at appropriate levels. Cai et al. demonstrated
65 that an alpha-carboxysome shell protein complements the deletion of a beta-carboxysome shell protein.²⁵
66 Sturms et al. observed that co-expression of the Eut and Pdu operons in *S. enterica* from a synthetic
67 promoter system resulted in co-purification of Eut and Pdu components, albeit with a disruption of Pdu
68 MCP function.²⁶ Such observations are consistent with another finding: among many shell protein
69 homologs, there are essential asparagine, arginine and lysine residues at the edge of each tile that are
70 important for assembly, indicating a common interaction mechanism (For PduA these residues are K26,
71 N29, and R79).^{4, 27} In addition, shell assembly does not always depend on cargo loading. The protein shell
72 can form when no cargo is expressed, and foreign proteins may also be encapsulated, or localized to the
73 interior of the shell.^{28–31} In fact, peptide sequences that direct native enzymes to the interior of the Eut

74 MCP also function to encapsulate proteins within the Pdu MCP.³² Thus, it may be possible to bring Pdu
75 and Eut shell proteins together to form chimeric MCP assemblies with new properties.

76 In this work, we explore two strategies for altering MCP function: mutations to the pore residues
77 and generation of chimeric MCPs. We study the well-characterized *S. enterica* Pdu MCP, with a focus on
78 its PduA shell protein and its paralog EutM, also from *S. enterica*. We find that EutM behaves as a PduA
79 substitute in the Pdu MCP, and we describe the phenotypic changes resulting from EutM substitution. A
80 $\Delta pduA::eutM$ chromosomally modified strain incorporates EutM into the Pdu MCP in place of PduA.
81 Interestingly, the $\Delta pduA::eutM$ strain exhibits improved growth on 1,2-PD over cells expressing the
82 native Pdu MCP, and this new phenotype is recapitulated by a single amino acid mutation in the PduA
83 pore, PduA^{K37Q}. A saturation mutagenesis of position 37 of PduA reveals that the pore is highly amenable
84 to mutation, and that the hydrophobicity and charge of the amino acid at this locus are the most important
85 factors affecting growth on 1,2-PD. This study thus demonstrates the potential for success in altering
86 metabolite transport properties with both MCP engineering strategies, enabling MCP design and
87 optimization for synthetic biology applications.

88 **RESULTS AND DISCUSSION**

89 **The pore of PduA is different from the pore of EutM**

90 PduA forms one of three canonical shell protein types: a hexameric tile complex containing a
91 central pore approximately 6 Å in diameter.²² The residues defining the shape and properties of the pore
92 include those contained in a loop at the hexamer symmetry axis with sequence Lys-Ile-Gly-Ser-Gly-Leu.
93 The importance of the K37 and S40 loci of this sequence was previously demonstrated; certain mutations
94 to either of these residues result in diminished growth of *S. enterica* LT2 on 1,2-PD as a sole carbon
95 source.^{23, 27} EutM, a close PduA homolog from the Eut operon, also forms a hexamer with nearly identical
96 predicted structure (RMSD EutM/PduA = 1.004 Å, PDB ID 3MPY (EutM) and 3NGK (PduA)) and a
97 highly similar amino acid sequence (67% identity).^{20, 22} Refer to the protein sequence alignment for PduA

98 and EutM (Figure S1). Key differences between PduA and EutM are found in the pore, particularly at the
99 residues aligning with PduA K37 and S40. The EutM pore loop has the sequence Gln-Ile-Gly-Gly-Gly-
100 Leu (Figure 1b, c). In Figure 1c, the pore-lining lysine residue in PduA and the glutamine in EutM are
101 highlighted in green.

102 **The growth of *S. enterica* on 1,2PD is improved in $\Delta pduA::eutM$ strain**

103 We substituted *eutM* for *pduA* to test whether the structural similarity of *eutM* is sufficient to
104 complement the deletion of *pduA* and maintain proper Pdu MCP assembly. We also wanted to learn if
105 small differences between the pore structures would manifest in significant differences in phenotypes
106 attributed to metabolite diffusion. In order to minimize disruption of the native regulation of the Pdu
107 operon, we constructed a genomic integration of *eutM* at the *pduA* locus (Figure 1d). To test MCP
108 formation in this $\Delta pduA::eutM$ strain, the expression of the Pdu operon was induced in minimal NCE
109 media containing 55 mM 1,2-PD as the sole carbon source, and 150 nM coenzyme B₁₂. Growth of the
110 $\Delta pduA::eutM$ strain in this medium is an indicator of metabolite diffusion through the MCP structure. For
111 instance, a strain harboring a *pduA* deletion without a compensating substitution exhibits decreased cell
112 growth due to increased shell permeability to the toxic intermediate propionaldehyde, despite producing a
113 seemingly closed MCP that may be purified and appears to contain all other necessary shell proteins and
114 enzymes.^{21, 27} Furthermore, occlusion or other changes to the architecture of the pores of the MCP shell
115 also lead to growth defects due to restriction of diffusion of the substrate 1,2-PD.⁷ Our results show this is
116 not the case for the $\Delta pduA::eutM$ strain. Not only did the strain harboring the substitution grow without a
117 noticeable defect on 1,2-PD, we observed improved growth over wild type under these growth conditions
118 (Figure 3, Table 1). To control for changes in cell morphology, we also quantified colony forming units at
119 various growth time points to confirm improved growth of the $\Delta pduA::eutM$ strain (Figure S2). The
120 $\Delta pduA::cat/sacb$ mutant strain was used as a control for both growth measurements. This strain grew
121 poorly compared to the wild-type (Figure 3, Figure S2). Moreover, no gross changes to cell morphology
122 or physiology were observed in the $\Delta pduA::eutM$ strain by phase contrast (Figure S3).

123 **EutM is localized to the Pdu MCP in the *S. enterica* $\Delta pduA::eutM$ strain**

124 We next set out to demonstrate that EutM is incorporated into the shell, and to determine if the
125 formation of the shell was altered by this incorporation. We found that substituting *eutM* for *pduA*
126 allowed the formation of chimeric MCPs that were morphologically similar to the wildtype, as assessed
127 by a number of assays. First, we targeted green fluorescent protein (GFP) for localization within the
128 MCPs produced in each of the variants and analyzed assembly formation *in vivo* using fluorescence
129 microscopy. GFP was expressed as a fusion to the PduD¹⁻²⁰ encapsulation tag and an SsrA degradation tag
130 such that only encapsulated GFP remains protected from proteolysis by the ClpXP machinery.^{31, 33, 34} This
131 results in bright puncta by fluorescence microscopy. Fluorescent puncta were observed in both
132 $\Delta pduA::eutM$ and wild-type strains, confirming MCP formation and protein encapsulation *in vivo* (Figure
133 2a, e, b, f, and Figure S4). Automated image analysis by Cell Profiler yielded a count of the number of
134 puncta per cell and indicated that the number of chimeric MCPs per cell is similar to the number of wild-
135 type MCPs per cell (Figure S5).³⁵ Variability in puncta brightness is often observed between cells as well
136 as between puncta within a cell. While we have not investigated the cause of this feature, it appears
137 consistent across both the wild-type and $\Delta pduA::eutM$ strains.

138 We next purified MCPs by sedimentation to compare the morphology of chimeric MCPs to wild-
139 type *in vitro*. Transmission electron microscopy (TEM) analysis showed that MCPs purified from the
140 $\Delta pduA::eutM$ strain are morphologically similar to wild-type MCPs (Figure 2i, j, and Figure S4). An
141 analysis of MCP diameter demonstrated that MCPs purified from the $\Delta pduA::eutM$ strain have a similar
142 mean diameter to the wild-type but with a somewhat wider distribution of sizes (Figure S6).

143 The protein content of purified MCP samples was analyzed by densitometry measurements of
144 coomassie-stained SDS-PAGE gels. Prominent bands consistently observed in MCP samples were
145 previously identified via mass spectrometry.¹⁹ Those identifications are used here. MCPs purified from
146 the $\Delta pduA::eutM$ strain exhibited shell protein bands in similar proportions to those found in wild-type
147 MCPs, except that the PduA band was missing and the 13 kDa band expected to be PduU roughly

148 doubled in density (Figures 2m, S4, S7). We hypothesized that this band represented both PduU and
149 EutM superimposed. To verify the localization of EutM, the purification was repeated with a FLAG
150 epitope tag fused to EutM. Western blotting indicated that EutM-FLAG runs at the apparent size of 13
151 kDa in a 12.5% acrylamide gel, and we confirmed that the 13 kDa band included EutM-FLAG by mass
152 spectrometry (Figures 2n, S8, Table S1).

153 Next, we constructed a non-polar $\Delta pduA::pduA-FLAG$ strain. We purified MCPs from this strain
154 along with the $\Delta pduA::eutM-FLAG$ strain and compared the levels of FLAG-tagged proteins using SDS-
155 PAGE and western blot. Loading of the purified MCP samples was normalized by total protein
156 concentration. We found that all samples contained similar amounts of PduA- or EutM-FLAG (Figure
157 S9).

158 To confirm that EutM is not merely co-purifying with Pdu MCPs, anti-FLAG immuno-gold
159 antibodies were used to label Pdu MCPs purified from the $\Delta pduA::eutM-FLAG$ strain. Immunostained
160 TEM images show that the anti-FLAG immuno-gold antibodies bind well-formed MCP structures. In
161 contrast, the wild-type MCP sample does not bind any gold particles (Figure 2 i, j). Taken together, all
162 evidence supports our hypothesis that EutM properly incorporates into the Pdu MCP shell and causes no
163 morphological defects.

164 **PduA^{K26A} and EutM^{K24A} mutations both abrogate MCP formation**

165 Sinha *et al.* demonstrated that a PduA^{K26A} mutant does not form MCPs due to disruption of
166 hexamer edge binding contacts, resulting in a growth defect on 1,2-PD even though the enzymes required
167 to metabolize 1,2-PD are present.²⁷ Also, MCP structures cannot be purified from the $pduA^{K26A}$ strain via
168 established methods for purifying the Pdu MCP. We replicated these results (Figures 2k, S10, 3), and in
169 addition used fluorescence microscopy with our encapsulation reporter PduD¹⁻²⁰-GFP-SsrA to observe
170 this mutant *in vivo*. We did not observe an absence of fluorescence as we expected; instead, we saw cells
171 linked by filaments resembling the nanotube or filament structures formed by overexpression of single

172 shell proteins or incomplete MCP operons as described previously (Figure 2c, g).^{4, 28, 36, 37} These structures
173 were accompanied by bright puncta varying in appearance and predominantly found at the cell poles.
174 These puncta are presumed to be aggregated or poorly formed MCPs since they cannot be purified in the
175 same manner as non-mutated MCPs. Attempts to purify MCPs from this strain gave only dilute aggregate
176 species as seen by SDS-PAGE analysis as well as by TEM (Figure 2k, Figure S10). The same phenotype
177 was observed when a similar mutation was made to *eutM* in the $\Delta pduA::eutM$ strain (Figure 2d, h, i, and
178 Figure S10). Growth on 1,2-PD is decreased for a $\Delta pduA::eutM^{K24A}$ strain, likely due to mis-formed
179 MCPs like those formed in the $pduA^{K26A}$ strain (Figure 3). $\Delta pduA::eutM^{K24A}$ and $pduA^{K26A}$ mutants were
180 used as non-MCP forming controls in subsequent experiments. Their behaviors coincide in all tests. This
181 is further evidence that EutM is able to complement PduA and fully integrate into the Pdu MCP.
182 Moreover, our results indicate that the same lysine binding partner is important for assembly of both the
183 wild-type Pdu MCP and the EutM Pdu MCP chimera.

184 **Limiting coenzyme B₁₂ condition confirms $\Delta pduA::eutM$ MCPs form a complete compartment shell**

185 Coenzyme B₁₂ is a required cofactor for the first step of 1,2-PD metabolism in the Pdu pathway.
186 This cofactor is supplied in the growth medium because *S. enterica* cannot produce it under our aerobic
187 conditions. Growth at low coenzyme B₁₂ concentrations (20 nM) has been used previously as a method to
188 sensitively detect whether mutations to MCP proteins result in non-assembly or defective MCP shells.^{21,}
189 ^{23, 27, 38} The limiting coenzyme B₁₂ condition is advantageous for non-assembling MCP mutants, such as
190 the $pduA^{K26A}$ mutant, and allows improved growth. Conversely, the wild type and mutant strains that
191 assemble well grow poorly at 20 nM coenzyme B₁₂ compared to mutants that do not assemble MCPs. We
192 exploited this growth condition as an additional assay to detect if the mutations we made to PduA
193 disrupted MCP assembly. We found that the $\Delta pduA::eutM$ mutant grows as poorly as the wild-type strain
194 in this condition (Figure S11), indicating that the chimeric MCP shells form a complete diffusion barrier.

195 **The effects of EutM substitution are recapitulated by a point mutation in the PduA pore**

196 The above experiments showed that there were no observable differences in shell protein
197 expression or MCP assembly between wild-type and $\Delta pduA::eutM$ strains. Thus, we set out to investigate
198 whether the behavior of the EutM-Pdu chimera is indeed due to a change in diffusion properties of the
199 shell. We hypothesized that the difference in growth caused by the EutM substitution was the result of a
200 change in diffusion through the central pore of the shell protein. To support this idea, we incorporated a
201 chromosomal point mutation to encode PduA^{K37Q}, to more closely mimic the EutM pore sequence.

202 The analyses performed for the $\Delta pduA::eutM$ substitution mutant were repeated with $pduA^{K37Q}$.
203 Strikingly, the $pduA^{K37Q}$ mutation confers improved growth (Figure 3, S2) over wild type while still
204 maintaining normal MCP morphology and protein content (Figures 2m, S5, S6, S7, S9, S12, S13). A
205 single point mutation to the pore causes the same effect as a complete shell protein substitution. This is
206 strong evidence indicating that diffusion of molecules through the pore of the MCP shell is altered,
207 causing the observed changes in growth and metabolite levels. A modification of lysine to glutamine
208 changes the charge environment in the pore, and may also change the polarity or hydrogen bonding
209 configuration. Both of these factors likely alter the diffusion barrier small molecules experience.^{23, 39} It is
210 convenient for MCP engineering purposes that such small mutations to this region can confer improved
211 function of the whole organelle structure.

212 **The EutM substitution and the PduA^{K37Q} pore mutant change the accumulation of downstream** 213 **metabolites**

214 HPLC analysis was performed to measure metabolite levels in the media over the course of
215 growth on 1,2-PD (Figure 4). Fast growth on 1,2-PD has previously been associated with decreased
216 aldehyde levels in the media.¹⁰ Therefore, we expected that the $\Delta pduA::eutM$ and $pduA^{K37Q}$ strains would
217 accumulate less aldehyde in the media. Indeed, we found that the accumulation of downstream products
218 propionaldehyde, 1-propanol, and propionate are decreased in both the $\Delta pduA::eutM$ and $pduA^{K37Q}$
219 mutants. We also found that the consumption of 1,2-PD occurred at roughly the same rate in wild-type,
220 $\Delta pduA::eutM$, and $pduA^{K37Q}$ strains (the differences in calculated rate are not statistically significant).

221 Non-assembling mutants (*pduA*^{K26A} and Δ *pduA::eutM*^{K24A}) show the highest levels of
222 propionaldehyde, at 8.6 and 10 mM, respectively, compared to 2.1 mM for the wild-type MCP (Figure 5).
223 Propionaldehyde is toxic to *S. enterica* when exogenously added at concentrations of about 8 mM.¹⁴ Also,
224 since propionaldehyde is volatile, accumulation of this molecule in the media results in some unmeasured
225 amount of propionaldehyde evaporation to the head space. The differences in propionaldehyde
226 concentrations among wild type, Δ *pduA::eutM*, and *pduA*^{K37Q} are more subtle. Wild-type *S. enterica*
227 maintains about 2 mM propionaldehyde for much of the exponential phase. We observed no detectable
228 propionaldehyde in Δ *pduA::eutM* cultures apart from a single spike usually occurring between 27 and 30
229 hours reaching a level averaging 2 mM. The *pduA*^{K37Q} strain behaved similarly to the Δ *pduA::eutM* strain
230 with respect to propionaldehyde accumulation. Therefore, while these strains reach similar maximum
231 concentrations of propionaldehyde in the media, the Δ *pduA::eutM*, and *pduA*^{K37Q} mutants sustain this
232 concentration for a shorter period.

233 In wild-type cultures the 1-propanol concentration level stabilizes at around 25 mM from 36
234 hours onward. In contrast, Δ *pduA::eutM* and *pduA*^{K37Q} strains reach a maximum of about 20 mM 1-
235 propanol, and the concentration decreases after 36 hours. This indicates that the accumulation rate of 1-
236 propanol is slower in the Δ *pduA::eutM* and *pduA*^{K37Q} mutants. Non-assembling mutants accumulate even
237 less 1-propanol, likely due to the greatly increased escape of the aldehyde before conversion. After the
238 supply of 1,2-PD is exhausted, 1-propanol in the media is likely converted back to propionaldehyde and
239 then consumed.²³ This process in addition to some evaporation from the headspace results in the decline
240 in 1-propanol concentration at the end of each growth period. A surprising trend is seen in the
241 concentration of propionate observed in the media over time. We observed increased propionate
242 accumulation in strains that grew poorly. Non-assembling mutants ranked highest in propionate
243 accumulation, followed by the wild type, then *pduA*^{K37Q} and Δ *pduA::eutM*.

244 To offer an explanation for these behaviors, we propose that the Δ *pduA::eutM* and *pduA*^{K37Q}
245 mutants restrict the diffusion of intermediate propionaldehyde, the negatively charged propionyl-

246 phosphate product, or both. In the following discussion, we consider how this hypothesis is supported by
247 the observed metabolite concentrations and growth phenotype.

248 The restriction of propionaldehyde diffusion by the $\Delta pduA::eutM$ and $pduA^{K37Q}$ mutants would
249 directly result in the observed decreases in propionaldehyde accumulation compared to the wild-type.
250 Since propionaldehyde is volatile, increased accumulation in the media leads not only to toxicity, but also
251 to a loss of carbon to the gas phase. Preventing aldehyde escape may also result in a higher local
252 propionaldehyde concentration and a kinetic enhancement of pathway throughput. Both of these effects
253 contribute to the differences in final cell density observed between strains, leading to fastest growth in the
254 $\Delta pduA::eutM$ and $pduA^{K37Q}$ mutants.

255 Restriction of propionyl-phosphate diffusion could result in the decreased accumulation of both
256 propionate and 1-propanol by the $\Delta pduA::eutM$ and $pduA^{K37Q}$ mutants. To illustrate, non-assembling
257 mutants allow free diffusion of propionyl-phosphate, facilitating its conversion to propionate by the
258 PduW kinase, which is not localized to the MCP. In contrast, propionyl-phosphate molecules face a
259 moderate diffusion barrier in the wild-type pore, resulting in decreased propionate release compared to
260 non-assembling mutants. Thus, the $\Delta pduA::eutM$ and $pduA^{K37Q}$ mutants may form a greater diffusion
261 barrier since even less propionate is released than in the wild-type strain. The mitigation of 1-propanol
262 accumulation in $\Delta pduA::eutM$ and $pduA^{K37Q}$ mutants might also be due to restricted propionyl-phosphate
263 diffusion. Decreased product escape could cause a pathway back-up reducing the rate of NADH
264 production, leading to slower turnover in the 1-propanol branch of the pathway involved in NADH
265 recycling. Considering pore structure, it is intuitive that a positively charged lysine in the wild-type PduA
266 pore would facilitate the interaction of a negative product such as propionyl-phosphate more than the
267 neutral glutamine, and allow a greater rate of egress.^{39,40} Propionate escape in the wild-type could lead
268 directly to decreased growth rate, or slower propionate uptake into central metabolism could be a result of
269 aldehyde toxicity.

270 We recognize that there is still much to learn about the fate of propionate and the relationship
271 between the 1,2-PD utilization pathway and the rest of central metabolism. According to our current
272 understanding, the effect of decreased propionaldehyde leakage should result in flux enhancement.
273 However, though strains that appear to restrict aldehyde leakage grow better, they also release less of the
274 downstream products, propionate and 1-propanol. Further study is necessary resolve these two effects.

275 **Shell protein pore residues can be altered to control encapsulated enzyme function**

276 We further explored the relationship between the pore and 1,2-PD metabolism by creating a
277 genomically-encoded library of PduA variants that differ at residue 37. We checked each mutant for
278 MCP formation by fluorescence microscopy using our encapsulation reporter PduD¹⁻²⁰-GFP-ssrA (Figure
279 S12). All mutants exhibited fluorescent puncta, indicating the formation of MCPs. Growth on 1,2-PD
280 with a limiting concentration of coenzyme B12 further demonstrated that all mutants did not form “leaky”
281 MCP structures (Figure S11). Finally, we selected a few of the library members for purification. Purified
282 MCPs from these strains are all similar to wild-type MCPs when examined by SDS-PAGE banding
283 pattern and by TEM. (Figure S10 and S13).

284 The library was then screened for growth in NCE with 1,2-PD and 150 nM coenzyme B12
285 (Figure S14). The end-point cell density for each mutant is summarized in Figure 6a. To analyze these
286 data we compared average end-point cell density to the characteristics of the amino acid substituted in
287 each mutant. The variables considered were the predicted amino acid pK_a, hydrophathy index, and residue
288 molecular weight. We found a moderate negative correlation between hydrophathy index and final cell
289 density (correlation coefficient of -0.47), indicating more hydrophobic side chains generally caused a
290 decrease in growth (Figure 6c). Final cell density was also plotted versus pK_a for several charged amino
291 acid side chains. There was a negative correlation (correlation coefficient of -0.6) suggesting that a more
292 positively charged pore is detrimental to growth on 1,2-PD (Figure 6b). There was no clear trend when
293 growth was compared to size of the amino acid side chain alone (Figure S15). This indicates that, of the

294 three characteristics considered, the hydrophobicity and charge of the amino acids in the pore were most
295 important for retention of small molecules like propionaldehyde within the MCP.

296 **Pore residues of PduA evolve rapidly**

297 Upon analysis of the PduA K37 mutagenesis library, we were surprised to observe that the pore
298 loop of PduA is very tolerant to mutation; indeed, all mutants formed MCPs similar to wild type. It is
299 remarkable that this locus of PduA can accommodate any amino acid without a compensating mutation.
300 For future engineering of the MCP shell, it is desirable to know if other residues may be mutated to
301 change MCP function or to add chemical handles for the conjugation of useful moieties. We therefore
302 used the Rate4Site program (version 2.01) to calculate an evolutionary rate scaling factor for each residue
303 of PduA relative to the overall evolution rate of the protein.⁴¹ The input of this program is a multiple
304 sequence alignment of 193 non-redundant PduA homologs from a variety of MCP systems. The resulting
305 output is mapped over the structure of PduA (Figure 7a). Residues of a darker red color have high
306 evolutionary rates, while white residues evolve slowly relative to the average rate for the entire protein.
307 The evolutionary rate parameter represents the number of non-synonymous replacements expected per
308 unit time for a specific residue. This parameter is related to conservation score but the rate calculation
309 also uses the architecture and branch lengths of the phylogenetic tree in order to account for evolutionary
310 time. If we detect residues that change often over time, but greatly affect the protein function, this could
311 indicate a structural feature important for adapting function to environmental change.

312 The results indicate that pore positions 37 and 40 have a higher evolutionary rate than the average
313 evolution rate for the entire protein, while the isoleucine and glycine at positions 38 and 39 evolve more
314 slowly than the protein average. This suggests that a hydrophobic motif at the narrowest part of the pore
315 is essential for MCP function in general, whichever pathway is encapsulated within. On the other hand,
316 while the residues forming rings on either side of the hydrophobic motif evolve more quickly, they
317 evidently also affect encapsulated pathway function, most likely by inhibiting or facilitating diffusion of

318 one or more metabolites. The mutability of certain pore residues may allow the flexibility to encapsulate a
319 number of pathways involving a diversity of small molecules.

320 Interestingly, the residues K90-I92 near the C-terminus of PduA, as well as several other residues
321 on both concave and convex surfaces, are also very rapidly evolving (labeled on Figure 7a). These areas
322 of the protein structure are less constrained and may serve as good targets for protein modifications,
323 particularly for control of cargo loading or for labeling the outer surface for engineering applications.

324 **PduA pore mutants conferring faster growth on 1,2-PD are not abundant among PduA homologs**

325 Among the strains in the K37 mutation library, the strains with Gln, Glu, Ser, and Trp at position
326 37 grew to a greater final OD₆₀₀ than the wild-type strain under laboratory conditions. Curiously, this
327 indicates that the *S. enterica* serovar lab strain used in this study is not optimized for flux through the Pdu
328 pathway and other factors may affect the fitness of LT2 in the host gut. We searched among PduA
329 homologs to find out if any of the amino acid variants that resulted in highest growth were among those
330 commonly found in nature. A distribution of the frequencies of each amino acid aligned with position 37
331 is given in Figure 7b. Together, strains with Gln, Glu, Ser, or Trp aligned with position 37 only represent
332 9.38% of over 1000 related sequences. The population of PduA homologs in nature is not enriched for
333 these high performers.

334 **Conclusions**

335 In this work, we demonstrate the ability to assemble Pdu MCPs with a foreign shell protein
336 substitution, and the flexibility to use protein engineering to alter the pore structures of PduA without
337 abrogating MCP structure. For the first time, mutations to the Pdu MCP shell have resulted in enhanced
338 function of the enclosed pathway. PduA shell protein mutants, including those with only a single amino
339 acid substitution of a pore-lining residue, have improved growth phenotypes with 1,2-PD as a carbon
340 source. We uncovered significant evidence that these phenotypes are due to modification of small
341 molecule diffusion rates through the pores of the MCP, leading to escape of the intermediate aldehyde

342 and possibly preventing the escape of the product propionyl-phosphate. These trends are observable
343 through HPLC analysis of metabolites and are supported by our analysis of protein structure and MCP
344 assembly. An investigation of the relationship between amino acid chemistry in the pore and the resulting
345 growth performance indicates that charge and hydrophobicity of the side chain play key roles. Shells with
346 more negatively charged pores tended to grow poorly on 1,2-PD, while an increase in the hydrophobicity of
347 the residue at position 37 also decreased the fitness of the strain. The PduA K37 mutagenesis library
348 produced a range of growth phenotypes with strains that grew on 1,2-PD faster or slower than the wild
349 type. Generation of chimeras and/or modification of shell protein pore residues are both promising
350 techniques to optimize MCP shell characteristics for an encapsulated pathway of interest and to achieve
351 useful phenotypes not found in nature.

352

353 **METHODS**

354 *Bacterial strains, media, and growth conditions*

355 The bacterial strains used in this study are *Salmonella enterica* serovar Typhimurium LT2 and
356 *Escherichia coli* DH10B. For mutations to *pduA*, the gene was amplified from the *S. enterica*
357 chromosome and inserted into a pBAD33, p15a, cm^R plasmid by golden gate assembly with BsaI
358 restriction endonuclease.⁴² This plasmid was transformed into *E. coli*. *pduA* and was mutated as described
359 above using PCR site-directed mutagenesis. *EutM* K24A was created in the same manner. Primers for
360 each mutation are shown in supplemental Table S2.

361 Chromosomal modifications were made by recombineering as described by Court.⁴³ First, the
362 native *pduA* gene was replaced with the *cat/sacb* cassette amplified from the TUC01 genome with
363 primers mfsp105 and either mfsp106 or mfsp195 for *eutM* or *pduA* insertions, respectively (primer
364 sequences listed in Table S2). This amplicon was used to create $\Delta pduA::cat/sacb$. The *eutM* gene was
365 amplified with primers mfsp199 and mfsp200 and used to make $\Delta pduA::eutM$. To make mutations in

366 *pduA*, the gene was first subcloned into a plasmid vector and mutated by site directed mutagenesis as
367 described above. The inserts were then amplified from plasmids by PCR with primers mfsp197 and
368 mfsp104 and used to make all chromosomal *pduA* mutants. For FLAG tag fusions, the non-polar
369 $\Delta pduA::cat/sacb$ strain constructed with mfsp106 was used as a starting point to maintain the same RBS
370 region for the subsequent *pduB* gene. The FLAG sequence was added by PCR to *pduA* and *eutM*
371 sequences using primers mfsp304 and mfsp218, respectively before insertion by recombineering. Each
372 strain was confirmed by Sanger sequencing of PCR amplicons from the PduA region of the *S. enterica*
373 chromosome.

374 For growth of *S. enterica*, single colonies were picked from freshly streaked plates and were
375 grown in 5 mL of TB (Terrific Broth) medium at 30°C, 225 rpm, for 24 hours. For growth curves with a
376 1,2-PD carbon source, cultures were diluted to an OD₆₀₀ of 0.05 in 50 mL No-Carbon-E (NCE) minimal
377 medium supplemented with (unless stated otherwise) 55 mM 1,2-PD, 150 nM coenzyme B₁₂, 1 mM
378 magnesium sulfate, and 50 μM ferric citrate.³² Cultures were grown in 250 mL flasks at 37°C in an orbital
379 shaker at 225 rpm. At every time point, 1.5 mL of cell culture sample was taken for OD₆₀₀ measurement
380 and HPLC analysis. The doubling time of each growth curve was calculated by plotting the log of the
381 OD₆₀₀ vs. time. The linear region of each individual curve was identified (including a minimum of three
382 time points). The slope of the linear regression between time points was determined, and the doubling
383 time was calculated with $t_D = \log(2) / slope$.

384 For MCP purifications, flow cytometry, or microscopy experiments, single colonies were grown
385 in 5 mL LB (Lysogeny broth)-Lennox medium at 30°C, 225 rpm, for 24 hours. Cell cultures were diluted
386 from the primary culture 1:1000 into NCE supplemented with 42 mM sodium succinate, 1mM
387 magnesium sulfate, and 50 μM ferric citrate. For cultures containing the plasmid expressing PduD¹⁻²⁰-
388 GFP-ssrA, half the usual amount of antibiotic was used (17 mg/mL chloramphenicol). Culture volumes
389 were 400 mL of media in 2L flasks for MCP purification, and 5 mL of media in 24-well blocks
390 (Analytical Sales and Services, Inc., cat. no. 24108) for flow cytometry and microscopy. Cells were

391 grown at 37°C in an orbital shaker at 225 rpm. For cultures with PduD¹⁻²⁰-GFP-ssrA expressed from a
392 pBAD33 plasmid, 1.33 mM arabinose was added when subculture reached an OD₆₀₀ of 0.4. After five
393 additional hours of growth, samples were taken for fluorescence microscopy, flow cytometry, or Pdu
394 MCP purification.

395 *High Performance Liquid Chromatography (HPLC) media analysis*

396 At each time point, a 750 µL sample of whole cell culture was taken. This was centrifuged for 5
397 min at 13,000 g to remove cells. The supernatant was decanted and frozen until time of analysis. Before
398 analysis samples were thawed and filtered. The Shimadzu HPLC detection system was equipped with a
399 Bio-Rad Aminex HPX-87H (300 by 7.8 mm) ion exclusion column, LC-20AD solvent delivery system,
400 SIL-20AC auto sampler, RID-10A refractive index detector, and SPD-M20A diode array detector
401 (Shimadzu, Kyoto, Japan). Mobile phase is 5 mM H₂SO₄ with 0.4 mL min⁻¹ flow rate.¹⁴ The column
402 temperature is 35°C.

403 *MCP purification*

404 MCPs were purified by centrifugation of lysate as described previously.⁴⁴ They were stored at 4°C until
405 analysis by SDS-PAGE or TEM. Total protein concentration of pure MCP sample was determined by
406 Pierce BCA protein assay kit (Thermo Fisher).

407 *Transmission electron microscopy*

408 Samples were placed on 400 mesh formvar coated copper grids with a carbon film after the grids were
409 treated by glow discharge. 10 µL of purified MCPs at a concentration of about 0.1 mg/mL was placed on
410 each grid for 2 min. The grids were washed three times with deionized water before fixation. Then, 10 µL
411 of 2% glutaraldehyde in water was placed on each grid for 1 min. Grids were then washed an additional
412 three times with deionized water. Finally, samples were stained in 1.6% aqueous uranyl acetate for 1 min.
413 All solutions were centrifuged directly before use to avoid aggregate or particulate contact with grids.

414 Samples were imaged with an REI Tecnai T12 transmission electron microscope and a Gatan Ultrascan
415 1000 camera (Gatan, Pleasanton, CA).

416 *Immunostaining*

417 Samples were placed on 400 mesh formvar coated copper grids with a carbon film after glow discharge
418 treatment. 10 μ L of purified MCPs at a concentration of about 0.1 mg/mL was placed on grids for 2 min.
419 The grids were washed three times with deionized water. Blocking buffer was prepared with 25 mL of
420 phosphate buffered salt (PBS) at pH 7.4, 0.2 g bovine serum albumin (BSA), and 25 μ L of gelatin from
421 cold water fish skin (Sigma). Then, grids were incubated with blocking buffer for 15 min. Grids were
422 incubated with mouse anti-FLAG primary antibody (Sigma) diluted 1:1000 in the blocking buffer for 2
423 hours. Grids were washed four times with PBS for 2 min each. Grids were then incubated with anti-
424 mouse IgG –gold secondary antibody produced in goat (Sigma) diluted 1:20 in blocking buffer for 1 hour.
425 Grids were washed again four times with PBS for 2 min each time. Samples were fixed with 2%
426 glutaraldehyde in PBS for 5 min. Samples were washed for 2 min in PBS, then four times in deionized
427 water. Grids were stained with 1% uranyl acetate in water for 3 min before grid was allowed to dry.

428 *Mass spectroscopy*

429 Protein bands were excised manually and transferred to Eppendorf tubes (0.2 mL). Protein-containing gel
430 pieces were washed with a series of washes that included alternating additions of 100 μ L 0.1 M
431 ammonium bicarbonate followed by 100 μ L of 100% ACN in order to remove traces of coomassie
432 staining. Once gel slices were thoroughly destained, they were dried using a speed-vac. Gel pieces were
433 then subjected to 10 mM dithiothreitol for 20 minutes at 56 °C and subsequently alkylated with 20 mM
434 iodoacetamide for 1 h at room temperature in the dark. Acetonitrile was then added to shrink the gel
435 pieces and the pieces were thoroughly dried in the speed-vac prior to addition of 0.1 ml of a solution
436 containing 12.5 ng/ μ l trypsin in 50 mM ammonium bicarbonate. Gel pieces were allowed to swell at 4C
437 for 3 hours prior to overnight digestion at 37 °C. Liquid was collected following digestion and the gel

438 pieces were bathed in 100% acetonitrile to adequately shrink the gel pieces. All liquid was collected into
439 one tube and evaporated to dryness using a speed-vac. The peptide solution was reconstituted into 10 μ l
440 of 2% ACN with 0.1% TFA and assessed for peptide concentration based on the absorbance at 280 nm
441 using a Nanodrop spectrophotometer. One microgram total peptide was loaded onto the nano-ACQUITY
442 UPLC™ chromatographic system. Peptides were loaded and separated on a C18 Trizaic Naotile using a
443 60 min RP gradient at 450 nL/min (3–40% ACN over 40 min). The column temperature was set at 45
444 °C. Lock mass (Leucine enkephalin (556.2771 Da), 250 fmol/mL) was constantly infused by the
445 NanoAcquity auxiliary pump at a constant flow rate of 1 μ L/min with lockspray scans set at intervals of
446 45 seconds. The Xevo QToF™ mass spectrometer (Waters) was programmed to switch between low (6
447 eV) and high (18–42 eV) energies in the collision cell, with a scan time of 1 s per function over a mass
448 range of 50–2000 Da. LC-MS^E data were processed with ProteinLynx GlobalServer v2.3 (Waters) and
449 searched in the associated *S. enterica* protein database (UniProtKB/SwissProt Protein Knowledge Base).

450 ***Protein electrophoresis and western blot***

451 Denaturing protein electrophorese (SDS-PAGE) was performed on purified MCPs using a 12.5%
452 polyacrylamide gel unless otherwise noted. A 130V potential was applied for 70 minutes. Unless
453 otherwise noted, equal volumes (15 μ L) of each MCP sample were loaded onto the gel, and densitometry
454 quantifications were normalized by sample protein concentration. Gels were either stained with
455 Coomassie dye or further processed for western blot.⁴⁴ Western blotting was done with a PVDF
456 membrane according to standard protocols. Blotting was done with a mouse anti-FLAG primary antibody
457 (Sigma) 1:2000 dilution in 50 mM Tris 150 mM NaCl pH 7.6 with 0.05% Tween-20 (TBST) with 1%
458 w/v dry milk. The secondary antibody was HRP-conjugated goat anti-mouse antibody (Thermo) diluted
459 1:1000 in TBST. Labeling was visualized with west-pico chemiluminescent substrate (Thermo) using a
460 Bio-Rad ChemiDoc XRS+.

461 Densitometry analysis of Coomassie stained gels was performed using at least three different MCP
462 purifications for each strain. Each purified sample was run on three different gels. Images were analyzed

463 using the Image Lab software (Bio-Rad). The absolute signal measurement for each band was normalized
464 by the signal for the band containing proteins PduT and PduE (band 7 in Figure 2). This band had the
465 least variation in signal over all replicates.

466 *Fluorescence microscopy*

467 Bacteria were viewed using a Nikon Ni-U upright microscope with a 100x 1.45 n.a. plan apochromat
468 objective. Images were captured using an Andor Clara-Lite digital camera and Nikon NIS Elements
469 software. Fluorescence images were collected using a C-FL Endow GFP HYQ bandpass filter. All images
470 were taken with 400 ms exposure and adjusted identically for contrast in Adobe Photoshop software.

471 *Flow cytometry*

472 Cultures were grown as described above. Samples were diluted 1:40 in phosphate-buffered saline (PBS)
473 supplemented with 2 g/L kanamycin (to halt translation) in 96-well plates. 10,000 events were collected
474 for each sample on a Millipore Guava easyCyte 5HT instrument. Cells were distinguished from debris by
475 gating on the forward and side scatter channels using the FlowJo software. Reported fluorescence values
476 are the arithmetic mean of the geometric mean green fluorescence of three independent samples acquired
477 on three different days. Error bars represent one standard deviation.

478 *Multiple Sequence Alignment construction and evolutionary rate analysis*

479 Evolutionary rate of each amino acid locus of PduA was calculated using the Rate4Site program version
480 2.01.⁴¹ The input for this tool was a multiple sequence alignment (MSA) of 193 non-redundant PduA
481 homologs. The homologs were retrieved and the MSA assembled by the jackhammer web server
482 (<https://www.ebi.ac.uk/Tools/hmmer/search/jackhammer>). Rate4Site uses this data to construct a
483 phylogenetic tree. With both MSA and tree information, a rate-scaling factor is calculated that indicates
484 how rapidly each residue evolves relative to the mean protein rate. Default settings (including Bayesian
485 framework) were used here. A larger MSA of about 1000 sequences, also retrieved and aligned with
486 jackhammer, was used to calculate amino acid frequency.

487
488 **SUPPORTING INFORMATION** Supplementary information including tables and figures cited in the
489 text is available free of charge via the Internet at <http://pubs.acs.org>.

490
491 **ABBREVIATIONS**

492 MCP: microcompartment
493 1,2-PD: 1,2-propanediol
494 GFP: green fluorescent protein
495 Pdu: 1,2-propanediol utilization
496 Eut: ethanolamine utilization
497 NCE medium: No-Carbon-E medium
498 TEM: transmission electron microscopy
499 MSA: multiple sequence alignment
500 PBS: phosphate buffered salt

501
502 **AUTHOR INFORMATION**

503 **Corresponding Author**

504 Danielle Tullman-Ercek

505 *E-mail: dtercek@berkeley.edu

506 Department of Chemical and biomolecular Engineering, 201 Gilman Hall, University of California
507 Berkeley 94720, United States

508 **Author Contributions**

509 M.F.S and D.T.E. designed the project. M.F.S. performed the experiments. M.F.S, C.M.J, and D.T.E.
510 wrote the manuscript.

511 **Notes**

512 The authors declare no competing financial interest.

513

514 **ACKNOWLEDGEMENTS**

515 The authors wish to thank Chuchu Zhang for development of the Cell profiler pipeline used in this work,
516 as well as the Robert D. Ogg Electron Microscope Laboratory for equipment usage, training, and
517 materials. The authors would also like to thank the Tullman-Ercek lab for thoughtful discussions and
518 comments. The work was funded by the National Science Foundation (award MCB1150567 to D.T.E.).

519

520

521 **REFERENCES**

522 (1) Bobik, T. a, Havemann, G. D., Busch, R. J., Williams, D. S., and Aldrich, H. C. (1999) The
523 propanediol utilization (pdu) operon of *Salmonella enterica* serovar Typhimurium LT2 includes genes
524 necessary for formation of polyhedral organelles involved in coenzyme B(12)-dependent 1, 2-propanediol
525 degradation. *Journal of bacteriology* 181, 5967–75.

526 (2) Kofoid, E., Rappleye, C., Stojiljkovic, I., and Roth, J. (1999) The 17-gene ethanolamine (eut) operon
527 of *Salmonella typhimurium* encodes five homologues of carboxysome shell proteins. *Journal of*
528 *Bacteriology* 181, 5317–5329.

- 529 (3) Shively, J. M., Ball, F., Brown, D. H., and Saunders, R. E. (1973) Functional organelles in
530 prokaryotes: polyhedral inclusions (carboxysomes) of *Thiobacillus neapolitanus*. *Science* 182, 584–586.
- 531 (4) Heldt, D., Frank, S., Seyedarabi, A., Ladikis, D., Parsons, J. B., Warren, M. J., and Pickersgill, R. W.
532 (2009) Structure of a trimeric bacterial microcompartment shell protein, EtuB, associated with ethanol
533 utilization in *Clostridium kluyveri*. *The Biochemical journal* 423, 199–207.
- 534 (5) Bobik, T. A. (2006) Polyhedral organelles compartmenting bacterial metabolic processes. *Applied*
535 *Microbiology and Biotechnology* 70, 517–525.
- 536 (6) Kerfeld, C. A., Heinhorst, S., and Cannon, G. C. (2010) Bacterial microcompartments. *Annual review*
537 *of microbiology* 64, 391–408.
- 538 (7) Chowdhury, C., Sinha, S., Chun, S., Yeates, T. O., and Bobik, T. A. (2014) Diverse bacterial
539 microcompartment organelles. *Microbiology and Molecular Biology Reviews* 78, 438–468.
- 540 (8) Kerfeld, C. A., and Erbilgin, O. (2015) Bacterial microcompartments and the modular construction of
541 microbial metabolism. *Trends in Microbiology* 23, 22–34.
- 542 (9) Abdul-Rahman, F. (2013) The distribution of polyhedral bacterial microcompartments suggests
543 frequent horizontal transfer and operon reassembly. *Journal of Phylogenetics & Evolutionary Biology* 1,
544 118.
- 545 (10) Jorda, J., Lopez, D., Wheatley, N. M., and Yeates, T. O. (2013) Using comparative genomics to
546 uncover new kinds of protein-based metabolic organelles in bacteria. *Protein Science* 22, 179–195.
- 547 (11) Petit, E., LaTouf, W. G., Coppi, M. V., Warnick, T. a, Currie, D., Romashko, I., Deshpande, S.,
548 Haas, K., Alvelo-Maurosa, J. G., Wardman, C., Schnell, D. J., Leschine, S. B., and Blanchard, J. L.
549 (2013) Involvement of a bacterial microcompartment in the metabolism of fucose and rhamnose by
550 *Clostridium phytofermentans*. *PLoS one* 8, e54337.
- 551 (12) Axen, S. D., Erbilgin, O., and Kerfeld, C. A. (2014) A taxonomy of bacterial microcompartment loci
552 constructed by a novel scoring method. *PLoS Computational Biology* (Tanaka, M. M., Ed.) 10, e1003898.
- 553 (13) Penrod, J. T., and Roth, J. R. (2006) Conserving a volatile metabolite : a role for carboxysome-like
554 organelles in *Salmonella enterica*. *Journal of Bacteriology* 188, 2865–2874.

- 555 (14) Sampson, E. M., and Bobik, T. a. (2008) Microcompartments for B12-dependent 1,2-propanediol
556 degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate.
557 *Journal of bacteriology* 190, 2966–71.
- 558 (15) Huseby, D. L., and Roth, J. R. (2013) Evidence that a metabolic microcompartment contains and
559 recycles private cofactor pools. *Journal of bacteriology* 195, 2864–79.
- 560 (16) Cheng, S., Fan, C., Sinha, S., and Bobik, T. a. (2012) The PduQ enzyme is an alcohol dehydrogenase
561 used to recycle NAD⁺ internally within the Pdu microcompartment of *Salmonella enterica*. *PloS one* 7,
562 e47144.
- 563 (17) Liu, Y., Jorda, J., Yeates, T. O., and Bobik, T. A. (2015) The PduL phosphotransacylase is used to
564 recycle coenzyme A within the Pdu microcompartment. *Journal of Bacteriology* (Parkinson, J. S., Ed.)
565 197, 2392–2399.
- 566 (18) Yeates, T. O., Crowley, C. S., and Tanaka, S. (2010) Bacterial microcompartment organelles: protein
567 shell structure and evolution. *Annual review of biophysics* 39, 185–205.
- 568 (19) Havemann, G. D., and Bobik, T. A. (2003) Protein content of polyhedral organelles involved in
569 coenzyme B12 -dependent degradation of 1,2-propanediol in *Salmonella enterica* Serovar Typhimurium.
570 *Journal of bacteriology* 185, 5086-95.
- 571 (20) Takenoya, M., Nikolakakis, K., and Sagermann, M. (2010) Crystallographic insights into the pore
572 structures and mechanisms of the EutL and EutM shell proteins of the ethanolamine-utilizing
573 microcompartment of *Escherichia coli*. *Journal of bacteriology* 192, 6056–63.
- 574 (21) Cheng, S., Sinha, S., Fan, C., Liu, Y., and Bobik, T. a. (2011) Genetic analysis of the protein shell of
575 the microcompartments involved in coenzyme B12-dependent 1,2-propanediol degradation by
576 *Salmonella*. *Journal of bacteriology* 193, 1385–92.
- 577 (22) Crowley, C. S., Cascio, D., Sawaya, M. R., Kopstein, J. S., Bobik, T. a, and Yeates, T. O. (2010)
578 Structural insight into the mechanisms of transport across the *Salmonella enterica* Pdu microcompartment
579 shell. *The Journal of biological chemistry* 285, 37838–46.

- 580 (23) Chowdhury, C., Chun, S., Pang, A., Sawaya, M. R., Sinha, S., Yeates, T. O., and Bobik, T. A. (2015)
581 Selective molecular transport through the protein shell of a bacterial microcompartment organelle.
582 *Proceedings of the National Academy of Sciences* 112, 2990–2995.
- 583 (24) Dou, Z., Heinhorst, S., Williams, E. B., Murin, C. D., Shively, J. M., and Cannon, G. C. (2008) CO₂
584 fixation kinetics of *Halothiobacillus neapolitanus* mutant carboxysomes lacking carbonic anhydrase
585 suggest the shell acts as a diffusional barrier for CO₂. *Journal of Biological Chemistry* 283, 10377–
586 10384.
- 587 (25) Cai, F., Sutter, M., Bernstein, S. L., Kinney, J. N., and Kerfeld, C. A. (2014) Engineering bacterial
588 microcompartment shells: chimeric shell proteins and chimeric carboxysome shells. *ACS synthetic*
589 *biology* 4, 444–453.
- 590 (26) Sturms, R., Streauslin, N. A., Cheng, S., and Bobik, T. A. (2015) In *Salmonella enterica*,
591 ethanolamine utilization is repressed by 1,2-propanediol to prevent detrimental mixing of components of
592 two different bacterial microcompartments. *Journal of Bacteriology* (Metcalf, W. W., Ed.) 197, 2412–
593 2421.
- 594 (27) Sinha, S., Cheng, S., Sung, Y. W., McNamara, D. E., Sawaya, M. R., Yeates, T. O., and Bobik, T. A.
595 (2014) Alanine scanning mutagenesis identifies an asparagine-arginine-lysine triad essential to assembly
596 of the shell of the pdu microcompartment. *Journal of molecular biology* 426, 2328–45.
- 597 (28) Parsons, J. B., Frank, S., Bhella, D., Liang, M., Prentice, M. B., Mulvihill, D. P., and Warren, M. J.
598 (2010) Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and
599 evidence of filament-associated organelle movement. *Molecular Cell* 38, 305–315.
- 600 (29) Choudhary, S., Quin, M. B., Sanders, M. a, Johnson, E. T., and Schmidt-Dannert, C. (2012)
601 Engineered protein nano-compartments for targeted enzyme localization. *PloS one* 7, e33342.
- 602 (30) Lawrence, A. D., Frank, S., Newnham, S., Lee, M. J., Brown, I. R., Xue, W.-F., Rowe, M. L.,
603 Mulvihill, D. P., Prentice, M. B., Howard, M. J., and Warren, M. J. (2014) Solution structure of a
604 bacterial microcompartment targeting peptide and its application in the construction of an ethanol
605 bioreactor. *ACS Synthetic Biology* 3, 454–465.

- 606 (31) Kim, E. Y., and Tullman-Ercek, D. (2014) A rapid flow cytometry assay for the relative
607 quantification of protein encapsulation into bacterial microcompartments. *Biotechnology Journal* 9, 348–
608 354.
- 609 (32) Jakobson, C. M., Kim, E. Y., Slininger, M. F., Chien, A., and Tullman-Ercek, D. (2015) Localization
610 of proteins to the 1,2-propanediol utilization microcompartment by non-native signal sequences is
611 mediated by a common hydrophobic motif. *Journal of Biological Chemistry* 290, 24519–24533.
- 612 (33) Fan, C., Cheng, S., Sinha, S., and Bobik, T. A. (2012) Interactions between the termini of lumen
613 enzymes and shell proteins mediate enzyme encapsulation into bacterial microcompartments.
614 *Proceedings of the National Academy of Sciences* 109, 14995–15000.
- 615 (34) Savage, D. F., Afonso, B., Chen, A. H., and Silver, P. a. (2010) Spatially ordered dynamics of the
616 bacterial carbon fixation machinery. *Science* 327, 1258–61.
- 617 (35) Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O., Guertin, D. A.,
618 Chang, J. H., Lindquist, R. A., Moffat, J., Golland, P., and Sabatini, D. M. (2006) CellProfiler: image
619 analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* 7, R100.
- 620 (36) Pang, A., Frank, S., Brown, I., Warren, M. J., and Pickersgill, R. W. (2014) Structural insights into
621 higher order assembly and function of the bacterial microcompartment protein PduA. *Journal of*
622 *Biological Chemistry* 289, 22377–22384.
- 623 (37) Noël, C. R., Cai, F., and Kerfeld, C. A. (2016) Purification and characterization of protein nanotubes
624 assembled from a single bacterial microcompartment shell subunit. *Advanced Materials Interfaces* 3,
625 1500295.
- 626 (38) Havemann, G. D., Sampson, E. M., and Bobik, T. A. (2002) PduA Is a shell protein of polyhedral
627 organelles involved in coenzyme B₁₂-dependent degradation of 1,2-propanediol in *Salmonella enterica*
628 Serovar Typhimurium LT2. *Journal of Bacteriology* 184, 1253–1261.
- 629 (39) Glasgow, J. E., Asensio, M. A., Jakobson, C. M., Francis, M. B., and Tullman-Ercek, D. (2015)
630 Influence of electrostatics on small molecule flux through a protein nanoreactor. *ACS Synthetic Biology* 4,
631 1011–1019.

- 632 (40) Welte, W., Nestel, U., Wacker, T., and Diederichs, K. (1995) Structure and function of the porin
 633 channel. *Kidney International* 48, 930–940.
- 634 (41) Mayrose, I. (2004) Comparison of site-specific rate-inference methods for protein sequences:
 635 empirical bayesian methods are superior. *Molecular Biology and Evolution* 21, 1781–1791.
- 636 (42) Engler, C., Gruetzner, R., Kandzia, R., and Marillonnet, S. (2009) Golden gate shuffling: a one-pot
 637 dna shuffling method based on type II restriction enzymes. *PLoS ONE* (Peccoud, J., Ed.) 4, e5553.
- 638 (43) Datta, S., Costantino, N., and Court, D. L. (2006) A set of recombineering plasmids for gram-
 639 negative bacteria. *Gene* 379, 109–115.
- 640 (44) Sinha, S., Cheng, S., Fan, C., and Bobik, T. a. (2012) The PduM protein is a structural component of
 641 the microcompartments involved in coenzyme B(12)-dependent 1,2-propanediol degradation by
 642 *Salmonella enterica*. *Journal of bacteriology* 194, 1912–8.

643

644 **Tables**

645 Table 1. Doubling times of PduA mutants for growth on 1,2-PD. Both $\Delta pduA::eutM$ and $pduA^{K37Q}$ strains
 646 are statistically different than the wildtype strain by two tailed T-test. The $\Delta pduA::eutM$ and $pduA^{K37Q}$
 647 strains are not significantly different from each other by two tailed T-test. n=7 for all cases.

Strain	Doubling time (hr)	Standard error
MFSS044, Wildtype	7.85	±0.89
MFSS252, $\Delta pduA::eutM$	5.73	±0.29
MFSS238, $pduA^{K37Q}$	5.75	±0.49

648

649 Table 2. Bacterial strains used in this study

650

Strain	Organism	Genotype
MFSS044	<i>S. enterica</i> serovar Typhimurium	LT2 wild type
MFSS263	<i>S. enterica</i> serovar Typhimurium	LT2 $pduA^{K26A}$
MFSS236	<i>S. enterica</i> serovar Typhimurium	LT2 $\Delta pduA::cat/sacB$
MFSS252	<i>S. enterica</i> serovar Typhimurium	LT2 $\Delta pduA::eutM$

MFSS273	<i>S. enterica</i> serovar Typhimurium	LT2 $\Delta pduA::eutM$ -FLAG
MFSS238	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37Q}
MFSS393	<i>S. enterica</i> serovar Typhimurium	LT2 $\Delta pduA::pduA$ ^{K37Q} -FLAG
MFSS392	<i>S. enterica</i> serovar Typhimurium	LT2 $\Delta pduA::pduA$ -FLAG
MFSS337	<i>S. enterica</i> serovar Typhimurium	LT2 $\Delta pduA::eutM$ ^{K24A}
MFSS320	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37H}
MFSS321	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37C}
MFSS322	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37R}
MFSS323	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37G}
MFSS324	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37P}
MFSS325	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37V}
MFSS326	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37F}
MFSS327	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37T}
MFSS328	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37L}
MFSS329	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37D}
MFSS330	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37E}
MFSS331	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37S}
MFSS335	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37N}
MFSS336	<i>S. enterica</i> serovar Typhimurium	LT2 <i>pduA</i> ^{K37W}
TUC01 ^a	<i>E. coli</i>	DH10B <i>cat/sacB</i>

651 ^aDatta et al.³⁶

652 **Figure Captions**

653 Figure 1. Organization and modifications of Pdu MCP

654 (a) Diagram illustrating the Pdu pathway including enzymes, cofactors, substrates, intermediates and
655 products. The black hexagonal outline represents the boundary of the MCP. Enzymes outside of this
656 boundary are not localized to the interior of the MCP.³⁹ (b) Schematic demonstrating the hypothesized
657 overall structure of the MCP. Arrows indicate replacement of the PduA shell protein (red) with EutM
658 (yellow). Crystal structures (PDB IDs: 3ngk, 3mpy) to the right illustrate structural similarity.^{14,17} (c) Left
659 hand images are ribbon structures of the pores of PduA (red) and EutM (yellow). Residues PduA K37 and
660 the corresponding glutamine of the EutM pore are highlighted in green. To the right are the surface
661 models of the respective pore regions colored by the electrostatic potential. Blue is positive potential,
662 white is neutral, and red is negative potential. (d) A diagram of the Pdu operon and the operon resulting

663 from recombineering to substitute *eutM* for *pduA* on the *S. enterica* chromosome. Other Pdu shell
664 proteins are blue arrows, enzymes of the pathway and proteins of unknown function are grey arrows.

665

666 Figure 2. EutM incorporates into the Pdu MCP

667 (a, b, c, d) Phase contrast images of *S. enterica* strains LT2 (*wild type*), LT2 $\Delta pduA::eutM$, LT2 $pduA^{K26A}$,
668 and LT2 $\Delta pduA::eutM^{K24A}$, respectively. All strains are co-expressing Pdu MCPs and PduD¹⁻²⁰-GFP-ssrA.
669 (e, f, g, h) GFP fluorescence microscopy images corresponding to phase contrast images a-d. Large
670 filamentous aggregates are indicated with blue arrows in panels g and h. (i, j) Immunostained
671 transmission electron microscopy of MCPs purified from wild type LT2 and LT2 $\Delta pduA::eutM-FLAG$,
672 respectively. Bound IgG-conjugated gold particles are indicated with red arrows. (k, l) Transmission
673 electron microscopy images of MCPs purified from the strains depicted in images c, d. For convenience,
674 larger versions of the images in panels (i-l) are provided in Supplemental Figure S16. (m) Coomassie-
675 stained SDS-PAGE of purified MCPs. Lane (i) is a purification from strain LT2 wild type, lane (ii) is
676 from strain LT2 $pduA^{K37Q}$, lane (iii) is from strain LT2 $\Delta pduA::eutM-FLAG$. (n) α -FLAG western blot of
677 purified compartments.

678

679 Figure 3. Strains harboring shell protein variants grow on 55 mM 1,2-PD and 150 nM coenzyme B₁₂

680 MCP variant strains were grown on 55 mM 1,2-PD with 150 nM coenzyme B₁₂. Wild-type strain LT2 is
681 compared to LT2 $\Delta pduA::eutM$, $pduA^{K37Q}$, $pduA^{K26A}$, $\Delta pduA::eutM^{K24A}$, and $\Delta pduA::cat/sacb$ strains.
682 Mean optical density at 600 nm is plotted vs time. Error bars represent 95% confidence interval from six
683 measurements started on four different days.

684

685 Figure 4. HPLC analysis of MCP-forming strains during growth on 1,2-PD and coenzyme B₁₂

686 The concentrations of 1,2-PD and products of the Pdu pathway were monitored during growth on 55 mM
687 1,2-PD and 150 nM coenzyme B₁₂. Mean values of metabolites are plotted vs time. Metabolites measured
688 are 1,2-PD, 1-propanol, propionate, and propionaldehyde, as labeled. In the left hand column (a, c, e, g)
689 the wild-type *S. enterica* LT2 is compared to LT2 $\Delta pduA::eutM$. In the right hand column (b, d, f, h) the
690 wild-type strain is compared to LT2 *pduA*^{K37Q}. Error bars represent standard error from six measurements
691 started on four different days.

692

693 Figure 5. HPLC analysis of non-assembling shell protein mutant strains

694 The concentrations of 1,2-PD and products of the Pdu pathway were monitored during growth on 55 mM
695 1,2-PD and 150 nM coenzyme B₁₂. Mean values of metabolites are plotted vs time. Metabolites measured
696 are 1,2-PD, 1-propanol, propionate, and propionaldehyde, as labeled. The wild-type LT2 strain is
697 compared to LT2 *pduA*^{K26A} and LT2 $\Delta pduA::eutM$ ^{K24A} strains. Error bars represent standard error from six
698 measurements started on four different days.

699

700 Figure 6. A library of amino acids at residue 37 of PduA confers a variety of growth effects

701 (a) Mean values of final OD (600nm) are plotted for various PduA mutants from growth on 1,2-PD with
702 150 nm coenzyme B₁₂. Error bars represent standard error from four different measurements started on
703 separate days. The red line highlights wild-type growth for comparison. (b) Plot of final OD vs pK_a of
704 charged amino acids substituted at position 37 of PduA. (c) Plot of final OD vs hydropathy index of
705 amino acids substituted at position 37 of PduA.

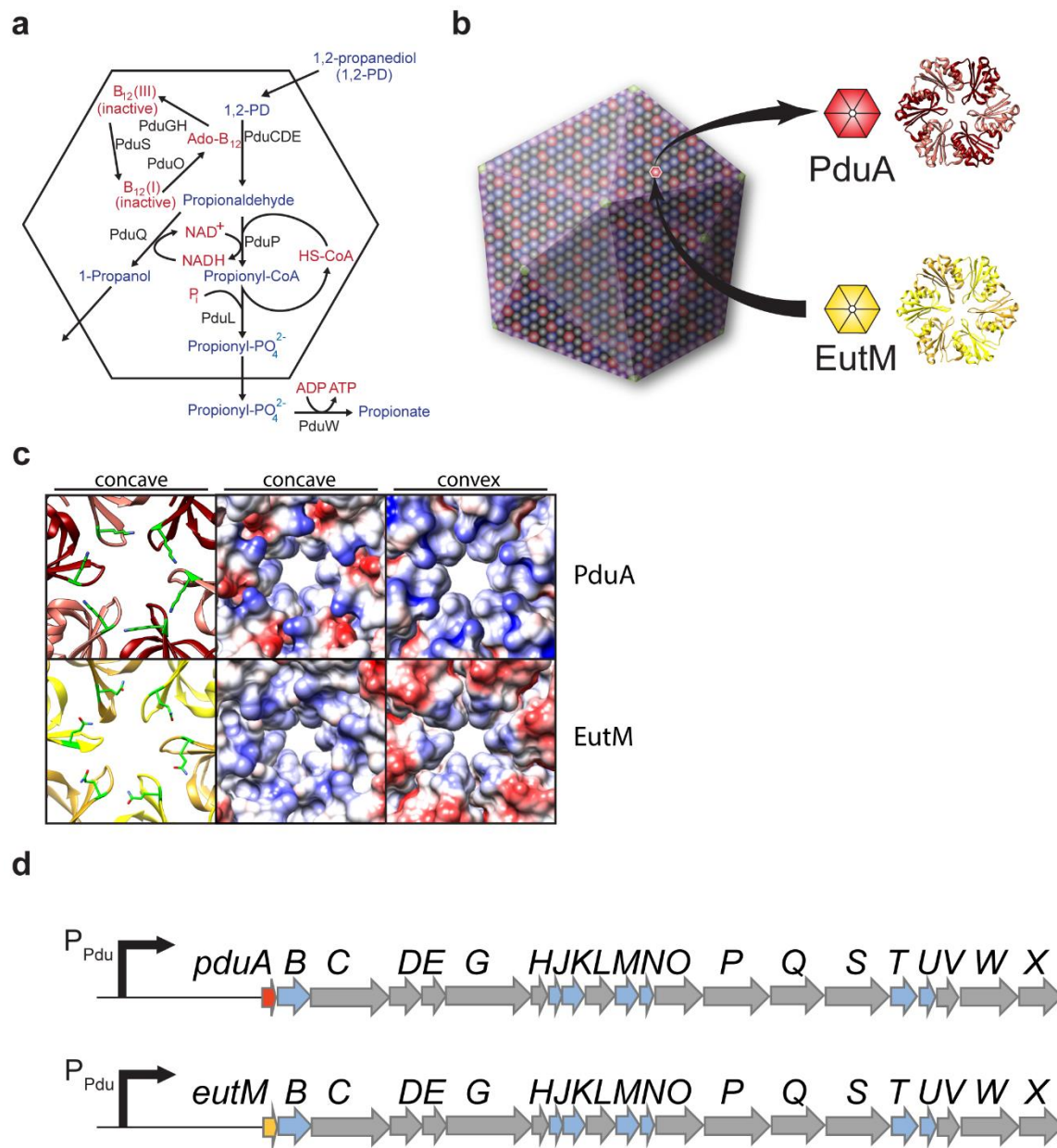
706

707 Figure 7. Phylogenetic analysis of PduA homologs

708 (a) Rate4Site evolutionary rate scaling factor mapped over the crystal structure of PduA. White represents
709 most conserved residues. Dark red represents residues with high evolution rates. Some residues with
710 highest and lowest evolutionary rates are labeled for reference. (b) Histogram showing the % frequency of
711 each amino acid aligning with PduA K37 in an MSA of n=1000 homolog sequences.

712 **Figures**

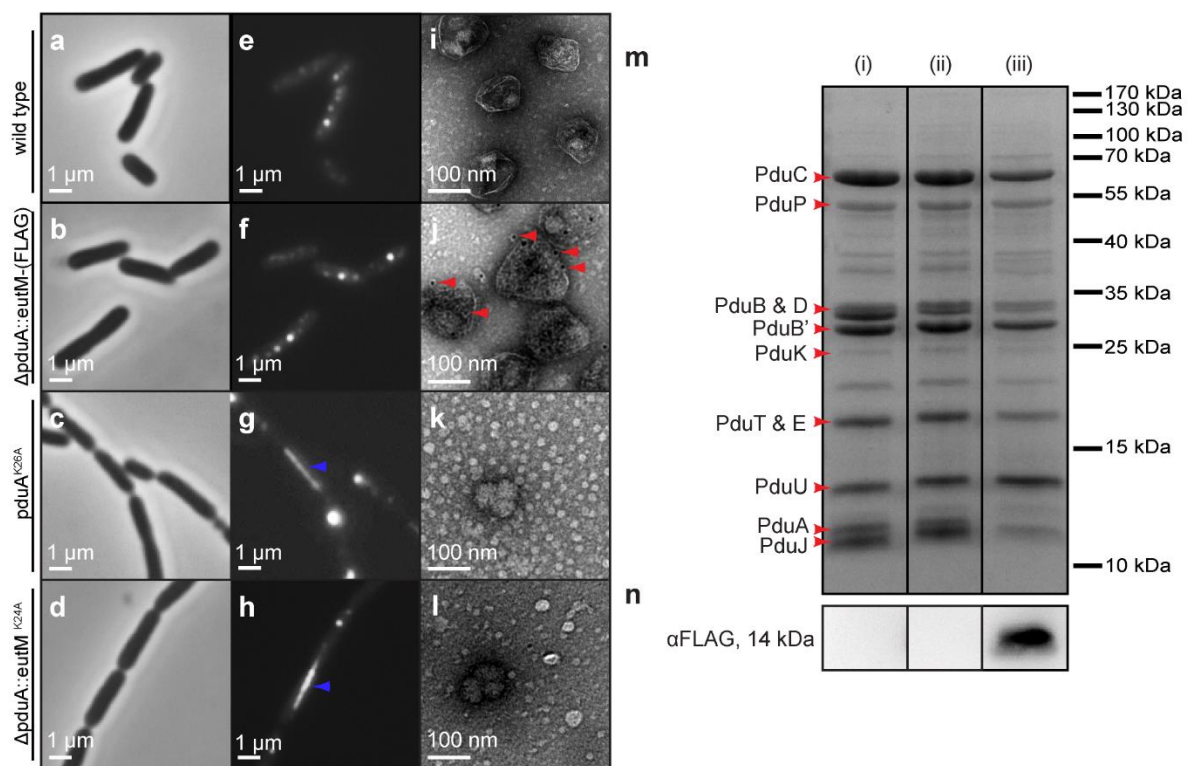
713 Figure 1.



714

715

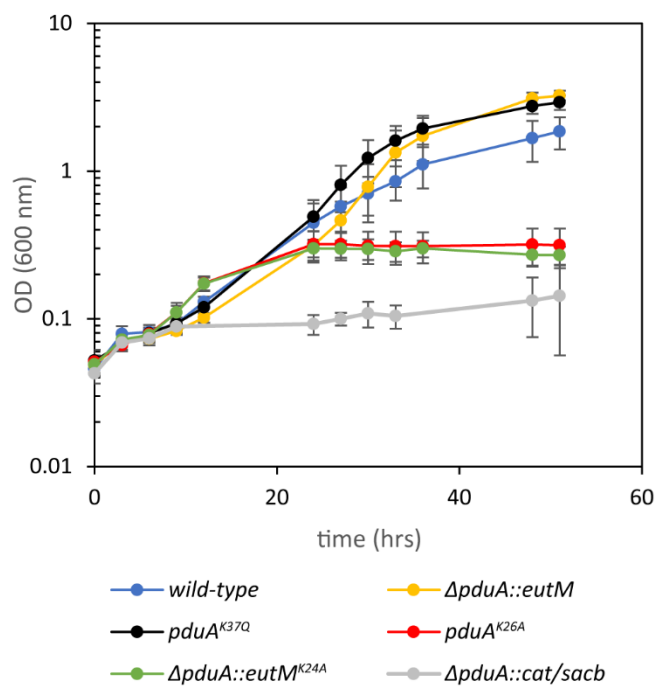
716 Figure 2.



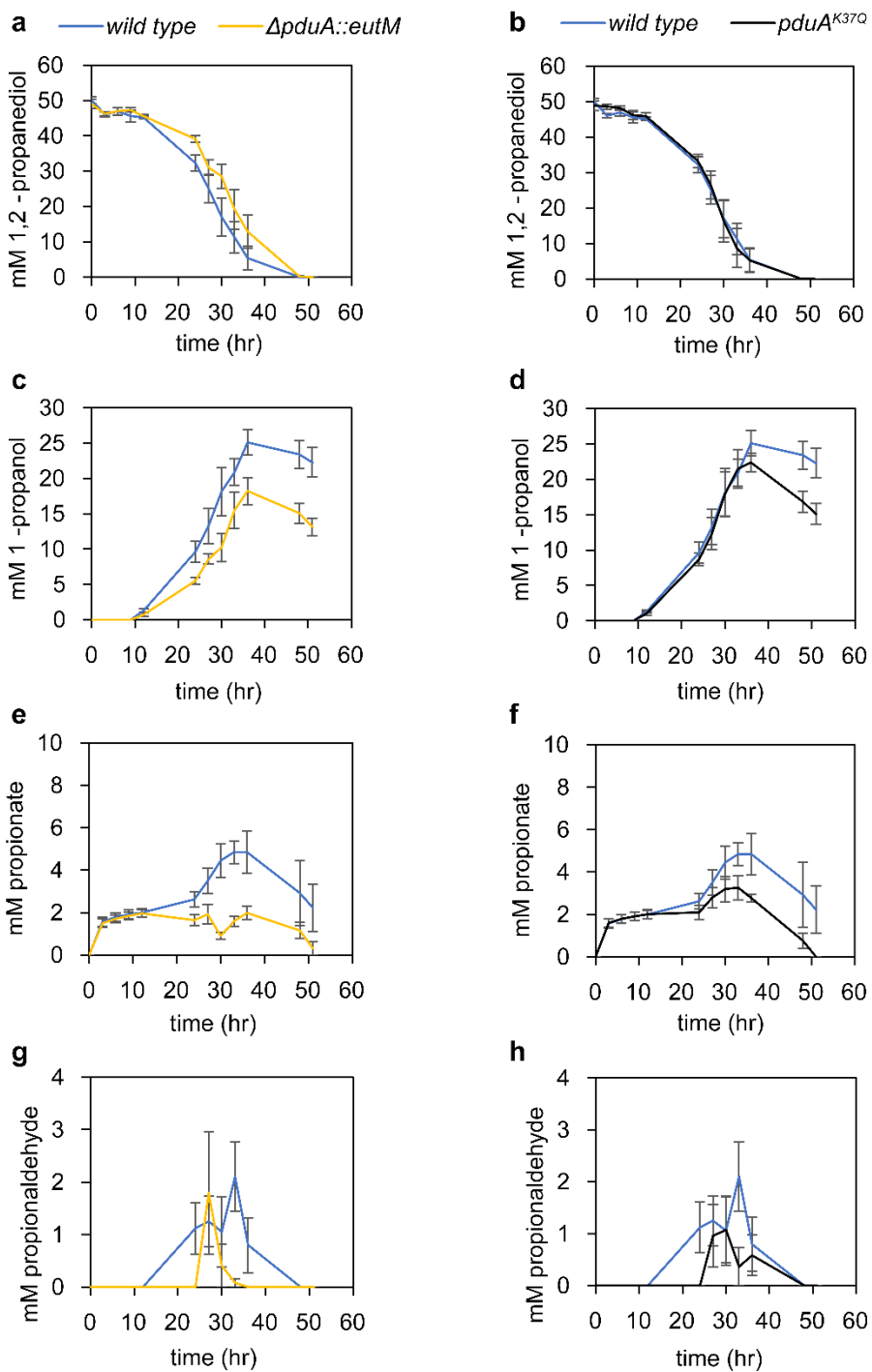
717

718

719 Figure 3.



721 Figure 4.

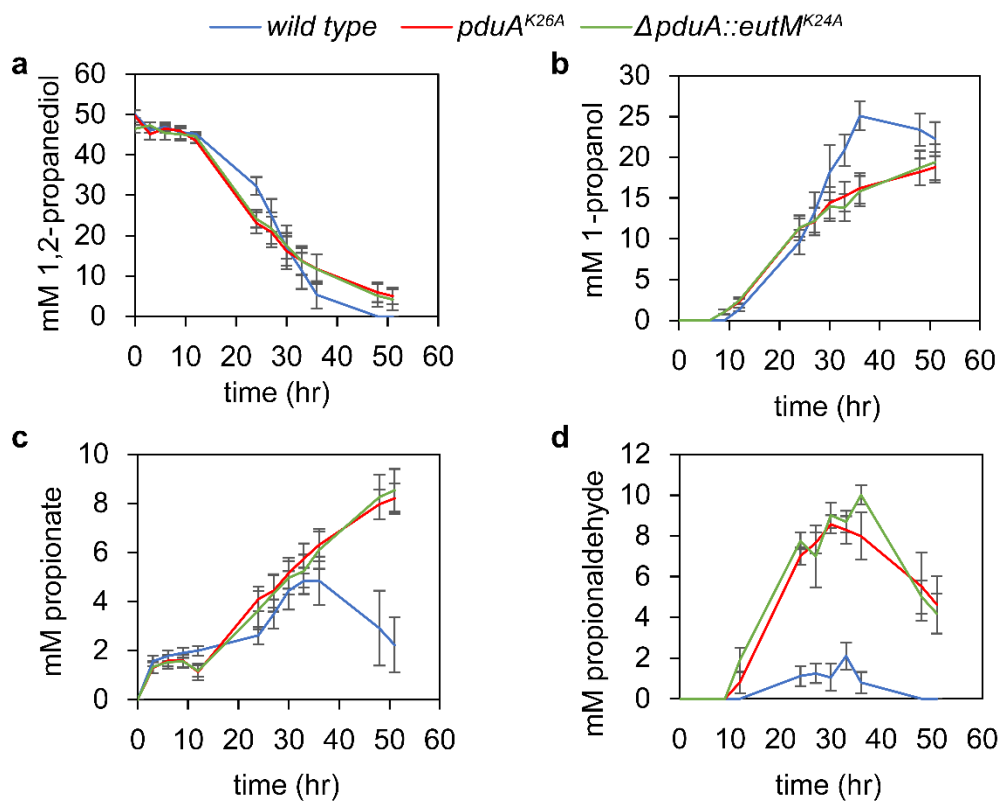


722

723

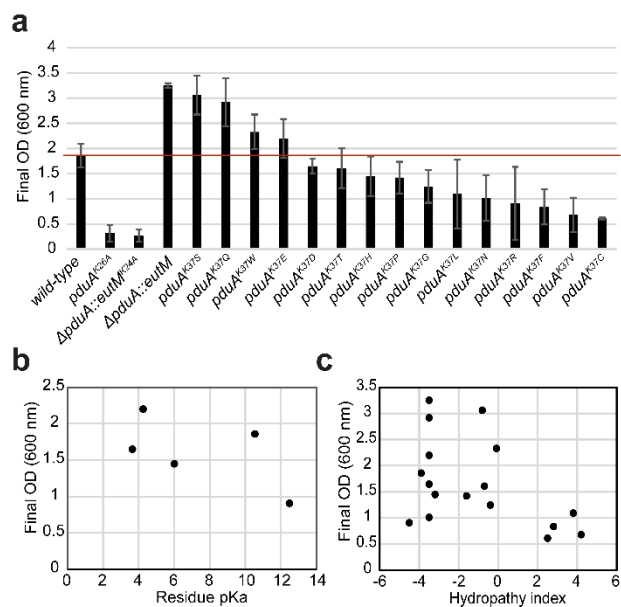
724

725 Figure 5.



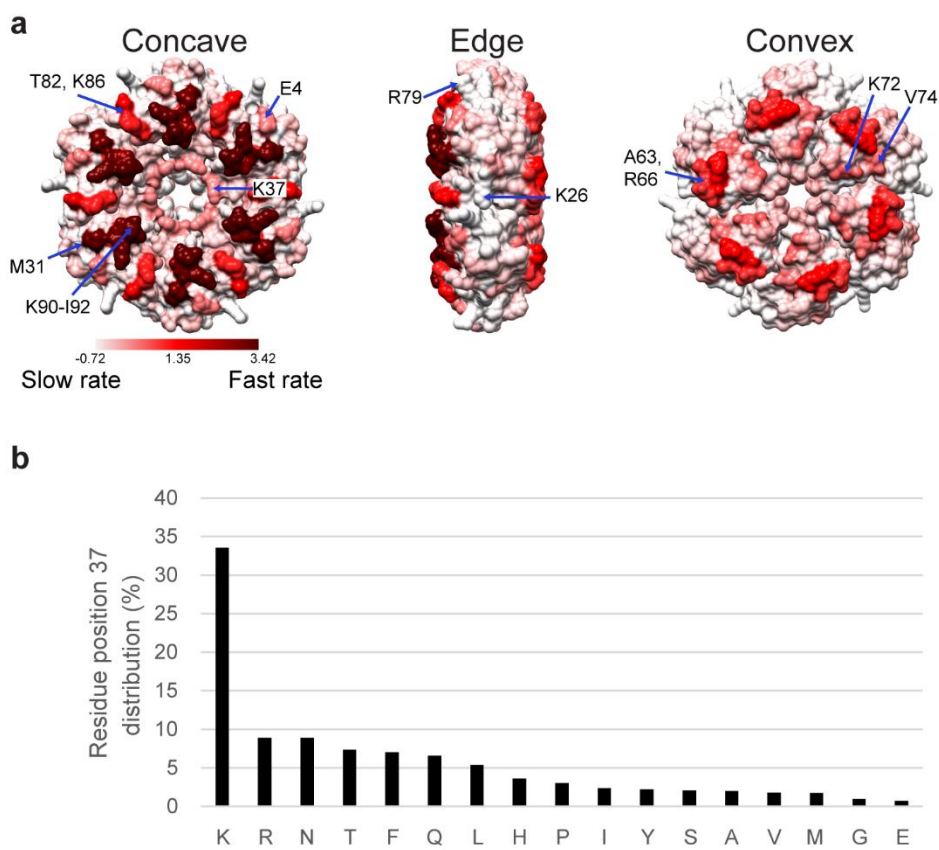
726

727 Figure 6.



728

729 Figure 7.



730