

Engineering and modulating functional cyanobacterial CO₂-fixing organelles

1 Yi Fang¹, Fang Huang¹, Matthew Faulkner¹, Qiuyao Jiang¹, Gregory F. Dykes¹, Mengru Yang¹, and
2 Lu-Ning Liu^{1*}

3 ¹Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB, United
4 Kingdom

5 * **Correspondence:**

6 Dr Lu-Ning Liu

7 luning.liu@liverpool.ac.uk

8 **Keywords:** bacterial microcompartment, carboxysome, carbon fixation, cyanobacteria,
9 encapsulation, Rubisco, self-assembly, synthetic engineering

10

11 Abstract

12 Bacterial microcompartments (BMCs) are proteinaceous organelles widespread among bacterial phyla and
13 provide a means for compartmentalizing specific metabolic pathways. They sequester catalytic enzymes
14 from the cytoplasm, using an icosahedral proteinaceous shell with selective permeability to metabolic
15 molecules and substrates, to enhance metabolic efficiency. Carboxysomes were the first BMCs discovered
16 and their unprecedented capacity of CO₂ fixation allows cyanobacteria to make a significant contribution
17 to global carbon fixation. There is an increasing interest in utilizing synthetic biology to construct
18 synthetic carboxysomes in new hosts, i.e. higher plants, to enhance carbon fixation and productivity. Here,
19 we report the construction of a synthetic operon of β -carboxysomes from the cyanobacterium
20 *Synechococcus elongatus* PCC7942 to generate functional β -carboxysome-like structures in *Escherichia*
21 *coli*. The protein expression, structure, assembly, and activity of synthetic β -carboxysomes were
22 characterized in depth using confocal, electron and atomic force microscopy, proteomics, immunoblot
23 analysis, and enzymatic assays. Furthermore, we examined the *in vivo* interchangeability of β -
24 carboxysome building blocks with other BMC components. To our knowledge, this is the first production
25 of functional β -carboxysome-like structures in heterologous organisms. It provides important information
26 for the engineering of fully functional carboxysomes and CO₂-fixing modules in higher plants. The study
27 strengthens our synthetic biology toolbox for generating BMC-based organelles with tunable activities
28 and new scaffolding biomaterials for metabolic improvement and molecule delivery.

29

30 1 Introduction

31 Subcellular compartmentalization is a hallmark of eukaryotic cells. It allows cells to perform and confine
32 various chemical reactions in space and time and provide a means for eliminating metabolic crosstalk and
33 enhancing the efficiency of compartmentalized metabolic pathways (Giessen and Silver, 2016; Hammer
34 and Avalos, 2017). Compartmentalization also occurs in the prokaryotic cytoplasm and membranes, such
35 as photosynthetic membranes (Liu, 2016). A particular example is bacterial microcompartments (BMCs),

36 which are distributed in at least 23 different bacterial phyla, sequestering enzymes within the specialized
37 niches from the cytoplasm (Axen et al., 2014; Chowdhury et al., 2014; Kerfeld and Erbilgin, 2015). All
38 BMCs are made entirely of proteins, including interior enzymes that catalyze sequential reactions and an
39 encapsulating shell. The architecture of the BMC shell resembles an icosahedral viral capsid, with facets
40 composed of hexameric and trimeric proteins and vertices capped by pentameric proteins (Kerfeld et al.,
41 2005; Tanaka et al., 2008; Klein et al., 2009; Sutter et al., 2017). These multiple protein paralogs have
42 specific permeability for the passage of metabolites. Three types of BMCs have been extensively studied:
43 the carboxysomes for CO₂ fixation, the PDU BMCs for 1,2-propanediol utilization, and the EUT BMCs
44 for ethanolamine utilization (Bobik et al., 2015; Kerfeld and Erbilgin, 2015).

45 The first BMCs identified were the carboxysomes, the central machinery for CO₂ fixation in cyanobacteria
46 (Shively et al., 1973). Composed of thousands of protein subunits, carboxysomes encapsulate the CO₂-
47 fixation enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) and carbonic anhydrase in a
48 protein-based shell. Compartmentalization of these enzymes and the selective permeability of the shell
49 that allows for the diffusion of HCO₃⁻ and diminishes CO₂ leakage ensure a high CO₂ concentration near
50 Rubisco (Cai et al., 2009), thereby favouring the higher carboxylation rates of Rubisco within
51 carboxysomes and enhanced carbon fixation. According to the Rubisco phylogeny, protein composition
52 and assembly, carboxysomes can be classified as α -carboxysomes (possessing Form 1A Rubisco) and β -
53 carboxysomes (containing plant-like Form 1B Rubisco)(Rae et al., 2013). It has been deduced that α -
54 carboxysomes and β -carboxysomes embrace distinct assembly pathways. The α -carboxysome shell
55 assembles concomitantly with aggregation of Rubisco (Iancu et al., 2010) or even without Rubisco
56 packing (Baker et al., 1998; Menon et al., 2008), whereas *de novo* assembly of β -carboxysomes follows
57 the “inside out” mode: the β -carboxysome shell begins assembly after RuBisCO aggregation (Chen et al.,
58 2012; Cameron et al., 2013).

59 The CO₂-fixing organelles in the model freshwater cyanobacterium *Synechococcus elongatus* PCC7942
60 (Syn7942) are β -carboxysomes, which have specific spatial location and regulation in cyanobacterial cells
61 (Savage et al., 2010; Sun et al., 2016). The β -carboxysome shell in Syn7942 is composed of the hexameric
62 proteins CcmK2-4 constructing shell facets (Kerfeld et al., 2005), CcmO (Rae et al., 2012) and the CcmL
63 pentamers located at the vertices (Tanaka et al., 2008). Rubisco enzymes form a densely-packed
64 paracrystalline array in the β -carboxysome lumen (Faulkner et al., 2017). The packing of Rubisco is
65 mediated by the 35 kDa truncated version of CcmM (CcmM35)(Long et al., 2010; Long et al., 2011). The
66 longer form of CcmM, Ccm58, contains an N-terminal domain with homology to γ -carbonic anhydrase
67 and a C-terminal domain comprising three small subunit-like domains (SSUs) with homology to RbcS,
68 the small subunit of Rubisco (Peña et al., 2010). CcmM interacts with the carboxysomal β -carbonic
69 anhydrase (CcaA), Rubisco, and CcmN that acts as a bridge between CcmM and the shell (Long et al.,
70 2007; Long et al., 2010; Kinney et al., 2012). Though structurally resembling icosahedral virus capsids,
71 the β -carboxysome from Syn7942 possesses a soft mechanical fingerprint (Faulkner et al., 2017).

72 The self-assembly, modularity and metabolic enhancement of BMCs make these bacterial organelles an
73 ideal engineering objective (Frank et al., 2013). BMC shell proteins have been shown to self-assemble to
74 form flat sheets or tubular structures (Pang et al., 2014; Noël et al., 2016; Sutter et al., 2016), and empty
75 shells (Lassila et al., 2014; Cai et al., 2015; Held et al., 2016; Sutter et al., 2017). Expression of *ccmK1*,
76 *ccmK2*, *ccmL*, and *ccmO* from the cyanobacterium *Halotheca* sp. PCC 7418 in *Escherichia coli* could
77 generate the synthetic β -carboxysome shells, ~ 25 nm in diameter (Cai et al., 2016). Likewise, empty
78 BMC shells with the diameter of ~ 40 nm were obtained by expressing a synthetic operon of *Haliangium*
79 *ochraceum* BMC shell genes (Lassila et al., 2014; Sutter et al., 2017). The empty BMC shells, without
80 encapsulated enzymes, are notably smaller than the native BMCs. Furthermore, previous studies have

81 demonstrated the possibilities of engineering entire PDU BMCs from *Citrobacter freundii* and EUT
82 BMCs from *Salmonella enterica* in *E. coli* (Parsons et al., 2008; Parsons et al., 2010; Choudhary et al.,
83 2012). Expressing the α-carboxysome operon from a chemoautotroph *Halothiobacillus neapolitanus* has
84 also led to the production of recombinant carboxysome-like structures with CO₂ fixation activity in *E. coli*
85 (Bonacci et al., 2012) and in a Gram-positive bacterium *Corynebacterium glutamicum* (Baumgart et al.,
86 2017). Apart from engineering BMCs in bacterial expression systems, efforts have been made to express
87 β-carboxysome components in the chloroplasts of the model plant tobacco *Nicotiana benthamiana*. It was
88 illustrated that transient expression of β-carboxysome proteins CcmK2, CcmM, CcmL, CcmO and CcmN
89 could lead to the formation of carboxysome-like circular structures (Lin et al., 2014a). Moreover, Rubisco
90 enzymes with a high carboxylation rate could be produced in tobacco chloroplasts by importing Syn7942
91 Rubisco-CcmM35 complexes to replace endogenous plant Rubisco (Lin et al., 2014b; Occhialini et al.,
92 2016). Despite the recent advancement of carboxysome engineering, until now, heterologous production
93 of entire functional β-carboxysome structures in any non-native hosts has not been reported, to our
94 knowledge.

95 In this study, we generated a synthetic operon consisting of 12 β-carboxysome genes from Syn7942 and
96 expressed the construct in *E. coli* to produce synthetic β-carboxysome-like structures with CO₂ fixation
97 capacity. We further purified the synthetic β-carboxysomes and elucidated their structure, activity and
98 interchangeability. This study empowers our toolbox for engineering functional BMC structures in
99 heterologous organisms and emphasises the necessity of optimizing the expression of carboxysome
100 modules. It represents a step towards modulating BMC structures and functions using synthetic biology,
101 to produce new nanobioreactors and biomaterials with expanded biotechnological applications.

102

103 **2 Results**

104 **Expression of Syn7942 β-carboxysomes in *Escherichia coli***

105 In contrast to the genes of α-carboxysomes that are mainly clustered together into a single operon, β-
106 carboxysome genes typically appear in multiple dispersed gene clusters throughout the genome (Rae et
107 al., 2013). In Syn7942, the twelve genes encoding β-carboxysome proteins are distributed in five different
108 chromosomal loci and the major carboxysome genes are located in one locus containing the *ccmKLMNO*
109 and *rbclS* operons (Figure 1A). We generated a synthetic operon, by assembling all the twelve β-
110 carboxysomal genes, to express β-carboxysome proteins in *E. coli* (Supplementary Table 1). Additionally,
111 the *ccmKLMNO* and *rbclS* operons were assembled into a single synthetic operon, for producing β-
112 carboxysome structures with a reduced number of genes to be expressed. These two operons, containing
113 the native promoters of individual carboxysome genes, were then inserted into an *E. coli* expression
114 vector pETM11 under a T7 promoter, to generate the pLFbC901 and pLFbC601 constructs in the host *E.*
115 *coli* BL21DE3 cells (Figure 1A).

116 Expression of β-carboxysome components using pLFbC901 and pLFbC601 was induced by IPTG. The
117 formation of β-carboxysome structures was verified by thin-section electron microscopy. The
118 carboxysome-like structures (~ 200 nm in diameter) with a high internal protein density were observed in
119 both recombinant *E. coli* cells harbouring pLFbC901 and pLFbC601 but are invisible in the WT *E. coli*
120 cells without carboxysome-expressing vectors (Figure 1B). Rubisco assays were performed to determine
121 the *in vivo* carbon fixation activities of recombinant *E. coli* cells. Compared with WT cells, both
122 pLFbC601 and pLFbC901 exhibit high carbon fixation activities, suggesting the functioning of
123 recombinant carboxysome-like structures (Figure 1C). In addition, the recombinant *E. coli* cells

124 containing pLFbC901 present a higher carbon fixation activity than the pLFbC601 cells, revealing the
125 importance of CcaA, CcmK3, CcmK4, CcmP and RbcX in maintaining the carboxysome structure and
126 function. Considering the higher carbon fixation activity, we mainly focused on the pLFbC901 expression
127 in the following study.

128 Expression of β -carboxysome protein components in the recombinant *E. coli* pLFbC901 was characterised
129 by SDS-PAGE. Figure 1D illustrates that although the overall expression levels of β -carboxysome
130 proteins are low, some β -carboxysome proteins are identifiable from the total cell extracts (see also
131 Supplementary Figure 1). Proteomic analysis of the cell extracts further confirmed the presence of five β -
132 carboxysome components, including the shell proteins (CcmK2, CcmK4), shell-associated proteins
133 (CcmM) and internal proteins (RbcL, RbcS)(Supplementary Table 2). Some of the β -carboxysome
134 proteins, namely CcmL, CcmO, CcmK3, CcaA, CcmN, CcmP, and RbcX, were not detectable. These
135 downstream genes might be translated inefficiently, due to their distant locations to the inducible T7
136 promoter. Interestingly, the recent study illustrated that CcmN, CcmP, and RbcX were also not detectable
137 in isolated β -carboxysomes from Syn7942 (Faulkner et al., 2017).

138 To evaluate the assembly of β -carboxysomal proteins in *E. coli*, we generated another two plasmids. One
139 (RbcL-eGFP-RbcS) contains the *rbcLS* operon with the enhanced green fluorescence protein (eGFP) gene
140 fused at the 3' end of *rbcL*, which encodes the interior protein RbcL (Rubisco large chain) and the other
141 (CcmK4-eGFP) comprises the gene of the shell protein CcmK4 fused with *egfp* at the 3' end. The two
142 plasmids were transformed into *E. coli* strains, respectively, and were expressed in the absence or
143 presence of pLFbC901 (Figure 2). *E. coli* WT cells were treated in the same conditions as the control
144 (Figure 2, left). Without the expression of pLFbC901, the expressed RbcL-eGFP-RbcS and CcmK4-eGFP
145 are evenly distributed in the cytoplasm, and no detectable protein aggregation was visualized under this
146 induction condition (Figure 2, middle). Co-expression of pLFbC901 with RbcL-eGFP-RbcS or CcmK4-
147 eGFP induces the formation of protein assemblies (Figure 2, right), exhibiting strong fluorescence spots
148 preferably at the cell polar and relative faint fluorescence spots in the cytoplasm (Figure 2, right, arrows).
149 The cytoplasmic foci observed resemble the β -carboxysome structures observed in Syn7942 (Savage et
150 al., 2010; Cameron et al., 2013; Sun et al., 2016), indicative of the formation of β -carboxysome structures;
151 whereas the polar assemblies could be the inclusion bodies consisting of RbcL or CcmK4, as has been
152 discerned in the heterologous expression of PDU BMCs and α -carboxysomes (Parsons et al., 2010;
153 Bonacci et al., 2012).

154 **Purification and characterization of synthetic β -carboxysomes**

155 Purification of recombinant β -carboxysomes was performed by sucrose density gradient centrifugation,
156 following the induction of *E. coli* constructs (Figure 3A). We used the RbcL-eGFP-RbcS and pLFbC901
157 co-expression strain to facilitate the detection of synthetic β -carboxysomes during purification. The eGFP
158 tagging has been used in isolating natural β -carboxysomes from Syn7942 (Faulkner et al., 2017). The
159 majority of synthetic β -carboxysome-like structures were detected in the 50% sucrose fractions by
160 fluorescence imaging (Figure 3B). Rubisco assays of each β -carboxysome fraction also reveal that the
161 50% fraction presents the highest Rubisco activity (Figure 3C). By contrast, native β -carboxysomes from
162 Syn7942 were found in the 40% fraction (Faulkner et al., 2017). This is ascribed to the relatively larger
163 size of synthetic β -carboxysome-like structures (~ 200 nm, Figure 1B) in contrast to native β -
164 carboxysomes (~ 150 nm)(Faulkner et al., 2017).

165 The same purification strategy was applied to isolate synthetic β -carboxysomes without eGFP fusion.
166 Functional synthetic β -carboxysomes were also obtained in the 50% fraction, as confirmed by ^{14}C
167 radiometric Rubisco assays (Figure 3D). It is worth noting that the yield of synthetic β -carboxysomes was

low and the carboxysome proteins were not clearly identifiable in SDS-PAGE (data not shown). However, immunoblot analysis using anti-RbcL and anti-CcmK2 revealed the presence of interior and shell proteins in isolated synthetic β-carboxysomes (Figure 3E).

Electron microscopy images of purified β-carboxysomes describe that synthetic β-carboxysomes are 200-300 nm in diameter (Figures 4A and 4B), slightly larger and more heterologous than native β-carboxysomes from Syn7942 (Faulkner et al., 2017). Moreover, they do not exhibit manifestly a regular icosahedral shape and symmetry, though the straight edges and vertices are visible. Despite the possibility of being artifacts in sample preparation, the structural variations of native and synthetic β-carboxysomes may be a consequence of the change in protein abundance, ratio and organization, which requires further optimization.

Closer inspection of the electron micrographs reveals the highly-ordered paracrystalline arrays of interior proteins within the β-carboxysomal lumen (Figures 4B and 4C). Cross-section analysis illustrates that the center-to-center distance of the interior proteins is about 10 nm (Figure 4D). The specific protein arrangement observed in synthetic β-carboxysome-like particles is in good agreement with the Rubisco packing of the native β-carboxysome discriminated in previous studies (Kaneko et al., 2006; Faulkner et al., 2017). The individual proteins in the paracrystalline arrays depicted in the EM image were further analyzed using four-fold symmetrized cross-correlation single-particle averaging (Fechner et al., 2009; Liu et al., 2011b; Casella et al., 2017). The average protein structure matches well the atomic structure of Rubisco holoenzyme (Figure 4E). All of these results suggest that these proteins are Rubisco enzymes and are densely packed in the carboxysome-like structures.

Furthermore, we studied the architecture of synthetic β-carboxysomes in solution using atomic force microscopy (AFM). The identification and structural integrity of eGFP-fused β-carboxysomes (RbcL-eGFP-RbcS) were verified by simultaneous AFM-fluorescence imaging (Figure 5). The single synthetic β-carboxysome is ~ 250 nm large and ~ 180 nm high, consistent with TEM results (Figures 1B and 4A).

Expression of the synthetic β-carboxysome operon in Syn7942

Our study shows that the designed synthetic carboxysome operon can produce functional β-carboxysomes in *E. coli*. To verify if the synthetic β-carboxysome operon can be expressed in Syn7942 cells, we inserted the synthetic β-carboxysome operon into a Syn7942 expression vector pAM2991 under the control of IPTG-inducible Ptrc promoter (Ivleva et al., 2005). The generated vector was then transformed into the RbcL-eGFP Syn7942 cells grown in low light ($10 \mu\text{E} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$) and expressed in the presence of 100 μM IPTG (Figure 6A). The RbcL-eGFP Syn7942 cells express fluorescently tagged β-carboxysomes under endogenous genetic regulation (Sun et al., 2016). Confocal fluorescence images show that without IPTG induction, the Syn7942 cells harbour 1.7 ± 0.78 carboxysomes per cell (mean ± SD, $n = 250$, Figure 6B), consistent with the previous observation (Sun et al., 2016). IPTG induction results in an increase in carboxysome number per cell (2.6 ± 1.2 , mean ± SD, $n = 350$). The carboxysomes are spatially distributed along the long axis of the cell. Moreover, Rubisco assays elucidate a 13.7% increase in carbon fixation activity per cell (Figure 6C). These results demonstrated that the expression of synthetic β-carboxysome operon triggered the formation of extra β-carboxysomes that were physiologically positioned and active in native hosts.

Modularity and interchangeability of the β-carboxysome structure

The BMC shell is made of protein building blocks belonging to a family of homologous proteins (Yeates et al., 2011; Bobik et al., 2015). We evaluated the modularity of the β-carboxysome structure and

interchangeability of building blocks from distinct types of BMCs. The Form 1A Rubisco large subunit CbbL of α -carboxysomes from *Halothiobacillus neapolitanus* was fused with eYFP at the C-terminus and expressed in RbcL-CFP Syn7942 cells. Confocal images reveal the co-localization of CbbL-eYFP and RbcL-CFP (Figure 7A), corroborating that α -carboxysome CbbL can be encapsulated in the β -carboxysome lumen that comprises Form IB Rubisco enzymes, given the high sequence similarity between CbbL and RbcL (78%, Supplementary Figure 2). Likewise, the α -carboxysome shell protein CsoS1A was fused with CFP at the C-terminus and expressed in CcmK4-eYFP Syn7942 cells. The notable co-localization of CsoS1A-CFP and CcmK4-eYFP elucidates the incorporation of α -carboxysome shell proteins into β -carboxysome structures (Figure 7B), consistent with the previous finding (Cai et al., 2015).

In addition, the PDU microcompartment shell protein PduA from *Salmonella enterica serovar Typhimurium* (*S. Typhimurium*) fused with superfolder GFP (sfGFP) was inserted in pAM2991 and expressed in Syn7942. PduA-sfGFP forms fluorescent puncta and present a typical distribution pattern of β -carboxysomes in the cytoplasm of Syn7942 cells, indicative of the integration of PDU BMCs shell proteins into the β -carboxysome structure (Figure 7C).

225

226 3 Discussion

The self-assembly and modularity features and the specific role for CO₂ fixation make carboxysomes a desirable engineering objective for metabolic enhancement and production of new nanomaterials. In this study, we show the construction of synthetic β -carboxysome operons and, to our knowledge, the first production of functional β -carboxysome-like structures in heterologous organisms. The structure and activity of the synthetic β -carboxysomes as well as the modularity of β -carboxysome structures were characterized using a combination of molecular genetics, Rubisco assays, proteomics and immunoblot analysis, confocal, electron and atomic force microscopy.

The synthetic β -carboxysome operon was created to assemble 12 β -carboxysome genes. The recombinant β -carboxysomes, generated by the expression of the synthetic β -carboxysome operon, harbour densely packed Rubisco enzymes, CcmM and shell proteins (CcmK2, CcmK4), as confirmed by proteomic data (Supplementary Table 2), immunoblot analysis (Figure 3E), and Rubisco assays (Figure 3D). Our results reveal that these proteins are the main building blocks necessary for β -carboxysome formation and CO₂ fixation. The “inside-out” model of β -carboxysome biogenesis suggests that the assembly of β -carboxysome is initiated from the Rubisco-CcmM aggregation into procarboxysomes (Cameron et al., 2013), followed by the addition of CcmN which has a C-terminal encapsulation peptide (Kinney et al., 2012) to recruit shell proteins and result in the encapsulation of CcmK2 and CcmO. Seven of the 12 β -carboxysome proteins were not detected, likely suggesting their low protein abundance in the β -carboxysome structure or the impeded protein expression (i.e. CcaA), probably due to the distant gene location from the T7 promoter. Nevertheless, the yield of synthetic β -carboxysomes after purification is low (data not shown), confirming the overall expression of β -carboxysome proteins in the recombinant *E. coli* is restricted. Whether there are negative effects of the production of carboxysome-like structures on the growth of *E. coli* cells remains unknown. Moreover, the sizes of synthetic and native β -carboxysomes are relatively comparable (Figure 4), whereas the shape of synthetic β -carboxysomes is not well defined as that of the natural form, likely as a result of improper content and ratio of expressed carboxysome proteins. All these observations highlight the necessity of optimising the expression constructs and conditions in future work.

253 Despite the significance of Rubisco nucleation in shell encapsulation and as such β -carboxysome
254 formation (Cameron et al., 2013), it has been demonstrated that synthetic β -carboxysome shell structures
255 can be built by expressing the shell proteins CcmK1, CcmK2, CcmO, and CcmL (Cai et al., 2016). The
256 synthetic β -carboxysome shells are about 25 nm in diameter, notably smaller than the native β -
257 carboxysomes and synthetic β -carboxysomes produced in this study, suggesting the roles of interiors in
258 determining the overall shape of β -carboxysomes.

259 The majority of cyanobacteria possess Form 1B Rubisco, the same as plants. However, compared with β -
260 cyanobacteria Rubisco, plant Rubisco exhibits a relatively a low carboxylation activity but a high affinity
261 for CO₂ against O₂. Improving the catalytic activity of plant Rubisco has become a target of plant
262 engineering for enhanced photosynthesis and capacity (Carmo-Silva et al., 2015). One possible strategy is
263 to engineer functional cyanobacterial carboxysomes into plants and encapsulate plant Rubisco with the
264 carboxysome shell that has the selective permeability to HCO₃⁻ and CO₂. Prior to the present study, it has
265 been demonstrated that expressing β -carboxysome shell proteins and cyanobacterial Form 1B Rubisco in
266 tobacco chloroplasts is achievable (Lin et al., 2014a; Lin et al., 2014b; Occhialini et al., 2016). Here, we
267 describe the heterologous expression of functional β -carboxysomes in *E. coli*, which holds promise for
268 plant chloroplast engineering to construct entire functional carboxysome structures for enhanced crop
269 photosynthesis and productivity.

270 It is potentially challenging to perform multi-gene transformation in higher plants. The production of
271 functional β -carboxysomes requires not only the expression of 12 carboxysome proteins and possibly
272 other associated proteins, but also the manipulation of protein abundance and ratio. In a recent study, a
273 chimeric protein peptide CcmC, including three Rubisco small subunit-like domains, one CcaA and the
274 CcmN C-terminus encapsulation peptide, was designed to replace the endogenous β -carboxysome core
275 (CcmM, CcmN, CcaA)(Gonzalez-Esquer et al., 2015). The reengineered β -carboxysomes were shown to
276 support photosynthesis *in vivo*. In this study, we showed that the capacity of CO₂ fixation can be obtained
277 using a simplified carboxysome-expressing vector pLFbC601, which contains only seven β -carboxysome
278 genes in the *ccmKLMNO* and *rbclS* operons. This may represent an engineering objective to produce
279 functional carboxysomes in higher plants with a simplified building composition. Optimization of the
280 pLFbC601 expression to build functional CO₂-fixing modules awaits further investigation.

281 Despite the functional diversity, all BMCs share a common structural feature: an icosahedral protein shell
282 encapsulating internal catalytic enzymes (Bobik et al., 2015). The shell proteins have conserved sequences
283 and form typical BMC domains: the PF00936 domain for hexameric shell proteins, the PF03319 domain
284 for shell pentamers, and the PF00936 domain for trimeric shell proteins (Kerfeld et al., 2005; Tanaka et
285 al., 2008; Klein et al., 2009). We demonstrate that the α -carboxysome shell protein CsoS1A from
286 *Halothiobacillus neapolitanus*, the PDU shell hexameric protein PduA from *S. Typhimurium*, as well as
287 the Form 1A Rubisco large subunit CbbL from *Halothiobacillus neapolitanus*, can be integrated into the
288 native β -carboxysomes in Syn7942 cells, due to their sequence conservation and structural similarities
289 relative to CcmK2 and RbcL, respectively. Our results shed lights on the inherent modularity and
290 interchangeability of a range of BMC building blocks, which lay the foundation for generating chimeric
291 BMC structures in a predictive manner, to manipulate the metabolic activities of BMCs in native and
292 heterologous hosts and engineer BMC-based nanoreactors for specific functions. It will promote the “plug
293 and play” applications of BMC-like structures in biotechnology.

294

295 4 Material and Methods

296 Construction of synthetic operons, plasmids and strains

297 The strategy for constructing synthetic β -carboxysome operons was illustrated in Figure 1A. Genes were
298 cloned and synthesized from the genomic DNA of Syn7942 using the ligation-free method In-Fusion
299 cloning (Clontech, CA). Every operon contains 50 bp upstream and 20 bp downstream to preserve the
300 native promoter and ribosome-binding site sequence of all the genes. The operons were inserted into
301 either the pETM11 vector digested at the EcoRI and XhoI sites or the pAM2991 vector (a gift from Susan
302 Golden, Addgene plasmid # 40248 (Ivleva et al., 2005)) at the EcoRI and BamHI sites. Genes encoding
303 RbcL-eGFP-RbcS were cloned from the genomic DNA of the RbcL-eGFP Syn7942 mutant (Sun et al.,
304 2016) and inserted into the pTTQ18 vector at the EcoRI and XbaI sites; The same strategy was used for
305 cloning and inserting CcmK4-GFP at the EcoRI and KpnI sites. Genes encoding CsoS1A and CbbL were
306 cloned from the genomic DNA of *H. neapolitanus* using In-Fusion cloning, fused with the genes of CFP
307 and eYFP, respectively, and inserted into pAM2991. The *pduA* gene was cloned from the genome of
308 *Salmonella enterica serovar Typhimurium* LT2, fused with sfGFP, and inserted into pAM2991.
309 Restriction enzymes were purchased from New England Biolabs (UK). Reagents were purchased from
310 Sigma-Aldrich (US).

311 Construction of plasmids and synthetic operons was carried out in *E. coli* strain Top10 or BL21(DE3)
312 cells grown aerobically at 37°C in lysogeny broth (LB) medium. Medium supplements were used, where
313 appropriate, at the following final concentrations: 50 $\mu\text{g}\cdot\text{mL}^{-1}$ kanamycin, 50 $\mu\text{g}\cdot\text{mL}^{-1}$ spectinomycin, 100
314 $\mu\text{g}\cdot\text{mL}^{-1}$ ampicillin.

315 Syn7942 wildtype and RbcL-eGFP and CcmK4-eGFP mutants (Sun et al., 2016) were transformed with
316 the generated pAM2991 vectors described above. Cells from single colonies were grown in BG11 media
317 at a constant light until the middle of the exponential growth phase, and then induced with 100 μM IPTG
318 overnight at 30°C.

319 Expression of synthetic β -carboxysome operons and purification of synthetic β -carboxysomes

320 The transformed *E. coli* strain BL21DE3 was grown overnight at 37°C. The cultures were then diluted
321 1:100 in 1L flasks of LB and incubated at 37°C with shaking until OD₆₀₀ reaches 0.6, then induced with
322 50 μM IPTG and grow overnight at 18°C. Cells were harvested by centrifugation at 5,000 g for 10 min.
323 Synthetic carboxysome purification was carried out as described previously (Faulkner et al., 2017). The
324 cell pellet was washed once with TE buffer, and was then resuspended into TE buffer in the presence of
325 cell lytic B (Sigma Aldrich, US) and 1% protease inhibitor cocktails (Thermo-Fisher, UK), and then cell
326 breakage by Stansted pressure cell homogenizer (Stansted Fluid Power, UK). Cell lysate was then treated
327 with 1% *n*-dodecyl β -maltoside for 1 hour. Cell debris was removed by centrifugation, followed by a
328 centrifugation at 50,000 g to enrich synthetic β -carboxysomes. The generated pellet was resuspended in
329 TE buffer and followed by centrifugation using a step sucrose density gradient. The sucrose fractions were
330 then harvested and characterized to determine the presence and activities of carboxysome structures.

331 SDS-PAGE, immunoblot analysis, proteomic analysis, and Rubisco assay

332 SDS-PAGE, immunoblot analysis using anti-RbcL antibody were carried out as described previously (Sun
333 et al., 2016; Faulkner et al., 2017). Immunoblot analysis of CcmK2 was performed using anti-CcmK2
334 antibody (a gift from Cheryl Kerfeld) and horseradish peroxidase-conjugated goat anti-rabbit
335 immunoglobulin G secondary antibody (GE Healthcare, US).

336 Whole-cell proteomic analysis was performed as follows: 20 mL of *E. coli* cells expressing synthetic β -

337 carboxysome proteins were harvested. After washed and sonicated with 25 mM NH_4HCO_3 , the cell
338 lysates were used for mass spectrometry analysis, as described previously (Faulkner et al., 2017).

339 For *in vivo* Rubisco assays, 120 μL of *E. coli* cells ($\text{OD}_{600} = 4.0$) or 120 μL of Syn7942 cells ($\text{OD}_{750} = 4.0$)
340 in Rubisco assay buffer (100 mM EPPS, pH 8.0; 20 mM MgCl_2) were added into a 250 μL reaction.
341 Radiometric assays were performed following the previously described protocol (Price and Badger, 1989)
342 with additional cell permeabilization treatment (Schwarz et al., 1995). Rubisco activities of isolated
343 synthetic β -carboxysome structures were determined as previously described (Faulkner et al., 2017). 5 μL
344 of 1 $\text{mg}\cdot\text{mL}^{-1}$ isolated synthetic β -carboxysomes in Rubisco assay buffer were added. Samples were added
345 into scintillation vials containing $\text{NaH}^{14}\text{CO}_3$ at a final concentration of 25 mM and incubated at 30°C for 2
346 min before the addition of D-ribulose 1,5-bisphosphate sodium salt hydrate (RuBP, Sigma Aldrich, US) at
347 a final concentration of 5 mM. The reaction ensued for 5 min before being terminated by adding 2:1 by
348 volume 10% formic acid. Samples were dried for at least 30 min at 95°C to remove unfixed ^{14}C before
349 resuspending the fixed ^{14}C pellets with ultra-pure water and adding 2 ml of scintillation cocktail (Ultima
350 Gold XR, PerkinElmer, US). Radioactivity measurements were then taken using a scintillation counter
351 (Tri-Carb, PerkinElmer, US). Raw readings were used to calculate the amount of fixed ^{14}C , and then
352 converted to the total carbon fixation rates. Results are presented as mean \pm standard deviation (SD).

353 **Fluorescence microscopy**

354 *E. coli* BL21(DE3) cells possessing the *rbcL-egfp-rbcS-pTTQ18* and *ccmK4-pTTQ18* plasmids with or
355 without pLFbC901 were cultured in 37°C until OD_{600} reaches 0.6 and then induced with 50 μM IPTG
356 overnight at 18°C. Preparation of *E. coli* and Syn7942 cells for confocal microscopy was performed as
357 described earlier (Liu et al., 2012). The *E. coli* and Syn7942 cells were imaged using a Zeiss LSM710 or
358 LSM780 with a 63 \times or 100 \times oil-immersion objective and excitation at 488 nm. Live-cell images were
359 recorded from at least five different cultures. All images were captured with all pixels below saturation.
360 Image analysis was carried out using ImageJ software (NIH Image). Automated analysis of the number of
361 carboxysomes in Syn7942 cells was programmed into the image analysis software ImageSXM
362 (<http://www.imagesxm.org.uk/>), as described earlier (Sun et al., 2016).

363 **Transmission electron microscopy**

364 *E. coli* cells expressing β -carboxysome proteins were pelleted and fixed for 1 h with 2%
365 paraformaldehyde and 2% glutaraldehyde in 0.1 M sodium cacodylate buffer at pH 7.2. Samples were
366 stained for 1 hour with 2% osmium tetroxide and 1.5% potassium ferrocyanide, followed by staining with
367 1% thiocarbohydrazide for 20 min and 2% osmium tetroxide for 30 min. Afterwards, samples were
368 washed with ultrapure water, stained with 1% uranyl acetate for overnight at 4°C. Dehydration was
369 performed with a series of increasing alcohol concentrations (30 to 100%) before the samples were
370 embedded in resin. Thin sections of 70 nm were cut with a diamond knife. The structures of isolated
371 carboxysomes were characterized using negative staining TEM and AFM imaging as described previously
372 (Faulkner et al., 2017). Samples were stained with 3% uranyl acetate. Images were recorded using an FEI
373 Tecnai G2 Spirit BioTWIN transmission electron microscope. Image analysis was carried out using
374 ImageJ software (NIH Image). Particle averaging was performed using cross-correlation based Java
375 routines for the ImageJ image processing package (Fechner et al., 2009; Liu et al., 2011b; Casella et al.,
376 2017).

377 **Atomic force microscopy**

378 AFM imaging was carried out in solution to ensure the structural and functional integrity of β -

379 carboxysomes (Faulkner et al., 2017). Confocal-AFM images were captured using a NanoWizard 3 AFM
380 (JPK) integrated with a Zeiss LSM880 confocal microscope. Samples were adsorbed on glass slides in
381 adsorption buffer (10 mM Tris-HCl, 150 mM KCl, 25 mM MgCl₂, pH 7.5) for 10 mins, and then washed
382 with imaging buffer (10 mM Tris-HCl, 150 mM KCl, pH 7.5)(Liu et al., 2009; Liu et al., 2011a; Liu et al.,
383 2011b; Casella et al., 2017). Confocal images were captured using a 40× objective with 488 nm
384 excitation. Particles with high-intensity GFP signal were imaged by AFM in Quantitative Imaging (QI)
385 mode. The scanning force is ~100 pN. Image analysis was performed using JPK SPM Data Processing
386 (JPK).

387

388 **Acknowledgments**

389 The authors thank Prof. Cheryl Kerfeld (Lawrence Berkeley National Laboratory) for kindly providing the
390 anti-CcmK2 antibody. The authors also thank Dr. James Johnson and Dr. Hannah McCue (Liverpool
391 GeneMill, Biotechnology and Biological Sciences Research Council Grant, BB/M00094X/1) for advice
392 on synthetic biology; thank Mrs. Deborah Simpson for proteomic analysis; thank Mrs. Alison Beckett for
393 technical support of electron microscopy; thank the Liverpool Centre for Cell Imaging for technical
394 assistance and provision of Bio-AFM (Biotechnology and Biological Sciences Research Council ALERT
395 2014 Grant, BB/M012441/1).

396

397 **Funding**

398 This research was supported by Royal Society (UF120411, RG130442, IE131399, CH160004, and
399 RGF\EA\180233) and Biotechnology and Biological Sciences Research Council Grant (BB/M024202/1
400 and BB/M024202/1), Leverhulme Trust (ECF-2016-778, to FH) and Chinese Scholarship Council.

401

402 **Author Contributions**

403 YF and L-NL conceived the experiments; YF, FH, MF, QJ, GFD, and MY conducted the experiments;
404 YF, FH, MF, and L-NL analyzed the results; YF and L-NL wrote the manuscript. All authors reviewed
405 the manuscript.

406

407 **References**

- 408 Axen, S.D., Erbilgin, O., and Kerfeld, C.A. (2014). A taxonomy of bacterial microcompartment loci constructed by
409 a novel scoring method. *PLoS Comput Biol* 10(10), e1003898. doi: 10.1371/journal.pcbi.1003898.
- 410 Baker, S.H., Jin, S., Aldrich, H.C., Howard, G.T., and Shively, J.M. (1998). Insertion mutation of the form I cbbL
411 gene encoding ribulose biphosphate carboxylase/oxygenase (RuBisCO) in *Thiobacillus neapolitanus*
412 results in expression of form II RuBisCO, loss of carboxysomes, and an increased CO₂ requirement for
413 growth. *J Bacteriol* 180(16), 4133-4139.
- 414 Baumgart, M., Huber, I., Abdollahzadeh, I., Gensch, T., and Frunzke, J. (2017). Heterologous expression of the
415 *Halothiobacillus neapolitanus* carboxysomal gene cluster in *Corynebacterium glutamicum*. *J Biotechnol*
416 258, 126-135. doi: 10.1016/j.jbiotec.2017.03.019.

- 417 Bobik, T.A., Lehman, B.P., and Yeates, T.O. (2015). Bacterial microcompartments: widespread prokaryotic
418 organelles for isolation and optimization of metabolic pathways. *Mol Microbiol* 98(2), 193-207. doi:
419 10.1111/mmi.13117.
- 420 Bonacci, W., Teng, P.K., Afonso, B., Niederholtmeyer, H., Grob, P., Silver, P.A., et al. (2012). Modularity of a
421 carbon-fixing protein organelle. *Proc Natl Acad Sci U S A* 109(2), 478-483. doi: 10.1073/pnas.1108557109.
- 422 Cai, F., Bernstein, S.L., Wilson, S.C., and Kerfeld, C.A. (2016). Production and characterization of synthetic
423 carboxysome shells with incorporated luminal proteins. *Plant Physiol* 170(3), 1868-1877. doi:
424 10.1104/pp.15.01822.
- 425 Cai, F., Menon, B.B., Cannon, G.C., Curry, K.J., Shively, J.M., and Heinhorst, S. (2009). The pentameric vertex
426 proteins are necessary for the icosahedral carboxysome shell to function as a CO₂ leakage barrier. *PLoS*
427 *One* 4(10), e7521. doi: 10.1371/journal.pone.0007521.
- 428 Cai, F., Sutter, M., Bernstein, S.L., Kinney, J.N., and Kerfeld, C.A. (2015). Engineering bacterial
429 microcompartment shells: chimeric shell proteins and chimeric carboxysome shells. *ACS Synth Biol* 4(4),
430 444-453. doi: 10.1021/sb500226j.
- 431 Cameron, J.C., Wilson, S.C., Bernstein, S.L., and Kerfeld, C.A. (2013). Biogenesis of a bacterial organelle: the
432 carboxysome assembly pathway. *Cell* 155(5), 1131-1140. doi: 10.1016/j.cell.2013.10.044.
- 433 Carmo-Silva, E., Scales, J.C., Madgwick, P.J., and Parry, M.A. (2015). Optimizing Rubisco and its regulation for
434 greater resource use efficiency. *Plant Cell Environ* 38(9), 1817-1832. doi: 10.1111/pce.12425.
- 435 Casella, S., Huang, F., Mason, D., ZHAO, G.Y., Johnson, G.N., Mullineaux, C.W., et al. (2017). Dissecting the
436 native architecture and dynamics of cyanobacterial photosynthetic machinery. *Mol Plant* 10(11), 1434-
437 1448. doi: <http://dx.doi.org/10.1016/j.molp.2017.09.019>.
- 438 Chen, A.H., Afonso, B., Silver, P.A., and Savage, D.F. (2012). Spatial and temporal organization of chromosome
439 duplication and segregation in the cyanobacterium *Synechococcus elongatus* PCC 7942. *PLoS One* 7(10),
440 e47837. doi: 10.1371/journal.pone.0047837.
- 441 Choudhary, S., Quin, M.B., Sanders, M.A., Johnson, E.T., and Schmidt-Dannert, C. (2012). Engineered protein
442 nano-compartments for targeted enzyme localization. *PLoS One* 7(3), e33342. doi:
443 10.1371/journal.pone.0033342.
- 444 Chowdhury, C., Sinha, S., Chun, S., Yeates, T.O., and Bobik, T.A. (2014). Diverse bacterial microcompartment
445 organelles. *Microbiol Mol Biol Rev* 78(3), 438-468. doi: 10.1128/MMBR.00009-14.
- 446 Faulkner, M., Rodriguez-Ramos, J., Dykes, G.F., Owen, S.V., Casella, S., Simpson, D.M., et al. (2017). Direct
447 characterization of the native structure and mechanics of cyanobacterial carboxysomes. *Nanoscale* 9(30),
448 10662–10673. doi: 10.1039/C7NR02524F.
- 449 Fechner, P., Boudier, T., Mangenot, S., Jaroslowski, S., Sturgis, J.N., and Scheuring, S. (2009). Structural
450 information, resolution, and noise in high-resolution atomic force microscopy topographs. *Biophys J* 96(9),
451 3822-3831. doi: 10.1016/j.bpj.2009.02.011.
- 452 Frank, S., Lawrence, A.D., Prentice, M.B., and Warren, M.J. (2013). Bacterial microcompartments moving into a
453 synthetic biological world. *J Biotechnol* 163(2), 273-279. doi: 10.1016/j.jbiotec.2012.09.002.
- 454 Giessen, T.W., and Silver, P.A. (2016). Encapsulation as a strategy for the design of biological
455 compartmentalization. *J Mol Biol* 428(5 Pt B), 916-927. doi: 10.1016/j.jmb.2015.09.009.
- 456 Gonzalez-Esquer, C.R., Shubitowski, T.B., and Kerfeld, C.A. (2015). Streamlined construction of the
457 cyanobacterial CO₂-fixing organelle via protein domain fusions for use in plant synthetic biology. *Plant*
458 *Cell* 27(9), 2637-2644. doi: 10.1105/tpc.15.00329.
- 459 Hammer, S.K., and Avalos, J.L. (2017). Harnessing yeast organelles for metabolic engineering. *Nat Chem Biol*
460 13(8), 823-832. doi: 10.1038/nchembio.2429.
- 461 Held, M., Kolb, A., Perdue, S., Hsu, S.Y., Bloch, S.E., Quin, M.B., et al. (2016). Engineering formation of multiple
462 recombinant Eut protein nanocompartments in *E. coli*. *Sci Rep* 6, 24359. doi: 10.1038/srep24359.
- 463 Iancu, C.V., Morris, D.M., Dou, Z., Heinhorst, S., Cannon, G.C., and Jensen, G.J. (2010). Organization, structure,
464 and assembly of alpha-carboxysomes determined by electron cryotomography of intact cells. *J Mol Biol*
465 396(1), 105-117. doi: 10.1016/j.jmb.2009.11.019.
- 466 Ivleva, N.B., Bramlett, M.R., Lindahl, P.A., and Golden, S.S. (2005). LdpA: a component of the circadian clock
467 senses redox state of the cell. *EMBO J* 24(6), 1202-1210. doi: 10.1038/sj.emboj.7600606.

- 468 Kaneko, Y., Danev, R., Nagayama, K., and Nakamoto, H. (2006). Intact carboxysomes in a cyanobacterial cell
469 visualized by hilbert differential contrast transmission electron microscopy. *J Bacteriol* 188(2), 805-808.
470 doi: 10.1128/JB.188.2.805-808.2006.
- 471 Kerfeld, C.A., and Erbilgin, O. (2015). Bacterial microcompartments and the modular construction of microbial
472 metabolism. *Trends Microbiol* 23(1), 22-34. doi: 10.1016/j.tim.2014.10.003.
- 473 Kerfeld, C.A., Sawaya, M.R., Tanaka, S., Nguyen, C.V., Phillips, M., Beeby, M., et al. (2005). Protein structures
474 forming the shell of primitive bacterial organelles. *Science* 309(5736), 936-938. doi:
475 10.1126/science.1113397.
- 476 Kinney, J.N., Salmeen, A., Cai, F., and Kerfeld, C.A. (2012). Elucidating essential role of conserved carboxysomal
477 protein CcmN reveals common feature of bacterial microcompartment assembly. *J Biol Chem* 287(21),
478 17729-17736. doi: 10.1074/jbc.M112.355305.
- 479 Klein, M.G., Zwart, P., Bagby, S.C., Cai, F., Chisholm, S.W., Heinhorst, S., et al. (2009). Identification and
480 structural analysis of a novel carboxysome shell protein with implications for metabolite transport. *J Mol*
481 *Biol* 392(2), 319-333. doi: 10.1016/j.jmb.2009.03.056.
- 482 Lassila, J.K., Bernstein, S.L., Kinney, J.N., Axen, S.D., and Kerfeld, C.A. (2014). Assembly of robust bacterial
483 microcompartment shells using building blocks from an organelle of unknown function. *J Mol Biol*
484 426(11), 2217-2228. doi: 10.1016/j.jmb.2014.02.025.
- 485 Lin, M.T., Occhialini, A., Andralojc, P.J., Devonshire, J., Hines, K.M., Parry, M.A., et al. (2014a). β-Carboxysomal
486 proteins assemble into highly organized structures in Nicotiana chloroplasts. *Plant J* 79(1), 1-12. doi:
487 10.1111/tpj.12536.
- 488 Lin, M.T., Occhialini, A., Andralojc, P.J., Parry, M.A.J., and Hanson, M.R. (2014b). A faster Rubisco with
489 potential to increase photosynthesis in crops. *Nature* 513(7519), 547-550. doi: 10.1038/nature13776.
- 490 Liu, L.N. (2016). Distribution and dynamics of electron transport complexes in cyanobacterial thylakoid
491 membranes. *Biochim Biophys Acta* 1857(3), 256-265. doi: 10.1016/j.bbabi.2015.11.010.
- 492 Liu, L.N., Bryan, S.J., Huang, F., Yu, J.F., Nixon, P.J., Rich, P.R., et al. (2012). Control of electron transport routes
493 through redox-regulated redistribution of respiratory complexes. *Proc Natl Acad Sci U S A* 109(28), 11431-
494 11436. doi: 10.1073/pnas.1120960109.
- 495 Liu, L.N., Duquesne, K., Oesterhelt, F., Sturgis, J.N., and Scheuring, S. (2011a). Forces guiding assembly of light-
496 harvesting complex 2 in native membranes. *Proc Natl Acad Sci U S A* 108(23), 9455-9459. doi:
497 10.1073/pnas.1004205108.
- 498 Liu, L.N., Duquesne, K., Sturgis, J.N., and Scheuring, S. (2009). Quinone pathways in entire photosynthetic
499 chromatophores of *Rhodospirillum photometricum*. *J Mol Biol* 393(1), 27-35. doi:
500 10.1016/j.jmb.2009.07.044.
- 501 Liu, L.N., Sturgis, J.N., and Scheuring, S. (2011b). Native architecture of the photosynthetic membrane from
502 *Rhodobacter veldkampii*. *J Struct Biol* 173(1), 138-145. doi: 10.1016/j.jsb.2010.08.010.
- 503 Long, B.M., Badger, M.R., Whitney, S.M., and Price, G.D. (2007). Analysis of carboxysomes from *Synechococcus*
504 PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. *J Biol Chem*
505 282(40), 29323-29335. doi: 10.1074/jbc.M703896200.
- 506 Long, B.M., Rae, B.D., Badger, M.R., and Price, G.D. (2011). Over-expression of the beta-carboxysomal CcmM
507 protein in *Synechococcus* PCC7942 reveals a tight co-regulation of carboxysomal carbonic anhydrase
508 (CcaA) and M58 content. *Photosynth Res* 109(1-3), 33-45. doi: 10.1007/s11120-011-9659-8.
- 509 Long, B.M., Tucker, L., Badger, M.R., and Price, G.D. (2010). Functional cyanobacterial beta-carboxysomes have
510 an absolute requirement for both long and short forms of the CcmM protein. *Plant Physiol* 153(1), 285-293.
511 doi: 10.1104/pp.110.154948.
- 512 Menon, B.B., Dou, Z., Heinhorst, S., Shively, J.M., and Cannon, G.C. (2008). Halothiobacillus neapolitanus
513 carboxysomes sequester heterologous and chimeric RubisCO species. *PLoS One* 3(10), e3570. doi:
514 10.1371/journal.pone.0003570.
- 515 Noël, C.R., Cai, F., and Kerfeld, C.A. (2016). Purification and characterization of protein nanotubes assembled
516 from a single bacterial microcompartment shell subunit. *Advanced Materials Interfaces* 3(1), 1500295-n/a.
517 doi: 10.1002/admi.201500295.

- 518 Occhialini, A., Lin, M.T., Andralojc, P.J., Hanson, M.R., and Parry, M.A. (2016). Transgenic tobacco plants with
519 improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if
520 provided with elevated CO₂. *Plant J* 85(1), 148-160. doi: 10.1111/tpj.13098.
- 521 Pang, A., Frank, S., Brown, I., Warren, M.J., and Pickersgill, R.W. (2014). Structural insights into higher order
522 assembly and function of the bacterial microcompartment protein PduA. *J Biol Chem* 289(32), 22377-
523 22384. doi: 10.1074/jbc.M114.569285.
- 524 Parsons, J.B., Dinesh, S.D., Deery, E., Leech, H.K., Brindley, A.A., Heldt, D., et al. (2008). Biochemical and
525 structural insights into bacterial organelle form and biogenesis. *J Biol Chem* 283(21), 14366-14375. doi:
526 10.1074/jbc.M709214200.
- 527 Parsons, J.B., Frank, S., Bhella, D., Liang, M., Prentice, M.B., Mulvihill, D.P., et al. (2010). Synthesis of empty
528 bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated
529 organelle movement. *Mol Cell* 38(2), 305-315. doi: 10.1016/j.molcel.2010.04.008.
- 530 Peña, K.L., Castel, S.E., de Araujo, C., Espie, G.S., and Kimber, M.S. (2010). Structural basis of the oxidative
531 activation of the carboxysomal gamma-carbonic anhydrase, CcmM. *Proc Natl Acad Sci U S A* 107(6),
532 2455-2460. doi: 10.1073/pnas.0910866107.
- 533 Price, G.D., and Badger, M.R. (1989). Isolation and characterization of high CO₂-requiring-mutants of the
534 cyanobacterium *Synechococcus* PCC7942 : Two phenotypes that accumulate inorganic carbon but are
535 apparently unable to generate CO₂ within the carboxysome. *Plant Physiol* 91(2), 514-525.
- 536 Rae, B.D., Long, B.M., Badger, M.R., and Price, G.D. (2012). Structural determinants of the outer shell of beta-
537 carboxysomes in *Synechococcus elongatus* PCC 7942: roles for CcmK2, K3-K4, CcmO, and CcmL. *PLoS*
538 *One* 7(8), e43871. doi: 10.1371/journal.pone.0043871.
- 539 Rae, B.D., Long, B.M., Badger, M.R., and Price, G.D. (2013). Functions, compositions, and evolution of the two
540 types of carboxysomes: polyhedral microcompartments that facilitate CO₂ fixation in cyanobacteria and
541 some proteobacteria. *Microbiol Mol Biol Rev* 77(3), 357-379. doi: 10.1128/MMBR.00061-12.
- 542 Savage, D.F., Afonso, B., Chen, A.H., and Silver, P.A. (2010). Spatially ordered dynamics of the bacterial carbon
543 fixation machinery. *Science* 327(5970), 1258-1261. doi: 10.1126/science.1186090.
- 544 Schwarz, R., Reinhold, L., and Kaplan, A. (1995). Low activation state of ribulose-1,5-bisphosphate
545 carboxylase/oxygenase in carboxysome-defective *Synechococcus* mutants. *Plant Physiol* 108(1), 183-190.
- 546 Shively, J.M., Ball, F., Brown, D.H., and Saunders, R.E. (1973). Functional organelles in prokaryotes: polyhedral
547 inclusions (carboxysomes) of *Thiobacillus neapolitanus*. *Science* 182(4112), 584-586.
- 548 Sun, Y., Casella, S., Fang, Y., Huang, F., Faulkner, M., Barrett, S., et al. (2016). Light modulates the biosynthesis
549 and organization of cyanobacterial carbon fixation machinery through photosynthetic electron flow. *Plant*
550 *Physiol* 171(1), 530-541. doi: 10.1104/pp.16.00107.
- 551 Sutter, M., Faulkner, M., Aussignargues, C., Paasch, B.C., Barrett, S., Kerfeld, C.A., et al. (2016). Visualization of
552 bacterial microcompartment facet assembly using high-speed atomic force microscopy. *Nano Letters* 16(3),
553 1590-1595. doi: 10.1021/acs.nanolett.5b04259.
- 554 Sutter, M., Greber, B., Aussignargues, C., and Kerfeld, C.A. (2017). Assembly principles and structure of a 6.5-
555 MDa bacterial microcompartment shell. *Science* 356(6344), 1293-1297. doi: 10.1126/science.aan3289.
- 556 Tanaka, S., Kerfeld, C.A., Sawaya, M.R., Cai, F., Heinhorst, S., Cannon, G.C., et al. (2008). Atomic-level models
557 of the bacterial carboxysome shell. *Science* 319(5866), 1083-1086. doi: 10.1126/science.1151458.
- 558 Yeates, T.O., Thompson, M.C., and Bobik, T.A. (2011). The protein shells of bacterial microcompartment
559 organelles. *Curr Opin Struct Biol* 21(2), 223-231. doi: 10.1016/j.sbi.2011.01.006.

561 Figure Legend

562 **Figure 1. Synthetic operon design and heterologous expression of synthetic β -carboxysomes in *E.***
563 ***coli*.** (A) Schematic representation of the molecular organization of the natural β -carboxysome operons
564 spread in five loci within the Syn7942 genome and the synthetic β -carboxysome operons (pLFbC901 and
565 pLFbC601) inserted in the *E. coli* expression vector pETM11. Locus tags are indicated. (B) Thin-section
566 TEM images of *E. coli* WT and cells expressing β -carboxysome proteins using pLFbC601 and pLFbC901
567 plasmids. Arrows indicate β -carboxysome-like structures with polyhedral shapes observed in pLFbC601
568 and pLFbC901 *E. coli* cells. (C) *In vivo* carbon fixation assays of *E. coli* WT, pLFbC901 and pLFbC601
569 cells, indicating the CO₂-fixing activity of synthetic β -carboxysomes. A relatively higher CO₂-fixing
570 activity was determined in *E. coli* pLFbC901 cells compared with that of *E. coli* pLFbC601 cells. (D)
571 SDS-PAGE of the total cell extracts of the recombinant *E. coli* pLFbC901 and the *E. coli* WT cells.
572 Putative β -carboxysome proteins (RbcL, RbcS, CcmM, CcmK2, CcmK4) were identified based on their
573 molecular weights. The presence of these β -carboxysome protein blocks was further confirmed by
574 proteomic analysis (Supplementary Table 2). The gel image was spliced based on Supplementary Figure
575 1.

576 **Figure 2. Fluorescence microscopy images of *E. coli* WT and recombinant cells expressing β -**
577 **carboxysome components.** (A) *E. coli* cells expressing RbcL-eGFP-RbcS in the absence or presence of
578 the pLFbC901 vector. The *E. coli* WT cells were treated with the same conditions as the control. (B) *E.*
579 *coli* cells expressing CcmK4-eGFP in the absence or presence of the pLFbC901 vector. The appearance of
580 eGFP fluorescent puncta (arrows) indicates the assembly of β -carboxysome proteins in *E. coli*.

581 **Figure 3. Purification and characterization of synthetic β -carboxysomes produced in *E. coli*.** (A)
582 Step sucrose gradient separation of synthetic β -carboxysomes with and without RbcL-eGFP-RbcS from
583 pLFbC901. (B) Fluorescence detection of β -carboxysomes fused with eGFP in different sucrose fractions.
584 (C) Fluorescence intensities of eGFP signals in different sucrose fractions, indicating the aggregation and
585 abundance of β -carboxysome structures in each sucrose fraction. (D) Rubisco activities of β -
586 carboxysomes in 50% sucrose fraction determined by ¹⁴C radiometric assay. (E) Immunoblot analysis of
587 50% sucrose fraction using α -RbcL and α -CcmK2 antibodies, corroborating the presence of interior
588 (RbcL) and shell proteins (CcmK).

589 **Figure 4. Negative-staining TEM images of isolated recombinant β -carboxysomes.** (A)
590 Representative β -carboxysome structures showing the polyhedral shape. The diameter of synthetic β -
591 carboxysomes falls in the range of 200-300 nm. (B) Irregular β -carboxysome structures observed,
592 implying the heterogeneity of recombinant β -carboxysomes. However, the densely-packed paracrystalline
593 arrays of interior proteins could be discerned (dash lines). (C) Zoom-in inspection of the paracrystalline
594 arrays depicted in (B, rectangle). (D) Cross-section profile analysis of the paracrystalline arrays illustrates
595 the periodic arrangement of interior proteins. The center-to-center distance between neighboring proteins
596 is ~10 nm. (E) Four-fold symmetrized correlation average TEM image of the interior protein (top)
597 matches the atomic structure of Rubisco holoenzymes (bottom, PDB: 1UZH), indicating that the observed
598 densely-packed interior particles are Rubisco enzymes.

599 **Figure 5. Combined confocal and AFM imaging of isolated synthetic β -carboxysomes fused with**
600 **eGFP.** (A) Merged microscopic image captured using a hybrid JPK AFM-Zeiss 880 confocal microscope.
601 The triangular AFM cantilever is visible, and arrows indicate individual β -carboxysome particles
602 immobilized on mica surface. (B) Fluorescence image of a single recombinant β -carboxysome indicated
603 by the orange arrow in (A). (C) AFM topograph of the β -carboxysome captured simultaneously with the

604 fluorescence image (B). The combination of AFM-confocal fluorescence imaging ensures the
605 identification of intact recombinant β -carboxysomes. (D) Height profile of the β -carboxysome structure,
606 taken along the white arrowhead indicated in (B).

607 **Figure 6. Enhancement of β -carboxysome formation in Syn7942 by the expression of pLFbC901.**
608 (A) Confocal microscopy images illustrating the β -carboxysome formation and abundance in Syn7942
609 cells induced by IPTG. Chlorophyll fluorescence is shown in red and fluorescently tagged β -
610 carboxysomes are discerned as green puncta. (B) The number of β -carboxysomes per cell increases from
611 1.7 (no IPTG induction, $n = 250$) to 2.6 (100 μ M IPTG induction, $n = 350$). The whiskers show data from
612 minimum to maximum, “+” represents the average. Statistical analysis implies the significant difference
613 (***) $p < 0.001$, two-tailed t-test). (C) *In vivo* 14 C radiometric assays of Syn7942 cells induced by IPTG,
614 illustrating a 13.7% increase in Rubisco activity resulted in 100 μ M IPTG induction. Data are presented as
615 mean \pm SD. *** $p < 0.001$ ($n = 3$, two-tailed t-test).

616 **Figure 7. Interchangeability of BMC building blocks.** (A) Co-localization of RbcL-CFP and CbbL-
617 eYFP suggesting that α -carboxysome CbbL from *H. neapolitanus* can be encapsulated in the
618 cyanobacterial β -carboxysome structure. (B) Co-localization of CsoS1A-CFP and CcmK4-eYFP implying
619 the integration of α -carboxysome shell proteins from *H. neapolitanus* into the cyanobacterial β -
620 carboxysome shell. (C) Evenly distributed fluorescence puncta of PduA-sfGFP along the longitude axis of
621 Syn7942 cells indicating the incorporation of PDU shell proteins from *S. Typhimurium* in cyanobacterial
622 β -carboxysomes.