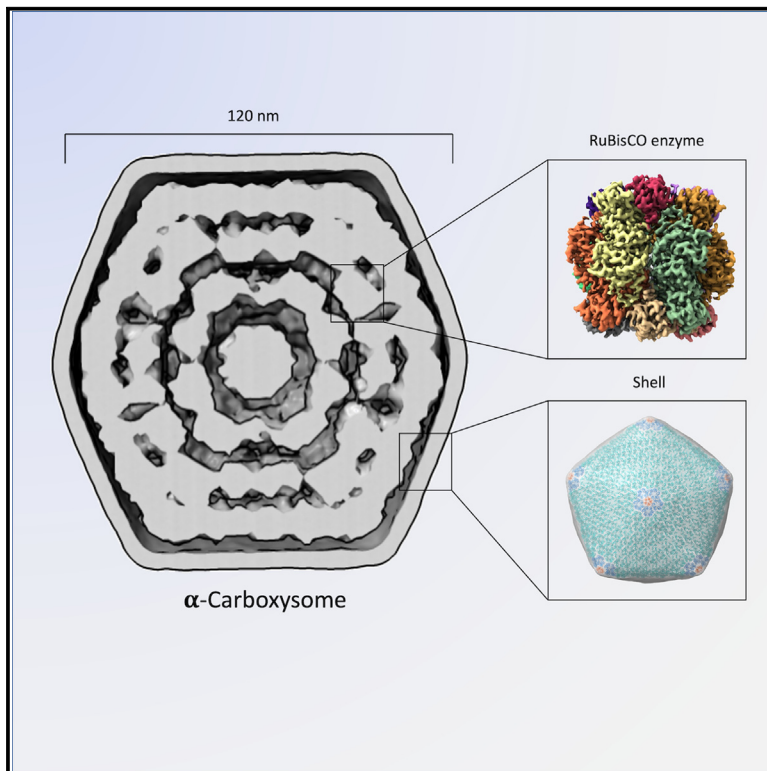


# Structure

## Single-particle cryo-EM analysis of the shell architecture and internal organization of an intact $\alpha$ -carboxysome

### Graphical abstract



### Authors

Sasha L. Evans,  
Monsour M.J. Al-Hazeem,  
Daniel Mann, ..., Gregory F. Dykes,  
Lu-Ning Liu, Julien R.C. Bergeron

### Correspondence

luning.liu@liverpool.ac.uk (L.-N.L.),  
julien.bergeron@kcl.ac.uk (J.R.C.B.)

### In brief

Carboxysomes are proteinaceous microcompartments found in cyanobacteria, which facilitate carbon fixation, the first step of photosynthesis, by sequestering the corresponding enzymes. Evans et al. used cryoelectron microscopy to characterize intact carboxysomes, which allowed them to establish a structural model for the carboxysome shell and for its internal enzyme organization.

### Highlights

- Carboxysomes are prokaryotic microcompartments that sequester photosynthetic enzymes
- We used single-particle cryo-EM analysis to characterize intact carboxysome complexes
- We have proposed a model of how enzymes are arranged within carboxysomes
- We also have generated a structural model of its shell architecture

Article

# Single-particle cryo-EM analysis of the shell architecture and internal organization of an intact $\alpha$ -carboxysome

Sasha L. Evans,<sup>1,7</sup> Monsour M.J. Al-Hazeem,<sup>2,7</sup> Daniel Mann,<sup>3</sup> Nicolas Smetacek,<sup>4,6</sup> Andrew J. Beavil,<sup>1</sup> Yaqi Sun,<sup>2</sup> Taiyu Chen,<sup>2</sup> Gregory F. Dykes,<sup>2</sup> Lu-Ning Liu,<sup>2,5,\*</sup> and Julien R.C. Bergeron<sup>1,4,8,\*</sup>

<sup>1</sup>Randall Centre for Cell and Molecular Biophysics, King's College London, London, UK

<sup>2</sup>Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

<sup>3</sup>Ernst-Ruska Centre 3, Forschungszentrum Jülich, Jülich, Germany

<sup>4</sup>Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield, UK

<sup>5</sup>College of Marine Life Sciences, and Frontiers Science Center for Deep Ocean Multispheres and Earth System, Ocean University of China, Qingdao, Shandong, China

<sup>6</sup>Present address: Boehringer Ingelheim RCV GmbH, Vienna, Austria

<sup>7</sup>These authors contributed equally

<sup>8</sup>Lead contact

\*Correspondence: [luning.liu@liverpool.ac.uk](mailto:luning.liu@liverpool.ac.uk) (L.-N.L.), [julien.bergeron@kcl.ac.uk](mailto:julien.bergeron@kcl.ac.uk) (J.R.C.B.)

<https://doi.org/10.1016/j.str.2023.03.008>

## SUMMARY

Carboxysomes are proteinaceous bacterial microcompartments that sequester the key enzymes for carbon fixation in cyanobacteria and some proteobacteria. They consist of a virus-like icosahedral shell, encapsulating several enzymes, including ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO), responsible for the first step of the Calvin-Benson-Bassham cycle. Despite their significance in carbon fixation and great bioengineering potentials, the structural understanding of native carboxysomes is currently limited to low-resolution studies. Here, we report the characterization of a native  $\alpha$ -carboxysome from a marine cyanobacterium by single-particle cryoelectron microscopy (cryo-EM). We have determined the structure of its RuBisCO enzyme, and obtained low-resolution maps of its icosahedral shell, and of its concentric interior organization. Using integrative modeling approaches, we have proposed a complete atomic model of an intact carboxysome, providing insight into its organization and assembly. This is critical for a better understanding of the carbon fixation mechanism and toward repurposing carboxysomes in synthetic biology for biotechnological applications.

## INTRODUCTION

Within cells, proteins tend to self-assemble and interact with other proteins or molecules to form active macromolecular machines,<sup>1–3</sup> which play central roles in many cellular processes.<sup>4,5</sup> Understanding the precise structures of natural protein assemblies is imperative for fundamental investigations of their biosynthesis and functions, and towards engineering artificial nanomaterials for new functions.<sup>6</sup>

Bacterial microcompartments (BMCs) are large macromolecular assemblies widespread in the bacterial domain.<sup>7,8</sup> BMCs serve as metabolic organelles, but unlike their eukaryotic counterparts, they have no lipid bilayer and are composed entirely of proteins.<sup>8</sup> By segregating metabolic enzymes from the cytosol, BMCs are thought to protect the cell from toxic intermediate metabolites and unwanted side reactions, and play pivotal roles in several enzymatic pathways, including autotrophic CO<sub>2</sub> fixation and catabolic processes.<sup>9–12</sup>

Despite their diverse range of functions, all BMCs possess a similar overall organization. They consist of a polyhedral proteinaceous shell, reminiscent of viral capsids. This shell encapsulates the enzymes involved in the corresponding metabolic pathway and acts as a semi-permeable physical barrier for molecule diffusion.<sup>1,13,14</sup> Structural studies of multiple BMC shell proteins in isolation have shown that they belong to three distinct categories: hexamers and pseudo-hexameric trimers, which tile the majority of the shell facets, and pentamers, which cap the vertices of the polyhedral shell.<sup>2,15–17</sup> Although our knowledge about the entire architecture of BMCs is still primitive, high-resolution cryoelectron microscopy (cryo-EM) structures of synthetic BMC minishells have provided insight into the organization of shell proteins and the dynamic nature of these proteinaceous shells for facilitating metabolite entry and exit.<sup>18–20</sup>

Carboxysomes were the first BMCs to be discovered.<sup>21</sup> They are found in cyanobacteria and some chemoautotrophs, and play a key role in carbon fixation.<sup>22</sup> Carboxysomes contain the

enzymes ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) and carbonic anhydrase (CA). CA catalyzes the conversion of cytosolic bicarbonate ( $\text{HCO}_3^-$ ) into  $\text{CO}_2$ , which is subsequently utilized by RuBisCO and fixed onto the 5-carbon sugar ribulose-1,5-bisphosphate (RuBP) as the first step in the Calvin-Benson-Bessham (CBB) cycle.<sup>23</sup> By generating carboxysomes to sequester these enzymes and allow the accumulation of  $\text{HCO}_3^-/\text{CO}_2$ , bacterial cells can provide an elevated level of  $\text{CO}_2$  around RuBisCO to enhance carbon fixation and overcome the competitive inhibition of RuBisCO carboxylation by  $\text{O}_2$ .<sup>22,24</sup> These intrinsic structural features allow carboxysomes to make a significant contribution to the global carbon fixation.<sup>22</sup> Notably, repurposing carboxysomes is an emerging discipline with applications in crop engineering, metabolic enhancement, bioenergy production, and therapeutics.<sup>22,25–27</sup>

Carboxysomes can be classified into two distinct groups:  $\alpha$ -carboxysomes, primarily encoded for by the *cso* operon, and  $\beta$ -carboxysomes, primarily encoded for by the *ccm* operon.<sup>28</sup> These two groups are distinct due to their protein composition and the types of RuBisCO encapsulated, belonging to form 1A and form 1B RuBisCO, respectively. Despite having been suggested to have evolved independently to adapt to different ecological niches, these two forms of RuBisCO demonstrate similar affinities for their substrates.<sup>29</sup> Unlike RuBisCO and shell proteins, the CA enzyme is largely evolutionarily distinct between  $\alpha$ -carboxysomes and  $\beta$ -carboxysomes. CsoSCA possesses a distinct fold from CcmM, the CA enzyme found in  $\beta$ -carboxysomes (and some  $\alpha$ -carboxysomes), although both proteins are essential for function, and both carry out similar roles.<sup>30</sup>

RuBisCO is a hexadecameric complex, comprised of eight large subunits and eight small subunits. The structures of RuBisCO from various cyanobacteria and plant species have been solved.<sup>31–35</sup> In  $\beta$ -carboxysomes, RuBisCO enzymes appear densely organized and form paracrystalline arrays that are important for  $\beta$ -carboxysome biogenesis.<sup>36–39</sup> In contrast, RuBisCO enzymes have been postulated to assemble concomitantly with the shell during  $\alpha$ -carboxysome biogenesis, a process promoted by the intrinsically disordered protein CsoS2, which induces the association between shell proteins and interiors.<sup>17,40</sup> The organization of RuBisCO inside  $\alpha$ -carboxysomes is poorly understood. Previous cryo-electron tomography analysis of  $\alpha$ -carboxysomes from the chemoautotrophic bacterium *Halothiobacillus neapolitanus*, and the cyanobacterial strains *Prochlorococcus marinus* MED4 and *Synechococcus* sp. WH8102 and WH8109, showed that the RuBisCO and CA enzymes appear to be packed densely and arranged into concentric layers.<sup>41–44</sup> However, no model has been proposed for the protein arrangement and interactions within the carboxysome and the architecture of the carboxysome shell. X-ray laser single-particle diffraction outlined the icosahedral shape of the  $\alpha$ -carboxysomes from *Halothiobacillus neapolitanus*, but no high-resolution structures of the intact carboxysome and the interior organization were reported.<sup>45</sup> These studies, together with those of other BMCs, have highlighted the challenges associated with structural characterization of these large heterogeneous macromolecular assemblies, specifically with great variations in the stoichiometric composition and interactions of individual building components that are adaptive to environmental changes.<sup>20,36,46</sup>

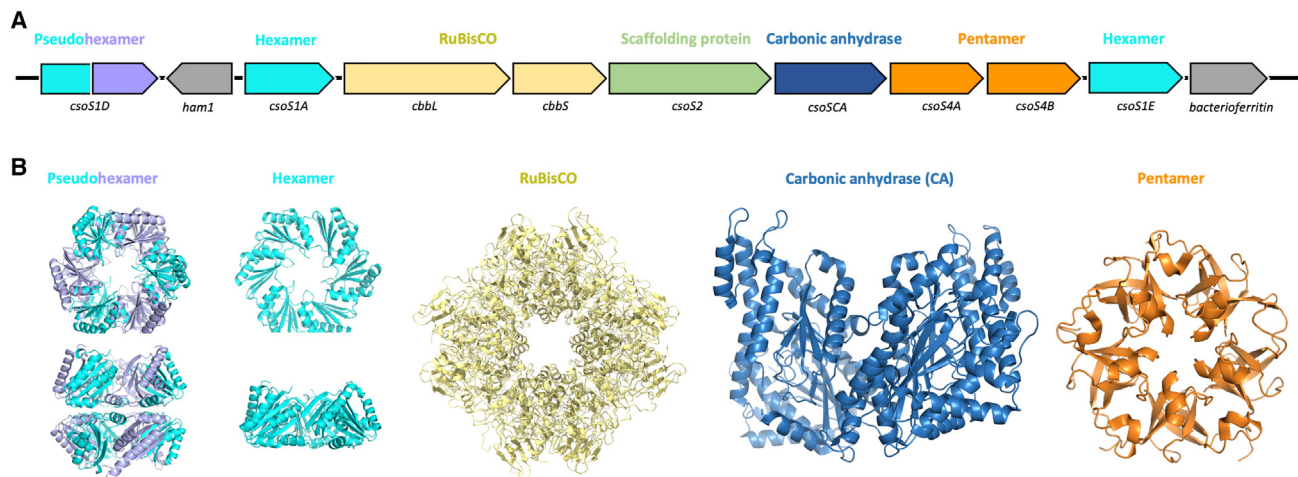
Here, we report the single-particle cryo-EM analysis of an intact  $\alpha$ -carboxysome from the marine  $\alpha$ -cyanobacterium *Cyanobium* sp. PCC 7001 (hereafter *Cyanobium*). We have determined the structure of its RuBisCO enzyme to 2.9 Å resolution, with the densities present for the substrate RuBP and for an unknown ligand. We also report a low-resolution structure of the icosahedral shell, demonstrating a range of dimensions, which precludes high-resolution analysis but nonetheless allows us to propose a hybrid structural model for the  $\alpha$ -carboxysome shell architecture. Moreover, 3D reconstruction combined with modeling allows us to propose a model for the arrangement of RuBisCO enzymes within the  $\alpha$ -carboxysome. The study provides insight into  $\alpha$ -carboxysome assembly, which will inform rational design and engineering of BMC-based nanostructures for diverse purposes.

## RESULTS

### Purification and single-particle analysis of $\alpha$ -carboxysomes from *Cyanobium*

The *Cyanobium*  $\alpha$ -carboxysome is encoded by a 9-gene operon, including 5 genes corresponding to shell proteins (*csoS1D*, *csoS1A*, *csoS4A*, *csoS4B*, and *csoS1E*), 3 genes encoding cargo enzymes RuBisCO (*cbbL* and *cbbS*) and CA (*csoSCA*), and 1 gene encoding the scaffolding protein CsoS2 (*csoS2*) (Figure 1A). CsoS1A and CsoS1E contain one Pfam00936 domain, homologous to the prototypical BMC shell hexamer, that tiles the majority of the  $\alpha$ -carboxysome shell; in addition, CsoS1E also possesses ~80 residues at its N terminus that are predicted to be unstructured. CsoS1D, containing two fused Pfam00936 domains, shows similarity to pseudo-hexamer trimers (sequence identity: 67%), corresponding to two stacked hexamers and presumably responsible for the passage of large molecules in and out of the carboxysome (we note, however, that the formation of pseudo-hexamers for this protein has not been confirmed experimentally). CsoS4A and CsoS4B have one Pfam03319 domain and belong to the family of BMC shell pentamers that cap the vertices of the polyhedral shell (Figure 1B).

To isolate functional carboxysomes, we grew *Cyanobium* photosynthetically in BG-11 freshwater medium until the late exponential phase (Figure S1A). Native  $\alpha$ -carboxysomes were isolated from *Cyanobium* using sucrose gradient ultracentrifugation and were enriched at the 30%–40% sucrose fraction (Figure S1B). SDS-PAGE (Figure 2A) and immunoblot analysis (Figure 2B) of the 40%–50% fraction demonstrated the presence of major  $\alpha$ -carboxysome components CbbL, CsoS2, and CsoS1A. Mass spectrometry analysis further indicated that the isolated  $\alpha$ -carboxysomes comprise all 9 building proteins (Table S1). Among them, RuBisCO subunits (CbbL and CbbS), CsoS2, and CsoS1A are highly abundant proteins, while CsoS4A, CsoS4B, and CsoS1D have low abundance in the  $\alpha$ -carboxysome, in good agreement with the mass spectrometry data of  $\alpha$ -carboxysomes from *Halothiobacillus neapolitanus*.<sup>47</sup> Nonetheless, due to the large errors for these measurements, the exact stoichiometry for each component could not be established unambiguously. Negative-stain EM showed that the isolated  $\alpha$ -carboxysomes have a canonical polyhedral BMC shape, with an average diameter of ~120 nm (Figure S1C), comparable to previous observations.<sup>29,48</sup> A <sup>14</sup>C-based assay of RuBisCO



**Figure 1. The *Cyanobium* sp. PCC7001  $\alpha$ -carboxysome**

(A) Gene organization of the  $\alpha$ -carboxysome operon, including genes encoding the shell hexamers (cyan) and pentamers (orange), the scaffolding protein (green), and the cargo enzymes RuBisCO (yellow) and CA (blue).

(B) Structural models of the corresponding  $\alpha$ -carboxysome proteins, based on the previously determined structures of homologous proteins.

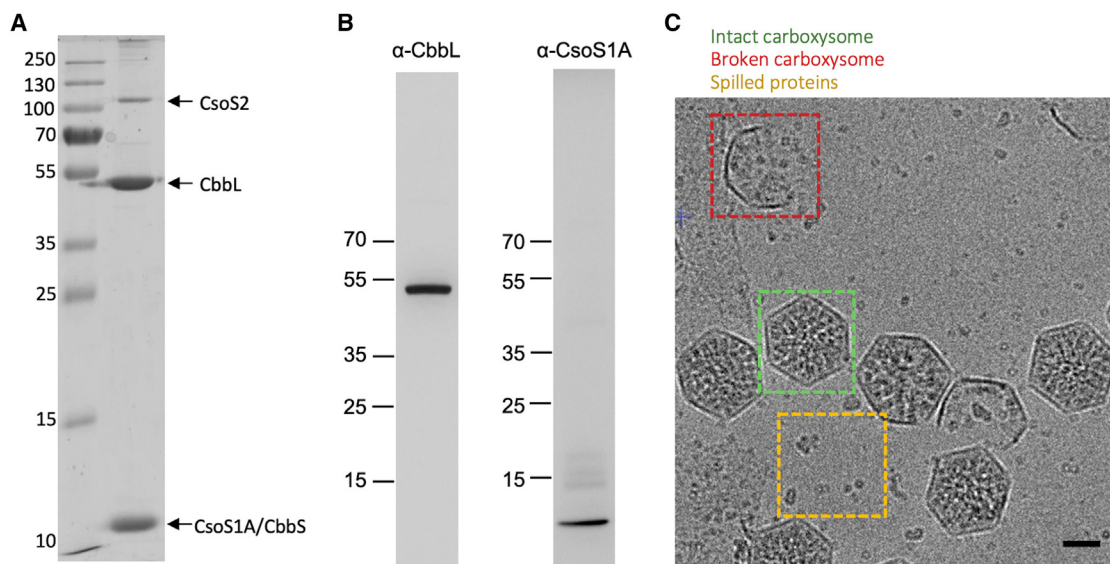
activity confirmed that the isolated  $\alpha$ -carboxysomes are catalytically active for carbon fixation (Figure S1D).

These intact, functional  $\alpha$ -carboxysomes were then subjected to single-particle cryo-EM analysis to study their 3D architecture. Initial screening showed a heterogeneous sample, containing intact carboxysomes with proteinaceous content, but also broken carboxysome shell fragments without any cargo inside, and disassembled proteins outside of the carboxysomes (Figure 2C). This deviates from negative-stain EM results (Fig-

ure S1C). We postulate that the broken shells largely result from sample handling and/or freezing, and that the disassembled proteins correspond to the proteins that spilled from the broken carboxysomes, including the enzymes RuBisCO and CA, as well as isolated shell components.

#### Structure of RuBisCO from native $\alpha$ -carboxysomes

To gain structural insights into the  $\alpha$ -carboxysomes, we collected a cryo-EM dataset of the sample described above using a



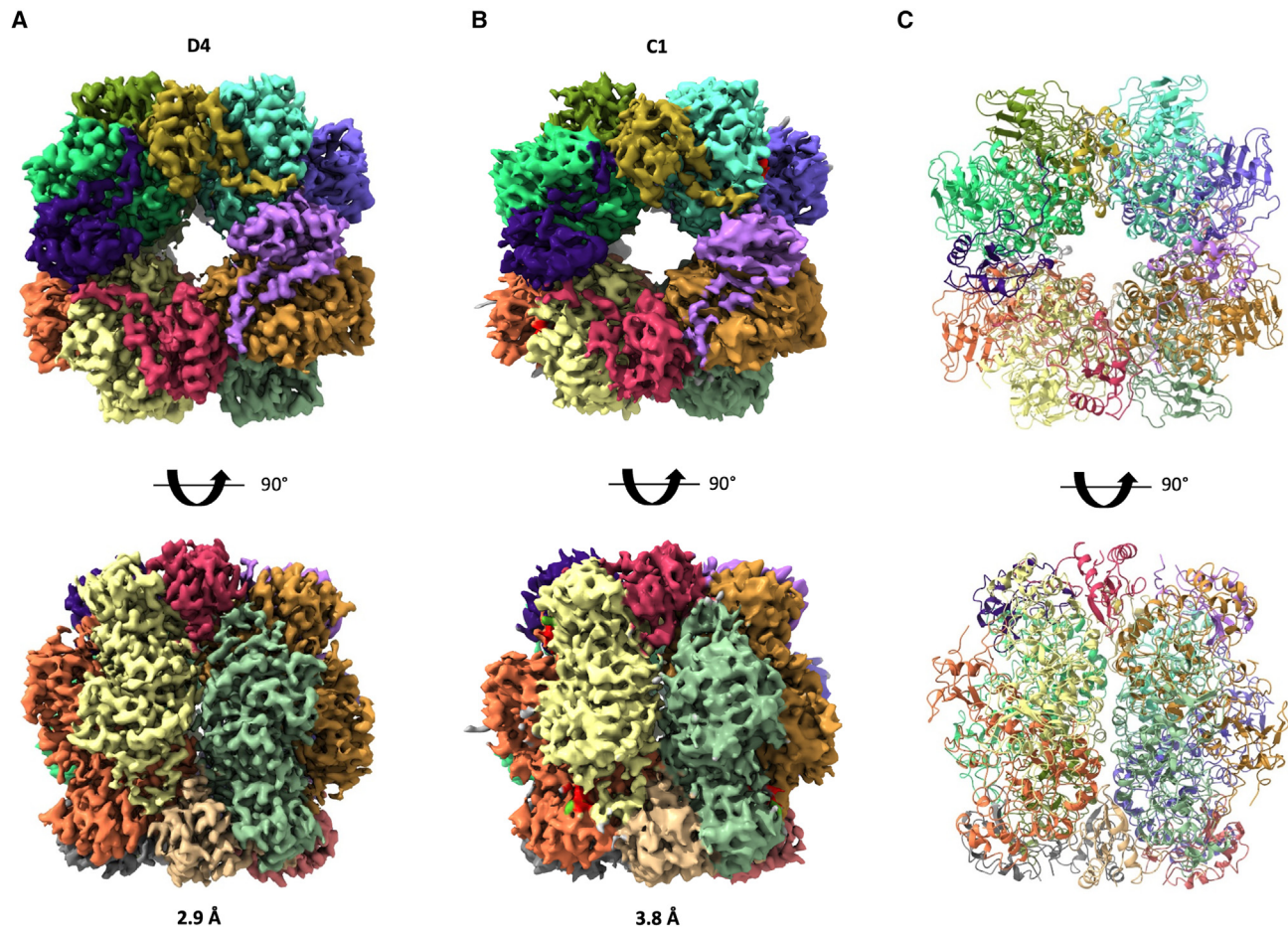
**Figure 2. Purification and cryo-EM analysis of the *Cyanobium*  $\alpha$ -carboxysome**

(A) SDS-PAGE of the purified carboxysomes. Bands for proteins CsoS2, CbbL, CbbS, and CsoS1A could be identified.

(B) Western blotting of the purified  $\alpha$ -carboxysome complex, using antibodies raised against peptides from CbbL and CsoS1, confirming the presence of both proteins.

(C) Cryo-electron micrograph of frozen-hydrated  $\alpha$ -carboxysome samples. Intact BMCs, with incorporated proteins, are visible (green box), along with broken ones (red box). Smaller protein complexes, presumably spilled from these, are also visible (yellow box). Scale bar: 50 nm.

See also Figure S1 and Table S1.



**Figure 3. Cryo-EM structure of RuBisCO isolated from *Cyanobium*  $\alpha$ -carboxysomes**

(A and B) The Coulomb potential map of RuBisCO, obtained from the particles of spilled RuBisCO proteins, colored and segmented by chain, is shown for the map with D4 symmetry (A) and without symmetry (B).

(C) Atomic model of the *Cyanobium* RuBisCO enzyme, in cartoon representation, and colored as in (A) and (B).

See also [Figures S2–S4](#) and [Video S1](#).

standard, high-magnification ( $\sim 1 \text{ \AA}/\text{pix}^2$ ) data collection approach. Because of the size of the complex, and its propensity to break (see above), we only obtained a few intact carboxysomes fully visible within each micrograph in this dataset. This largely precluded any analysis of the carboxysome complex. However, the spilled particles were readily visible on ice, and we were able to pick these, leading to a set of  $\sim 3,000,000$  particles.

Following initial 2D classifications, clear classes of two distinct molecular species could be identified. Specifically, several classes showed clear 4-fold symmetry and were visually identified as the RuBisCO (CbbL<sub>8</sub>-CbbS<sub>8</sub>) holoenzyme ([Figure S2A](#)). Additional classes were obtained for smaller protein(s), but these were featureless, and the corresponding protein(s) could not be identified based on 2D classes ([Figure S2B](#)). We hypothesize that these proteins correspond to a mixture of CA and shell proteins; however, this would require further validation.

We next conducted 3D refinement using the set of particles that could be identified as RuBisCO in the 2D classes. This yielded a 2.9 Å resolution Coulomb potential map ([Figures 3A, S2C, and S2E; Table 1; Video S1](#)), with eight large subunits

(CbbL) and eight small subunits (CbbS) readily identifiable. Using this map, we were able to build an atomic model of the *Cyanobium* RuBisCO enzyme ([Figure 3C; Table 1](#)). Notably, density is present in the active site, in a position suitable to be the substrate RuBP and magnesium ion, which allowed us to model these molecules in the active site ([Table 1](#)). This observation demonstrates that most RuBisCO enzymes within the carboxysome are active and bound to the substrates, in agreement with the RuBisCO assay results ([Figure S1D](#)).

However, we note that this density for the substrates is at a significantly lower contour level compared with the protein density, indicative of partial occupancy. To resolve which subunit(s) contained the substrate, we performed a refinement of the RuBisCO map without imposing any symmetry, leading to a second map at 3.8 Å resolution ([Figure 3B; Table 1](#)). As shown on [Figure S3](#), we observed varied levels of density in all eight active sites: for one chain, the full RuBP density can be readily identified (chain G), whereas in two chains (A and E), no density is present at all. Finally, we observe partial substrate density in five chains (C, I, K, M, and O), which could correspond to one or two

**Table 1. Cryo-EM data collection and structure refinement parameters**

	Dataset 1		Dataset 2
<b>Data collection</b>			
Voltage (kV)	300		300
Exposure (e/Å <sup>2</sup> )	30		29.7
Fractions	44		33
Defocus range (μm)	−0.5 to −1.5		−1 to −2.2
Pixel size (Å pix <sup>−1</sup> )	1.11		2.23
No. of micrographs	4,593		5,429
Initial particle no.	2,800,000		15,545
<b>Map refinement</b>			
Final particle no.	131,356	131,356	3,533
Resolution (Å)	2.87	3.79	18.25
Symmetry	D4	C1	I
<b>Structure refinement</b>			
Non-hydrogen atoms	35,138	35,070	−
Protein residues	4,432	4,424	−
Ligands	16	10	−
Protein B factor	30.19	81.71	−
Ligand B factor	68.53	113.28	−
Bond length RMSD (Å)	0.007	0.002	−
Bond angle RMSD (°)	0.790	0.668	−
MolProbity score	2.02	1.78	−
Clash score	8.23	11.83	−
Poor rotamers	2.51	0.11	−
Ramachandran favored (%)	96.11	96.81	−
Ramachandran allowed (%)	3.70	3.03	−
Ramachandran disallowed (%)	0.18	0.16	−
RMSD, root-mean-square deviation.			

molecules of the reaction product, 3-phosphoglycerate, and/or a mix of states that are averaged out through alignment. This confirms the activity of the RuBisCO enzymes present in carboxysome, with various stages of the reaction cycle present within one different subunit of the same complex. While we do not think that there is specific cooperativity between the eight subunits of the complex, this might indicate some correlation in the catalytic reaction between subunits, perhaps regulated by interacting proteins such as CA and/or CSoS2.

To further investigate the catalytic state of the RuBisCO complex, we analyzed the conformation of the large subunit loop 6, which had previously been shown to adopt two distinct conformations, corresponding to apo (open) or substrate-bound (closed) states of the enzymes.<sup>32,49</sup> In our RuBisCO map, the density of this loop is poorly resolved, suggesting that it is likely dynamic. Nonetheless, we were able to build an atomic model for the entire loop. When compared with the crystal structures of the spinach RuBisCO (71% identity to CbbL, 29% to CbbS) in both open and closed states, we observed that on our map, this loop corresponds to the open conformation (Figure S4A), likely providing an explanation for the range of substrate-/product-bound states observed (see above).

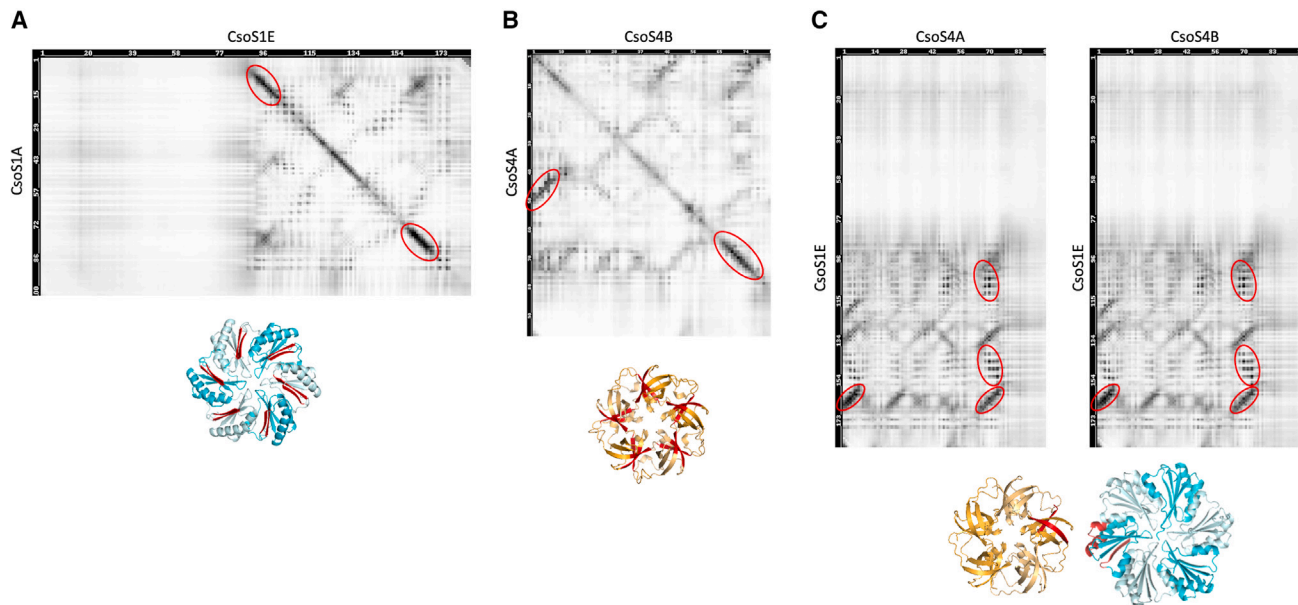
We also note that in the map obtained in the absence of symmetry, some unattributed, diffuse density at a lower contour level is present at the surface of the complex for one of the subunits (Figure S4B). This suggests that, for some of the particles, other proteins are bound in this location. This region of the map is at a much lower resolution and did not allow us to identify what protein this might be based on the density alone. Nonetheless, this finding provides evidence that there are other proteins bound to RuBisCO, which likely originates from the broken carboxysomes. Further investigation is required to determine which carboxysomal protein the density represents. We note, nonetheless, that this density is not in the region of the complex where CsoS2 had previously been shown to bind, suggesting that either CsoS2 has multiple modes of interaction with RuBisCO or that this density corresponds to a different protein.<sup>40</sup>

### Cryo-EM analysis of the $\alpha$ -carboxysome shell

The current structural information on  $\alpha$ -carboxysomes is limited to low-resolution tomography data.<sup>41,42,44,50,51</sup> We therefore attempted to use single-particle cryo-EM to gain insight into the overall architecture of the *Cyanobium*  $\alpha$ -carboxysome shell. As mentioned above, both the manual handling and freezing of the complex led to significant breakage, which prevented large data collection of intact carboxysomes. To address this, we froze grids immediately after purification, leading to a higher proportion of intact carboxysomes. In addition, we collected data at lower magnification (Table 1), allowing a larger field of view to include more intact particles. Collectively, these strategies allowed us to collect a second dataset with an average of 2–3 intact complexes per micrographs, leading to a set of 15,545 shell particles.

Initial 2D classification of the intact carboxysome complexes was carried out (Figure S5). In these 2D classes, the cargos within the carboxysome shell are clearly ordered and organized into concentric layers, in line with the findings from previous  $\alpha$ -carboxysome studies by electron tomography.<sup>42,44</sup> 3D refinement attempts with this set of particles, without symmetry, failed to converge to interpretable models, with all the particles clustered in a small subset of angle assignments. We therefore carried out a masked 3D classification selectively for the shell (Figure S6), with icosahedral symmetry applied. This led to several classes of particles, of varying diameters from 119 to 123 nm (Figure S6), demonstrating the size heterogeneity of the *Cyanobium*  $\alpha$ -carboxysomes.

We next performed 3D refinement on the most populated class of particles, applying icosahedral symmetry with masking of the internal density. This led to a map of the carboxysome shell at  $\sim 18$  Å resolution (Figures 5A and S6B). At this resolution, the map is largely featureless but still allows us to clearly identify the edges of the icosahedron. We also note that previous studies on synthetic BMC shells have revealed that some pseudo-hexamers form double-layered complexes that protrude from the shell surface.<sup>20,52</sup> Such protrusions made of the pseudo-hexamers CsoS1D were not visible on our reconstruction, which could indicate that it is not the case for CsoS1D. Alternatively, it is possible that CsoS1D is distributed randomly on the shell surface, and therefore double layers are blurred out during reconstruction. A higher-resolution map, obtained without symmetry, would be required to verify this.



**Figure 4. Co-evolution maps of the  $\alpha$ -carboxysome shell proteins**

For each protein pair, the co-evolution map is shown at the top, and an atomic model is at the bottom, colored as in Figure 1. The residues with strong co-evolution correlation are indicated with a red circle and colored in red in the structural model. The co-evolution analysis for CsoS1A and CsoS1E (A), as well as CsoS4A and CsoS4B (B), strongly suggests interhomooligomer interactions. In contrast, in the case of CsoS1E and CsoS4A/B (C), intrahomooligomer interactions are likely identified.

See also Table S2.

### Modeling of the $\alpha$ -carboxysome shell architecture

As indicated above, the resolution of the map of the  $\alpha$ -carboxysome shell is not sufficient to build an atomic model *de novo*. Nonetheless, we used a hybrid approach, by combining this map with the previously elucidated structures of shell proteins and other modeling tools, to propose a structural model of the *Cyanobium*  $\alpha$ -carboxysome shell.

Specifically, we used co-evolution analysis (see STAR Methods for details) to determine the interactions between various shell proteins. We found a strong co-evolution correlation between CsoS1A and CsoS1E and between CsoS4A and CsoS4B (Table S2; Figures 4A and 4B). Mapping the regions with the strongest co-evolution links on the atomic models revealed that they correspond to the homo-oligomer interface (Figures 4A and 4B). These results suggest that  $\alpha$ -carboxysome shell proteins have strong tendencies to form hetero-oligomers, i.e., hexamers formed by a combination of CsoS1A and CsoS1E, and pentamers formed of both CsoS4A and CsoS4B, as demonstrated previously in  $\beta$ -carboxysomes.<sup>53,54</sup>

In addition, we observed a strong co-evolution correlation between CsoS1E and both CsoS4A and CsoS4B (Figure 4C). In contrast, the correlation between CsoS1A and CsoS4A/B was very limited (Table S2). This suggests that the interaction between hexamers and pentamers is formed specifically by CsoS1E, forming the first layer of hexamers around pentamers, while CsoS1A forms predominantly hexamer-hexamer interactions. We note that, alternatively, it is possible that this layer consists of a CsoS1E-CsoS1A hetero-oligomer, with CsoS1E exclusively at the interface with CsoS4A/B.

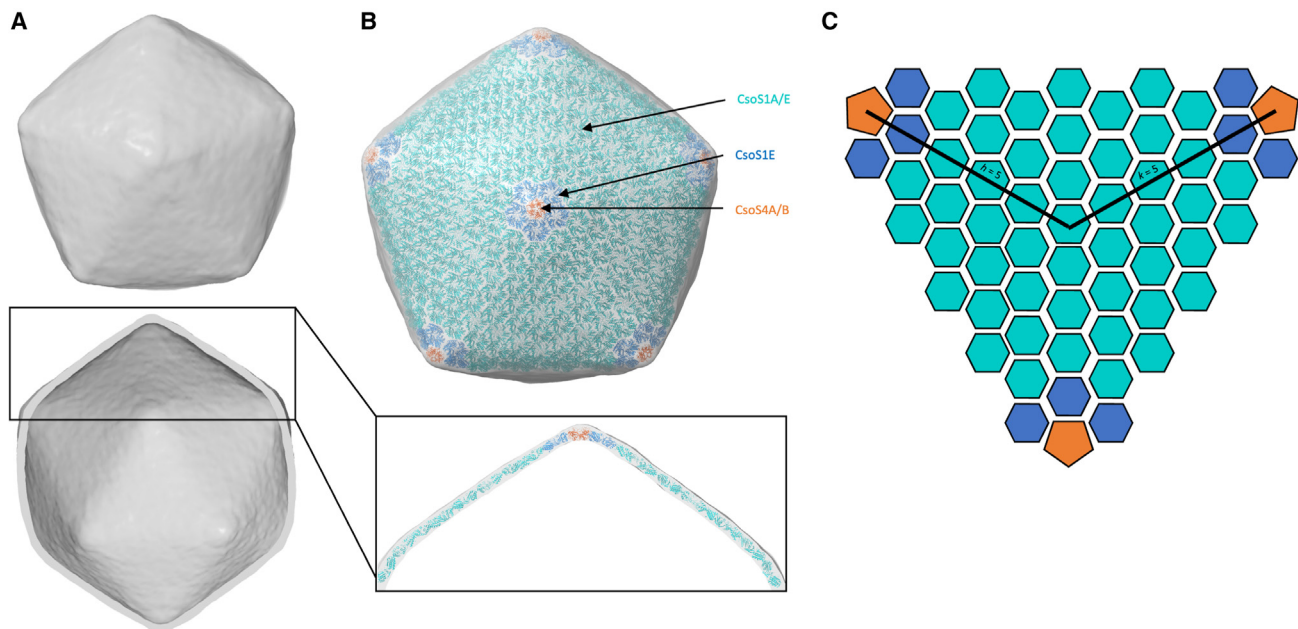
We next combined the hexamer and pentamer orientation derived from the previous structure of a synthetic  $\beta$ -carboxy-

some shell,<sup>16</sup> the low-resolution map of the *Cyanobium*  $\alpha$ -carboxysome shell (Figure 5A), and the co-evolution data (Table S2; Figure 4) to build an atomic model of the intact  $\alpha$ -carboxysome shell (Figure 5B; Video S2). In this model, the  $\alpha$ -carboxysome shell is comprised of 12 pentamers and 750 hexamers. This corresponds to a triangulation (T) number of 75 ( $h = 5$ ;  $k = 5$ ;  $T = h^2 + hk + k^2$ ) (Figure 5C).<sup>55</sup> As indicated above, there is variation in the dimensions of the shell, which likely corresponds to breathing between shell subunits. Further structural analysis, using a much larger number of intact  $\alpha$ -carboxysome particles, is required to verify this interpretation.

Intriguingly, we also observed a very limited co-evolution correlation between CsoS1D and any other shell proteins (Table S2). This could partially be due to its low abundance in the shell,<sup>47</sup> in agreement with SDS-PAGE and mass spectrometry analysis (Figure 2; Table S1), but likely indicates that it is randomly localized in the shell facets. As such, CsoS1D is not included in this structural model. However, this protein was explicitly present within the  $\alpha$ -carboxysome (Table S1). The role and position of CsoS1D within the shell merit further characterization.

### Internal arrangement of enzymes within the $\alpha$ -carboxysome

To further characterize the internal organization of the  $\alpha$ -carboxysomes, we carried out masked 3D refinement on the internal density (Figure S6). We initially attempted reconstructions using a range of symmetries (Figure S7); however, in most cases, this led to the blurring and distortion of features in the obtained maps. Subsequently, we applied masked 3D icosahedral refinements of individual rings of densities observed within the



**Figure 5. Architecture of the  $\alpha$ -carboxysome shell**

(A) Electron potential map of the carboxysome shell, to  $\sim 18$  Å resolution.

(B) Atomic model of the carboxysome shell. Pentamers (likely a mixture of CsoS4A and CsoS4B), located at the vertices of the shell, are shown in orange. They are surrounded by five dimers, probably consisting mostly of CsoS1E, in blue. Additional hexamer layers form the remaining icosahedral face, formed predominantly by CsoS1A, in cyan.

(C) Schematic representation of one face of the carboxysome shell model, colored as in (B). Black lines indicate the unit count for calculation of the T number. See also [Figure S5](#), and [Video S2](#).

carboxysomes. These yielded several reconstructions to  $\sim 18$  Å, with continuous density for each layer, which we termed the outmost, middle, inner, and core layers, respectively ([Figures S6A](#) and [S6B](#)). Notably, all these layers are of a thickness that is similar to the height of RuBisCO ( $\sim 10$  nm) and possess discernible features that broadly resemble RuBisCO's shape. We note, however, that features with 3- and 5-fold symmetry are present in this map but are likely artifacts of the imposed symmetry. Similarly, the fact that some layers, most notably the outmost layer, are significantly shorter ( $\sim 7$  nm) could reflect a symmetry artifact, heterogeneity in protein position, and/or the presence of CA, which is smaller but  $\sim 20\times$  less abundant than RuBisCO according to our mass spectrometry data (see [Table S1](#)).

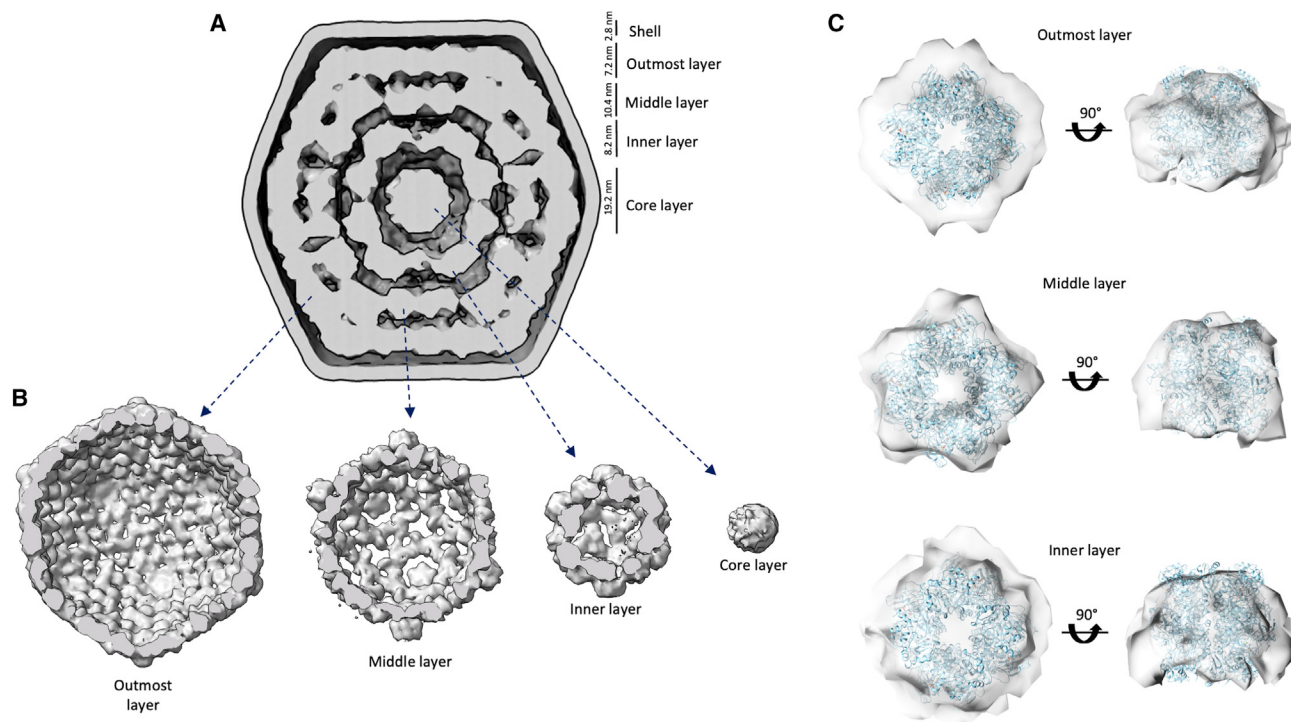
The thickness of each layer, and the presence of features that are compatible with RuBisCO, allowed us to manually place individual complexes in the corresponding density ([Figure 6C](#)), leading to an atomic model of its internal organization within the carboxysome ([Figure 7A](#); [Video S3](#)). In this model, RuBisCO forms concentric layers, and we were able to fit  $\sim 300$  RuBisCO within the internal density (4 in the core layer, 32 in the inner layer, 72 in the middle layer, and 192 in the outmost layer), roughly comparable with previous estimates.<sup>43</sup> Particularly in the middle and outermost layers, gaps with thinner densities are present between RuBisCO molecules, which were not accounted for in our model. It is likely that these gaps accommodate CsoS2 and/or CA proteins; however, the intrinsically disordered structure of CsoS2 and the much smaller size of CA (compared with RuBisCO) did not permit us to model them within the densities.

Our model of the  $\alpha$ -carboxysome internal organization shows two RuBisCO interfaces ([Figure 7B](#)). The first interface corresponds to the contacts between RuBisCO proteins within the same layer and involves interactions on the lateral side of RuBisCO ([Figure 7C](#)). This interaction is mainly mediated via contacts in the variable loop region of the large subunit CbbL ([Figure S8](#)), where the CsoS2 N terminus has been proposed to bind,<sup>40</sup> which awaits further validation. A second interface is formed by the interaction between RuBisCO proteins across the concentric layers in a top-to-bottom configuration ([Figure 7D](#)). In this case, the contacts appear to be largely mediated by two helices in the small subunit CbbS ([Figure S8](#)), although again, the limited resolution does not allow us to unambiguously resolve this. We note that both interfaces correspond to highly conserved regions of the RuBisCO proteins ([Figure S8](#)). This is most evident for the small subunit, where several residues in the top-to-bottom interface are conserved across organisms, and their conservation could indicate a common role in higher-order oligomerization.

## DISCUSSION

In this study, we present a single-particle cryo-EM analysis of an intact  $\alpha$ -carboxysome, purified from endogenous sources. Notably, we report the structure of its RuBisCO to 2.9 Å and observe the presence of unattributed densities on one side, suggesting that another protein is bound to some of the complexes. Using multistep classification, we obtained low-resolution maps of the icosahedral shell and of the internal cargo organization,





**Figure 6. Internal density of the  $\alpha$ -carboxysome structure**

(A) Slab through the overall density, revealing the different internal layers. The height of each layer is indicated.

(B) Individual maps for each layer, obtained from selective masked refinement of the particles used in (A).

(C) RuBisCO structure fitted into the cryo-EM maps generated for the internal density of the carboxysome. The quality of the fit for the outmost layer (top), middle layer (middle), and inner layer (bottom) are shown from the top and side.

See also [Figures S5](#) and [S7](#).

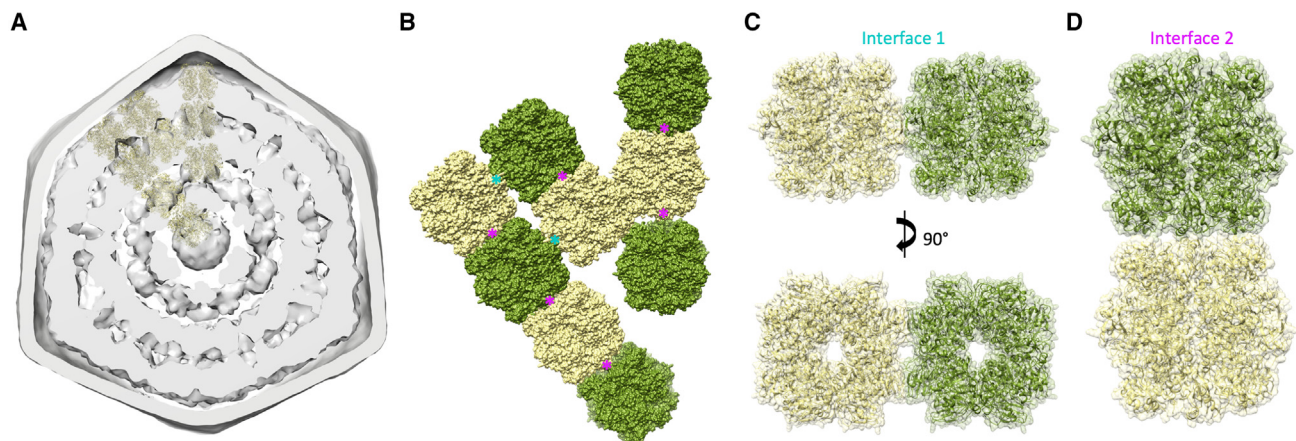
which allows us to propose an atomic model for their respective architecture, through integrative modeling. Collectively, this work provides insights into the architecture of BMCs and their internal organization.

We chose the *Cyanobium*  $\alpha$ -carboxysome as a model system in this study because its structure appears more homogeneous compared with other BMCs studied,<sup>36,41–46,56</sup> as demonstrated in this [\(Figure S1\)](#) and previous studies.<sup>29,48</sup> Our results demonstrate that its shell exhibits an icosahedral symmetry, albeit variable in shape and size, ranging from 119 to 123 nm in diameter ([Figures 2C](#) and [S6](#)). It confirms the common icosahedral architecture of carboxysomes in different species, as observed previously.<sup>42–44</sup>

The model of the internal RuBisCO organization within the  $\alpha$ -carboxysome highlights four concentric layers of cargo enzymes and two main forms (side by side and top to bottom) of RuBisCO-RuBisCO interfaces ([Figures 6](#) and [7](#)). In contrast, recent work using cryoelectron tomography showed that in a distinct  $\alpha$ -carboxysome from *Halothiobacillus neapolitanus*, RuBisCO forms filaments instead of concentric layers.<sup>51</sup> Similarly, a recent cryoelectron tomography analysis of the same *Cyanobium* carboxysome samples provided similar insights, and largely achieved similar conclusions (most notably on the RuBisCO organization within carboxysomes).<sup>57</sup> In those filaments, the interface is highly similar to one of the interfaces identified in our model. This strongly suggests that despite the diversity of  $\alpha$ -carboxysome species, this top-to-bottom interaction is

likely a conserved feature of RuBisCO-RuBisCO association. This conserved interaction is reminiscent of the recent discovery that many metabolic enzymes, such as CTP (cytidine triphosphate) synthase and IMPDH (inosine-5'-monophosphate dehydrogenase), are able to form higher-order assemblies to regulate their activities.<sup>58</sup> Whether the RuBisCO assembly patterns inside the carboxysome could modulate RuBisCO activity merits further investigation. Moreover, it is likely that these filaments aid in the assembly and encapsulation of the shell in collusion with CsoS2. In comparison, RuBisCO enzymes in  $\beta$ -carboxysomes form paracrystalline arrays and exhibit relatively denser packing.<sup>36,46</sup> The discrepancy in the internal organization and copy numbers of RuBisCO within  $\alpha$ - and  $\beta$ -carboxysomes sheds light on their different assembly pathways and encapsulation mechanisms.

The low resolution of the  $\alpha$ -carboxysome map reported here is partly due to the intrinsic heterogeneity and structural plasticity of natural carboxysome structures and internal RuBisCO packing. Given their dynamic and fast assembly, BMC structures are morphologically heterogeneous and vary in size and shape in their native host cells.<sup>2</sup> It has also been shown that the abundance of individual proteins in the  $\beta$ -carboxysome and the size of  $\beta$ -carboxysomes in cyanobacteria is dynamically regulated in response to changing growth conditions.<sup>59</sup> Moreover, the  $\beta$ -carboxysome shell appeared to be mechanically softer than virus capsids, highlighting the flexible nature of the shell architecture.<sup>36</sup> The structural plasticity of BMCs also occurred in



**Figure 7. Internal arrangement of proteins within the  $\alpha$ -carboxysome**

(A) Slab section of the  $\alpha$ -carboxysome electron potential map, with nine RuBisCO complexes fitted in the internal density, in cartoon representation. The height of the complex fits the width and features of the internal density.  
 (B) Surface representation of nine RuBisCO complexes from the internal organization model, in surface representation, and colored alternately in yellow and green. Two distinct inter-RuBisCO interfaces are present, indicated with a cyan and magenta star, respectively.  
 (C and D) Cartoon representation of two adjacent RuBisCO molecules forming the lateral (C) and longitudinal (D) interfaces, shown in cartoon through the transparent surface. The lateral contacts occur through loops in the CbbL subunit, while the longitudinal contacts are mediated by two helices in CbbL and CbbS. See also [Figure S7](#), and [Video S3](#).

protein-protein interactions, such as dynamic self-assembly and correlation between shell protein paralogs to form specific protein assemblies and hetero-oligomers in BMCs.<sup>15,53,54,60</sup> Consistently, our co-evolution analysis suggests that CsoS1A, CsoS1E, CsoS4A, and CsoS4B may form specific assemblies, in which CsoS4A and CsoS4B pentamers sit at the shell vertices, surrounded by CsoS1E proteins that then interact with CsoS1A hexamers ([Figure 5B](#)). It also suggests that the  $\alpha$ -carboxysome shell paralogs CsoS1A and CsoS1E, as well as CsoS4A and CsoS4B, are prone to form hetero-oligomers ([Table S2](#); [Figure 4](#)), as characterized in  $\beta$ -carboxysomes, which could function as a general mechanism for governing the passage of metabolites across the carboxysome shell. These flexible interactions may play vital roles in BMC shell assembly and permeability.

The power of single-particle cryo-EM should make it possible to determine the structure of an intact carboxysome at near-atomic resolution. However, there remains multiple practical challenges for this.<sup>61</sup> Because of the structural heterogeneity mentioned above, a very large number of particles will be required; due to the distinct symmetry between the shell and internal layers, ideally no symmetry would be applied, again increasing the number of particles required for structure determination to high resolution. Additionally, the size of the complex necessitates collecting data with a large field of view, both limiting the attainable resolution and the throughput of data collection. Nonetheless, with the most recent wide-field direct-electron cameras,<sup>62</sup> and with improved automation in data acquisition, this will likely be obtainable in the near future. We also emphasize that tomography approaches will also provide important new insights on the structural diversity of these complexes. While likely not reaching high-resolution structural information, tomography will notably be key to identify and characterize assembly intermediates and will provide a very important tool to understand the heterogeneous nature of carboxysomes.

Recently, extraordinary advances have been made in the acquisition of high-resolution characterization of synthetic BMC minishells.<sup>18–20,52,63</sup> These synthetic shells, with minimal components, exhibit more homogeneous structures and lack any of the internal enzymes, thereby facilitating the alignment of the particles. In contrast, our study on the intact  $\alpha$ -carboxysome structure provides insights into the carboxysome assembly as well as the diversity of BMC architectures and protein compositions. Further characterizations are expected to address how CsoS2 assists with the association of the outer layer of RuBisCO and shell proteins, how CsoS1D and CA are organized within the native  $\alpha$ -carboxysome, and how the internal packing of RuBisCO enzymes is physiologically regulated.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [RESOURCE AVAILABILITY](#)
  - Lead contact
  - Materials availability
  - Data and code availability
- [EXPERIMENTAL AND MODEL SUBJECT DETAILS](#)
- [METHOD DETAILS](#)
  - Cyanobacterial strain growth and carboxysome purification
  - SDS-PAGE and immunoblot analysis
  - RuBisCO assay
  - Mass spectrometry analysis
  - Thin-section electron microscopy
  - Negative-stain TEM grid preparation and screening
  - Cryo-EM grid preparation and data collection

● **QUANTIFICATION AND STATISTICAL ANALYSIS**

- Mass spectrometry data analysis
- Cryo-EM data processing
- Modelling and co-evolution analysis

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at <https://doi.org/10.1016/j.str.2023.03.008>.

**ACKNOWLEDGMENTS**

We acknowledge the Diamond Light Source for access and support of the cryo-EM facilities at the UK's national Electron Bio-imaging Center (eBIC) (under proposal EM-19832). We thank Prof. Ian Prior and Mrs. Alison Beckett for the support of EM and Dr. Deborah Simpson and Dr. Philip J. Brownridge for mass spectrometry analysis. The University of Sheffield FoS cryo-EM facility was used for initial grid preparation and optimization. We are grateful to Dr. Alex Parker for help with generating the supplemental videos. This work was supported by the National Natural Science Foundation of China (32070109); the National Key Research and Development Program of China (2021YFA 0909600); the Biotechnology and Biological Sciences Research Council Grant (BB/R019061/2 to J.R.C.B. and BB/M024202/1, BB/V009729/1, and BB/R003890/1 to L.-N.L.); the Royal Society (RGF/E/A\180233, URFAR\180030, and RGF/E/A\181061 to L.-N.L.); and the Leverhulme Trust (RPG-2021-286 to L.-N.L.).

**AUTHOR CONTRIBUTIONS**

S.L.E. performed the cryo-EM sample preparation and data processing for the shell and internal organization, with help from D.M., as well as the hybrid structural modeling, with help from J.R.C.B. M.M.J.A.-H. performed cell culturing, carboxysome isolation, biochemical characterization, and RuBisCO assays, with help from Y.S. and T.C. G.F.D. performed negative-stain TEM sample preparation and imaging. D.M. and N.S. performed the grid preparation and data processing for the RuBisCO structure. A.J.B. helped with cryo-EM data transfer and processing. S.L.E., L.-N.L., and J.R.C.B. wrote the manuscript, with contributions from all the authors.

**DECLARATION OF INTERESTS**

The authors declare no competing interests.

**INCLUSION AND DIVERSITY**

We support inclusive, diverse, and equitable conduct of research.

Received: July 11, 2022  
Revised: January 19, 2023  
Accepted: March 9, 2023  
Published: April 3, 2023

**REFERENCES**

1. Kerfeld, C.A., Aussignargues, C., Zarzycki, J., Cai, F., and Sutter, M. (2018). Bacterial microcompartments. *Nat. Rev. Microbiol.* *16*, 277–290. <https://doi.org/10.1038/nrmicro.2018.10>.
2. Liu, L.N., Yang, M., Sun, Y., and Yang, J. (2021). Protein stoichiometry, structural plasticity and regulation of bacterial microcompartments. *Curr. Opin. Microbiol.* *63*, 133–141. <https://doi.org/10.1016/j.mib.2021.07.006>.
3. Liu, L.N. (2021). Bacterial metabolosomes: new insights into their structure and bioengineering. *Microb. Biotechnol.* *14*, 88–93. <https://doi.org/10.1111/1751-7915.13740>.
4. Hamley, I.W. (2019). Protein assemblies: nature-inspired and designed nanostructures. *Biomacromolecules* *20*, 1829–1848. <https://doi.org/10.1021/acs.biomac.9b00228>.
5. Schmidt, C., and Urlaub, H. (2017). Combining cryo-electron microscopy (cryo-EM) and cross-linking mass spectrometry (CX-MS) for structural elucidation of large protein assemblies. *Curr. Opin. Struct. Biol.* *46*, 157–168. <https://doi.org/10.1016/j.sbi.2017.10.005>.
6. Nguyen, T.K., and Ueno, T. (2018). Engineering of protein assemblies within cells. *Curr. Opin. Struct. Biol.* *51*, 1–8. <https://doi.org/10.1016/j.sbi.2017.12.005>.
7. Sutter, M., Melnicki, M.R., Schulz, F., Woyke, T., and Kerfeld, C.A. (2021). A catalog of the diversity and ubiquity of bacterial microcompartments. *Nat. Commun.* *12*, 3809. <https://doi.org/10.1038/s41467-021-24126-4>.
8. Axen, S.D., Erbilgin, O., and Kerfeld, C.A. (2014). A taxonomy of bacterial microcompartment loci constructed by a novel scoring method. *PLoS Comput. Biol.* *10*, 1003898. <https://doi.org/10.1371/journal.pcbi.1003898>.
9. Erbilgin, O., McDonald, K.L., and Kerfeld, C.A. (2014). Characterization of a planctomycetal organelle: a novel bacterial microcompartment for the aerobic degradation of plant saccharides. *Appl. Environ. Microbiol.* *80*, 2193–2205. <https://doi.org/10.1128/AEM.03887-13>.
10. Havemann, G.D., and Bobik, T.A. (2003). Protein content of polyhedral organelles involved in coenzyme B 12-dependent degradation of 1,2-propanediol in *Salmonella enterica* serovar typhimurium LT2. *J. Bacteriol.* *185*, 5086–5095. <https://doi.org/10.1128/JB.185.17.5086-5095.2003>.
11. Petit, E., LaTouf, W.G., Coppi, M.V., Warnick, T.A., Currie, D., Romashko, I., Deshpande, S., Haas, K., Alvelo-Maurosa, J.G., Wardman, C., et al. (2013). Involvement of a bacterial microcompartment in the metabolism of fucose and rhamnose by Clostridium phytofermentans. *PLoS One* *8*, e54337. <https://doi.org/10.1371/journal.pone.0054337>.
12. Kofoed, E., Rappleye, C., Stojiljkovic, I., and Roth, J. (1999). The 17-gene ethanolamine (eut) operon of *Salmonella typhimurium* encodes five homologues of carboxysome shell proteins. *J. Bacteriol.* *181*, 5317–5329. <https://doi.org/10.1128/jb.181.17.5317-5329.1999>.
13. Mahinthichaichan, P., Morris, D.M., Wang, Y., Jensen, G.J., and Tajkhorshid, E. (2018). Selective permeability of carboxysome shell pores to anionic molecules. *J. Phys. Chem. B* *122*, 9110–9118. <https://doi.org/10.1021/acs.jpcc.8b06822>.
14. Faulkner, M., Szabó, I., Weetman, S.L., Sicard, F., Huber, R.G., Bond, P.J., Rosta, E., and Liu, L.N. (2020). Molecular simulations unravel the molecular principles that mediate selective permeability of carboxysome shell protein. *Sci. Rep.* *10*, 17501. <https://doi.org/10.1038/s41598-020-74536-5>.
15. Sutter, M., Faulkner, M., Aussignargues, C., Paasch, B.C., Barrett, S., Kerfeld, C.A., and Liu, L.N. (2016). Visualization of bacterial microcompartment facet assembly using high-speed atomic force microscopy. *Nano Lett.* *16*, 1590–1595. <https://doi.org/10.1021/acs.nanolett.5b04259>.
16. Klein, M.G., Zwart, P., Bagby, S.C., Cai, F., Chisholm, S.W., Heinhorst, S., Cannon, G.C., and Kerfeld, C.A. (2009). Identification and structural analysis of a novel carboxysome shell protein with Implications for metabolite transport. *J. Mol. Biol.* *392*, 319–333. <https://doi.org/10.1016/j.jmb.2009.03.056>.
17. Cai, F., Sutter, M., Cameron, J.C., Stanley, D.N., Kinney, J.N., and Kerfeld, C.A. (2013). The structure of CcmP, a tandem bacterial microcompartment domain protein from the  $\beta$ -carboxysome, forms a subcompartment within a microcompartment. *J. Biol. Chem.* *288*, 16055–16063. <https://doi.org/10.1074/jbc.M113.456897>.
18. Tan, Y.Q., Ali, S., Xue, B., Teo, W.Z., Ling, L.H., Go, M.K., Lv, H., Robinson, R.C., Narita, A., and Yew, W.S. (2021). Structure of a minimal  $\alpha$ -carboxysome-derived shell and its utility in enzyme stabilization. *Biomacromolecules* *22*, 4095–4109. <https://doi.org/10.1021/acs.biomac.1c00533>.
19. Kalnins, G., Cesle, E.E., Jansons, J., Liepins, J., Filimonenko, A., and Tars, K. (2020). Encapsulation mechanisms and structural studies of GRM2 bacterial microcompartment particles. *Nat. Commun.* *11*, 388. <https://doi.org/10.1038/s41467-019-14205-y>.
20. Greber, B.J., Sutter, M., and Kerfeld, C.A. (2019). The plasticity of molecular interactions governs bacterial microcompartment shell assembly. *Structure* *27*, 749–763.e4. <https://doi.org/10.1016/j.str.2019.01.017>.

21. Shively, J.M., Ball, F., Brown, D.H., and Saunders, R.E. (1973). Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of thio-bacillus neapolitanus. *Science* 182, 584–586. <https://doi.org/10.1126/science.182.4112.584>.
22. Liu, L.N. (2022). Advances in the bacterial organelles for CO<sub>2</sub> fixation. *Trends Microbiol.* 30, 567–580. <https://doi.org/10.1016/j.tim.2021.10.004>.
23. Raven, J.A., Cockell, C.S., and De La Rocha, C.L. (2008). The evolution of inorganic carbon concentrating mechanisms in photosynthesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 2641–2650. <https://doi.org/10.1098/rstb.2008.0020>.
24. Tcherkez, G.G.B., Farquhar, G.D., and Andrews, T.J. (2006). Despite slow catalysis and confused substrate specificity, all ribulose biphosphate carboxylases may be nearly perfectly optimized. *Proc. Natl. Acad. Sci. USA* 103, 7246–7251. <https://doi.org/10.1073/pnas.0600605103>.
25. Chen, T., Fang, Y., Jiang, Q., Dykes, G.F., Lin, Y., Price, G.D., Long, B.M., and Liu, L.N. (2022). Incorporation of functional rubisco activases into engineered carboxysomes to enhance carbon fixation. *ACS Synth. Biol.* 11, 154–161. <https://doi.org/10.1021/acssynbio.1c00311>.
26. Li, T., Jiang, Q., Huang, J., Aitchison, C.M., Huang, F., Yang, M., Dykes, G.F., He, H.L., Wang, Q., Sprick, R.S., et al. (2020). Reprogramming bacterial protein organelles as a nanoreactor for hydrogen production. *Nat. Commun.* 11, 5448. <https://doi.org/10.1038/s41467-020-19280-0>.
27. Fang, Y., Huang, F., Faulkner, M., Jiang, Q., Dykes, G.F., Yang, M., and Liu, L.N. (2018). Engineering and modulating functional cyanobacterial CO<sub>2</sub>-fixing organelles. *Front. Plant Sci.* 9, 739. <https://doi.org/10.3389/fpls.2018.00739>.
28. Rae, B.D., Long, B.M., Badger, M.R., and Price, G.D. (2013). Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO<sub>2</sub> fixation in cyanobacteria and some proteobacteria. *Microbiol. Mol. Biol. Rev.* 77, 357–379. <https://doi.org/10.1128/mmb.00061-12>.
29. Whitehead, L., Long, B.M., Price, G.D., and Badger, M.R. (2014). Comparing the in vivo function of  $\alpha$ -carboxysomes and  $\beta$ -carboxysomes in two model cyanobacteria. *Plant Physiol.* 165, 398–411. <https://doi.org/10.1104/pp.114.237941>.
30. Kupriyanova, E.v., Sinetova, M.A., Cho, S.M., Park, Y.I., Los, D.A., and Pronina, N.A. (2013). CO<sub>2</sub>-concentrating mechanism in cyanobacterial photosynthesis: organization, physiological role, and evolutionary origin. *Photosynth. Res.* 117, 133–146. <https://doi.org/10.1007/s11120-013-9860-z>.
31. Huang, F., Kong, W.W., Sun, Y., Chen, T., Dykes, G.F., Jiang, Y.L., and Liu, L.N. (2020). Rubisco accumulation factor 1 (Raf1) plays essential roles in mediating Rubisco assembly and carboxysome biogenesis. *Proc. Natl. Acad. Sci. USA* 117, 17418–17428. <https://doi.org/10.1073/pnas.2007990117>.
32. Taylor, T.C., and Andersson, I. (1997). Structure of a product complex of spinach ribulose-1,5-bisphosphate carboxylase/oxygenase. *Biochemistry* 36, 4041–4046. <https://doi.org/10.1021/bi962818w>.
33. Newman, J., and Gutteridge, S. (1994). Structure of an effector-induced inactivated state of ribulose 1,5-bisphosphate carboxylase/oxygenase: the binary complex between enzyme and xylulose 1,5-bisphosphate. *Structure* 2, 495–502. [https://doi.org/10.1016/S0969-2126\(00\)00050-2](https://doi.org/10.1016/S0969-2126(00)00050-2).
34. Matsumura, H., Shiomi, K., Yamamoto, A., Taketani, Y., Kobayashi, N., Yoshizawa, T., Tanaka, S.I., Yoshikawa, H., Endo, M., and Fukuyama, H. (2020). Hybrid rubisco with complete replacement of rice rubisco small subunits by sorghum counterparts confers C4 plant-like high catalytic activity. *Mol. Plant* 13, 1570–1581. <https://doi.org/10.1016/j.molp.2020.08.012>.
35. Banda, D.M., Pereira, J.H., Liu, A.K., Orr, D.J., Hammel, M., He, C., Parry, M.A.J., Carmo-Silva, E., Adams, P.D., Banfield, J.F., and Shih, P.M. (2020). Novel bacterial clade reveals origin of form I Rubisco. *Nat. Plants* 6, 1158–1166. <https://doi.org/10.1038/s41477-020-00762-4>.
36. Faulkner, M., Rodriguez-Ramos, J., Dykes, G.F., Owen, S.V., Casella, S., Simpson, D.M., Beynon, R.J., and Liu, L.N. (2017). Direct characterization of the native structure and mechanics of cyanobacterial carboxysomes. *Nanoscale* 9, 10662–10673. <https://doi.org/10.1039/c7nr02524f>.
37. Cameron, J.C., Wilson, S.C., Bernstein, S.L., and Kerfeld, C.A. (2013). Biogenesis of a bacterial organelle: the carboxysome assembly pathway. *Cell* 155, 1131–1140. <https://doi.org/10.1016/j.cell.2013.10.044>.
38. Wang, H., Yan, X., Aigner, H., Bracher, A., Nguyen, N.D., Hee, W.Y., Long, B.M., Price, G.D., Hartl, F.U., and Hayer-Hartl, M. (2019). Rubisco condensate formation by CcmM in  $\beta$ -carboxysome biogenesis. *Nature* 566, 131–135. <https://doi.org/10.1038/s41586-019-0880-5>.
39. Kaneko, Y., Danev, R., Nagayama, K., and Nakamoto, H. (2006). Intact carboxysomes in a cyanobacterial cell visualized by Hilbert differential contrast transmission electron microscopy. *J. Bacteriol.* 188, 805–808. <https://doi.org/10.1128/JB.188.2.805-808.2006>.
40. Oltrogge, L.M., Chaijarasphong, T., Chen, A.W., Bolin, E.R., Marqusee, S., and Savage, D.F. (2020). Multivalent interactions between CsoS2 and Rubisco mediate  $\alpha$ -carboxysome formation. *Nat. Struct. Mol. Biol.* 27, 281–287. <https://doi.org/10.1038/s41594-020-0387-7>.
41. Iancu, C.V., Morris, D.M., Dou, Z., Heinhorst, S., Cannon, G.C., and Jensen, G.J. (2009). Organization, structure, and assembly of  $\alpha$ -carboxysomes determined by electron cryotomography of intact cells. <https://doi.org/10.1016/j.jmb.2009.11.019>.
42. Schmid, M.F., Paredes, A.M., Khant, H.A., Soyer, F., Aldrich, H.C., Chiu, W., and Shively, J.M. (2006). Structure of Halothiobacillus neapolitanus carboxysomes by cryo-electron tomography. *J. Mol. Biol.* 364, 526–535. <https://doi.org/10.1016/j.jmb.2006.09.024>.
43. Iancu, C.V., Ding, H.J., Morris, D.M., Dias, D.P., Gonzales, A.D., Martino, A., and Jensen, G.J. (2007). The structure of isolated Synechococcus strain WH8102 carboxysomes as revealed by electron cryotomography. *J. Mol. Biol.* 372, 764–773. <https://doi.org/10.1016/j.jmb.2007.06.059>.
44. Dai, W., Chen, M., Myers, C., Ludtke, S.J., Pettitt, B.M., King, J.A., Schmid, M.F., and Chiu, W. (2018). Visualizing individual RuBisCO and its assembly into carboxysomes in marine cyanobacteria by cryo-electron tomography. *J. Mol. Biol.* 430, 4156–4167. <https://doi.org/10.1016/j.jmb.2018.08.013>.
45. Hantke, M.F., Hasse, D., Maia, F.R.N.C., Ekeberg, T., John, K., Svenda, M., Loh, N.D., Martin, A.V., Timneanu, N., Larsson, D.S.D., et al. (2014). High-throughput imaging of heterogeneous cell organelles with an X-ray laser. *Nat. Photonics* 8, 943–949. <https://doi.org/10.1038/nphoton.2014.270>.
46. Sun, Y., Wollman, A.J.M., Huang, F., Leake, M.C., and Liu, L.N. (2019). Single-organelle quantification reveals stoichiometric and structural variability of carboxysomes dependent on the environment. *Plant Cell* 31, 1648–1664. <https://doi.org/10.1105/tpc.18.00787>.
47. Sun, Y., Harman, V.M., Johnson, J.R., Chen, T., Dykes, G.F., Lin, Y., Beynon, R.J., and Liu, L.-N. (2022). Decoding the absolute stoichiometric composition and structural plasticity of  $\alpha$ -carboxysomes. Preprint at bioRxiv. <https://doi.org/10.1101/2021.12.06.471529>.
48. Long, B.M., Hee, W.Y., Sharwood, R.E., Rae, B.D., Kaines, S., Lim, Y.L., Nguyen, N.D., Massey, B., Bala, S., von Caemmerer, S., et al. (2018). Carboxysome encapsulation of the CO<sub>2</sub>-fixing enzyme Rubisco in tobacco chloroplasts. *Nat. Commun.* 9, 3570. <https://doi.org/10.1038/s41467-018-06044-0>.
49. Duff, A.P., Andrews, T.J., and Curmi, P.M. (2000). The transition between the open and closed states of rubisco is triggered by the inter-phosphate distance of the bound bisphosphate. *J. Mol. Biol.* 298, 903–916. <https://doi.org/10.1006/jmbi.2000.3724>.
50. Ting, C.S., Hsieh, C., Sundararaman, S., Mannella, C., and Marko, M. (2007). Cryo-electron tomography reveals the comparative three-dimensional architecture of Prochlorococcus, a globally important marine cyanobacterium. *J. Bacteriol.* 189, 4485–4493. <https://doi.org/10.1128/JB.01948-06>.
51. Metskas, L.A., Wilfong, R., and Jensen, G.J. (2022). Subtomogram averaging for biophysical analysis and supramolecular context. *J. Struct. Biol.* X 6, 100076. <https://doi.org/10.1016/J.YJSBX.2022.100076>.

52. Sutter, M., Greber, B., Aussignargues, C., and Kerfeld, C.A. (2017). Assembly principles and structure of a 6.5-MDa bacterial microcompartment shell. *Science* 356, 1293–1297. <https://doi.org/10.1126/science.aan3289>.
53. Sommer, M., Sutter, M., Gupta, S., Kirst, H., Turmo, A., Lechno-Yossef, S., Burton, R.L., Saechao, C., Sloan, N.B., Cheng, X., et al. (2019). Heterohexamers formed by CcmK3 and CcmK4 increase the complexity of beta carboxysome shells. *Plant Physiol.* 179, 156–167. <https://doi.org/10.1104/pp.18.01190>.
54. Garcia-Alles, L.F., Root, K., Maveyraud, L., Aubry, N., Lesniewska, E., Mourey, L., Zenobi, R., and Truan, G. (2019). Occurrence and stability of hetero-hexamer associations formed by  $\beta$ -carboxysome CcmK shell components. *PLoS One* 14, e0223877. <https://doi.org/10.1371/journal.pone.0223877>.
55. Caspar, D.L., and Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* 27, 1–24. <https://doi.org/10.1101/SQB.1962.027.001.005>.
56. Yang, M., Simpson, D.M., Wenner, N., Brownridge, P., Harman, V.M., Hinton, J.C.D., Beynon, R.J., and Liu, L.N. (2020). Decoding the stoichiometric composition and organisation of bacterial metabolosomes. *Nat. Commun.* 11, 1976. <https://doi.org/10.1038/s41467-020-15888-4>.
57. Ni, T., Sun, Y., Burn, W., Al-Hazeem, M.M.J., Zhu, Y., Yu, X., Liu, L.N., and Zhang, P. (2022). Structure and assembly of cargo Rubisco in two native  $\alpha$ -carboxysomes. *Nat. Commun.* 13, 4299. <https://doi.org/10.1038/s41467-022-32004-w>.
58. Lynch, E.M., Kollman, J.M., and Webb, B.A. (2020). Filament formation by metabolic enzymes—a new twist on regulation. *Curr. Opin. Cell Biol.* 66, 28–33. <https://doi.org/10.1016/J.CEB.2020.04.006>.
59. Menon, B.B., Heinhorst, S., Shively, J.M., and Cannon, G.C. (2010). The carboxysome shell is permeable to protons. *J. Bacteriol.* 192, 5881–5886. <https://doi.org/10.1128/JB.00903-10>.
60. Faulkner, M., Zhao, L.S., Barrett, S., and Liu, L.N. (2019). Self-assembly stability and variability of bacterial microcompartment shell proteins in response to the environmental change. *Nanoscale Res. Lett.* 14, 54. <https://doi.org/10.1186/s11671-019-2884-3>.
61. Burton-Smith, R.N., and Murata, K. (2021). Cryo-electron microscopy of the giant viruses. *Microscopy* 70, 477–486. <https://doi.org/10.1093/jmicro/dfab036>.
62. Nakane, T., Kotecha, A., Sente, A., McMullan, G., Masiulis, S., Brown, P.M.G.E., Grigoras, I.T., Malinauskaite, L., Malinauskas, T., Miehling, J., et al. (2020). Single-particle cryo-EM at atomic resolution. Preprint at bioRxiv. <https://doi.org/10.1101/2020.05.22.110189>.
63. Sutter, M., Laughlin, T.G., Sloan, N.B., Serwas, D., Davies, K.M., and Kerfeld, C.A. (2019). Structure of a synthetic  $\beta$ -carboxysome shell. *Plant Physiol.* 181, 1050–1058. <https://doi.org/10.1104/pp.19.00885>.
64. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). CryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* 14, 290–296. <https://doi.org/10.1038/nmeth.4169>.
65. Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* 74, 531–544. <https://doi.org/10.1107/S2059798318006551>.
66. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82. <https://doi.org/10.1002/pro.3943>.
67. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
68. Jing, X., Zeng, H., Wang, S., and Xu, J. (2020). A web-based protocol for interprotein contact prediction by deep learning. *Methods Mol. Biol.* 2074, 67–80.
69. Xu, J., McPartlon, M., and Li, J. (2021). Improved protein structure prediction by deep learning irrespective of co-evolution information. *Nat. Mach. Intell.* 3, 601–609. <https://doi.org/10.1038/s42256-021-00348-5>.
70. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). Brief Communication ColabFold: making protein folding accessible to all. *Nat. Methods* 19, 679–682. <https://doi.org/10.1038/s41592-022-01488-1>.
71. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. <https://doi.org/10.1002/jcc.20084>.
72. Schrodinger, L., and DeLano, W. (2020). PyMol. [www.pymol.org/pymol](http://www.pymol.org/pymol).

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Bacterial and virus strains</b>		
<i>Cyanobium</i> sp. PCC 7001	This study	<a href="https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=TaxonDetail&amp;page=taxonDetail&amp;taxon_oid=647533126">https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=TaxonDetail&amp;page=taxonDetail&amp;taxon_oid=647533126</a>
<b>Deposited data</b>		
Map of the carboxysome shell	This study	14379
Map of the carboxysome outermost layer	This study	14382
Map of the carboxysome middle layer	This study	14381
Map of the carboxysome innermost layer	This study	14380
Map of the carboxysome core layer	This study	14377
Map of RuBisCO with D4 symmetry	This study	14385
Coordinates of RuBisCO with D4 symmetry	This study	7YYO
Map of RuBisCO with C1 symmetry	This study	14376
Coordinates of RuBisCO with C1 symmetry	This study	8CMY
<b>Software and algorithms</b>		
CryoSPARC	Punjani et al., Nature Methods, 2017 <sup>64</sup>	<a href="https://cryosparc.com/">https://cryosparc.com/</a>
Phenix	Afonine et al., ACTA D, 2018 <sup>65</sup>	<a href="https://phenix-online.org/">https://phenix-online.org/</a>
ChimeraX	Pettersen et al., Protein Sci, 2021 <sup>66</sup>	<a href="https://www.cgl.ucsf.edu/chimerax/">https://www.cgl.ucsf.edu/chimerax/</a>
PyMol	The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.	<a href="https://pymol.org/2/">https://pymol.org/2/</a>
AlphaFold	Jumper et al., Nature, 2021 <sup>67</sup>	<a href="https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/AlphaFold2_advanced.ipynb">https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/AlphaFold2_advanced.ipynb</a>
RaptorX	Jing et al., Methods Mol Biol, 2020 <sup>68</sup>	<a href="http://raptorx.uchicago.edu/">http://raptorx.uchicago.edu/</a>

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Julien Bergeron ([julien.bergeron@kcl.ac.uk](mailto:julien.bergeron@kcl.ac.uk)).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

The structure of the *Cyanobium* sp. PCC 7001 RuBisCO enzyme with D4 symmetry has been deposited to the PDB (PDB: 7YYO), and the corresponding 2.9 Å cryo-EM map was deposited to the EMDB (EMDB: 14385). The structure obtained without imposing any symmetry has been deposited to the PDB (PDB: 8CMY), and the corresponding 3.8 Å map was also deposited to the EMDB (EMDB: 14376). The maps of the carboxysome shell, and of each individual internal layer, have been deposited to the EMDB (EMDB: 14379, 14382, 14381, 14380, 14377).

This paper does not report original code. Any additional information required to reanalyse the data reported in this paper is available from the [lead contact](#) upon request.

### EXPERIMENTAL AND MODEL SUBJECT DETAILS

*Cyanobium* sp. PCC 7001 strains were cultured in BG-11 medium as described in the corresponding [STAR Methods](#).

## METHOD DETAILS

### Cyanobacterial strain growth and carboxysome purification

*Cyanobium* sp. PCC 7001 (Pasteur Culture Collection of Cyanobacteria, PCC) cells were grown in 4 L of BG-11 medium under constant illumination at 30°C with constant stirring and bubbling with air. Carboxysomes were purified as described previously with modifications. Cells were collected by centrifugation (6000 g, 10 min) and resuspended in TEB buffer (5 mM Tris-HCL, pH 8.0, 1 mM EDTA, 20 mM NaHCO<sub>3</sub>) with additional 0.55 M mannitol and 60 kU rLysozyme (Sigma-Aldrich, United States). Cells were then incubated overnight (20 h) with gentle shaking at 30°C in the dark, and were collected via centrifugation (6000 g, 10 min). Cells were placed on ice and resuspended in 20 mL ice-cold TEB containing an additional 5 mL 1  $\mu$ m Silicone disruption beads. Cells were broken via bead beating for 8 mins in one-minute intervals of vortex, and 1 min on ice. Broken cells were separated from the beads, and the total resuspension volume was increased to 40 mL with TEB buffer containing an additional 4% IGEPAL CA-630 (Sigma-Aldrich, United States) were mixed on a rotating shaker overnight at 4°C. Unbroken cells were pelleted via centrifugation at 3,000 g for 5 mins, and the supernatant was centrifuged at 40,000 g for 20 mins. The pellet was then resuspended in 40 mL TEB containing 4% IGEPAL CA-630 and centrifuged again at 40,000 x g for 20 mins. The resulting pellet was then resuspended in 2 mL TEB + 10mM MgCl<sub>2</sub> (TEMB) (5 mM Tris-HCL, pH 8.0, 1 mM EDTA, 10 mM MgCl<sub>2</sub>, 20 mM NaHCO<sub>3</sub>) and centrifuged at 5000 x g for 5 mins before loading onto a 20-60% (v/v) sucrose gradient in TEMB buffer. Gradients were then centrifuged at 105,000 g for 60 mins at 4°C; the milky band at the 40%-50% interface was collected, diluted in 10 mL TEMB buffer and centrifuged again at 105,000 g for 60 mins. The final carboxysome pellet was then resuspended in 150  $\mu$ L TEMB for the following structural and biochemical analysis.

### SDS-PAGE and immunoblot analysis

Isolated carboxysomes were diluted to 5 mg·mL<sup>-1</sup> and denatured using 4X Bromophenol blue buffer (Fisher Scientific, United States). The samples were heated at 95°C for 10 mins, and insoluble debris was pelleted via short spin. Approximately 50  $\mu$ g proteins were loaded onto 15% (v/v) denaturing SDS-PAGE gels and stained using Coomassie Brilliant Blue G-250 (ThermoFisher Scientific, UK). Immunoblot analyses were performed using anti-CbbL (1:10,000 dilution, Agrisera, AS03 037, Sweden), anti-CsoS1 from *H. neapolitanus* (1:5000 dilution, Agrisera, AS14 2760, Sweden), and horseradish peroxidase-conjugated goat antirabbit immunoglobulin G secondary antibody (1:10,000 dilution, Agrisera AS101461, Sweden). Images were taken using a Quant LAS 4000 platform (GE Healthcare Life Sciences, USA).

### RuBisCO assay

RuBisCO activities of isolated carboxysomes were determined as described previously with minor modifications.<sup>25,31,36</sup> Isolated  $\alpha$ -carboxysomes were diluted to 0.5 mg mL<sup>-1</sup> in (100 mM EPPS, pH 8.0; 20 mM MgCl<sub>2</sub>) and 5  $\mu$ L was added to scintillation vials containing NaH<sup>14</sup>CO<sub>3</sub> with a range of concentrations (1.5-48 mM). and incubated at 37 °C for 2 mins before the addition of D-ribulose 1,5-bisphosphate sodium salt hydrate (RuBP, Sigma Aldrich, US) final concentration 0.04 mM. The reaction was carried out for 5 mins before being terminated by adding 2:1 by volume 10% formic acid. Samples were dried for at least 30 mins at 95 °C to remove unfixed <sup>14</sup>C before re-suspending the fixed <sup>14</sup>C pellets with ultra-pure water and adding 2 mL of scintillation cocktail (Ultima Gold XR, PerkinElmer, US). Radioactivity measurements were then performed using a scintillation counter (Tri-Carb, PerkinElmer, US). Raw readings were used to calculate the amount of fixed <sup>14</sup>C, and then converted to the total carbon fixation rates. RuBisCO activity. Data are presented as mean  $\pm$  standard deviation (SD) based on three biological replicates isolated from independent culture batches, and were analyzed using OriginPro 2020b (OriginLab, Massachusetts, USA).

### Mass spectrometry analysis

The isolated  $\alpha$ -carboxysome samples were washed with PBS buffer. Rapigest was added to a final concentration of 0.05% (w/v) into the sample for 10-min incubation at 80°C. The sample was then reduced with dithiothreitol (3 mM, final concentration) for 10 mins at 60°C, alkylated with iodoacetamide (9 mM, final concentration) for 30 min at room temperature in the dark, followed by digestion with trypsin at 37°C overnight. Digestion was terminated with 1  $\mu$ L of trifluoroacetic acid (TFA). Data-dependent LC-MS/MS analysis was conducted on a QExactive quadrupole-Orbitrap mass spectrometer coupled to a Dionex Ultimate 3000 RSLC nano-liquid chromatograph (Hemel Hempstead, UK). A 2  $\mu$ L sample digest was loaded onto a trapping column (Acclaim PepMap 100 C18, 75  $\mu$ m  $\times$  2 cm, 3  $\mu$ m packing material, 100 Å) in 0.1% TFA, 2% acetonitrile H<sub>2</sub>O, and set in line with the analytical column (EASY-Spray PepMap RSLC C18, 75  $\mu$ m  $\times$  50 cm, 2  $\mu$ m packing material, 100 Å). Peptides were eluted using a linear gradient of 96.2% buffer A (0.1% formic acid):3.8% buffer B (0.1% formic acid in water:acetonitrile 80:20, v/v) to 50% buffer A:50% buffer B over 30 mins at 300 nL min<sup>-1</sup>. The mass spectrometry analysis was operated in DDA mode with survey scans between  $m/z$  300-2000 acquired at a mass resolution of 70,000 (FWHM) at  $m/z$  200. The maximum injection time was 250 ms, and the automatic gain control was set to 1e<sup>6</sup>. Fragmentation of the peptides was performed by higher-energy collisional dissociation using a normalized collision energy of 30%. Dynamic exclusion of  $m/z$  values to prevent repeated fragmentation of the same peptide was used with an exclusion time of 20 seconds.

### Thin-section electron microscopy

Cyanobacterial cell cultures were pelleted by centrifugation (6,000 g, 10 min) and processed for thin section using a Pelco BioWave Pro laboratory microwave system. The cells are first fixed with 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer at pH 7.2 using

two steps of 100W. After agarose embedding, samples were then stained with 2% osmium tetroxide and 3% Potassium Ferrocyanide using three steps of 100W. The osmium stain was set using 1% thiocarbohydrazide and 2% osmium tetroxide. The samples were stained with 2% uranyl acetate, prior to dehydration by increasing alcohol concentrations (from 30 to 100%) and resin embedding. Thin sections of 70 nm were cut with a diamond knife and poststained with 3% lead citrate.

### Negative-stain TEM grid preparation and screening

Isolated  $\alpha$ -carboxysome samples were immobilized onto the glow-discharged grids and then were stained with 2% uranyl acetate. EM imaging was conducted using an FEI Tecnai G2 Spirit BioTWIN transmission electron microscope equipped with a Gatan Rio 16 camera.

### Cryo-EM grid preparation and data collection

For the structural characterisation of RuBisCO, 3  $\mu$ L aliquots of purified  $\alpha$ -carboxysomes at a concentration of  $\sim 1$  mg $\cdot$ mL $^{-1}$  were applied to Graphene Oxide coated, 300 mesh, 2/2  $\mu$ m hole/spacing, holey carbon grids (EMR). A Leica EM GP Automatic Plunge Freezer (Leica) was used to plunge freeze the sample, blotting for 3–6 s. Cryo-EM data was collected with a 300 kV Titan Krios TEM, equipped with a Falcon 3 direct electron detector (Thermo Fisher) operated in linear mode. 4593 micrographs were collected using the EPU software (Thermo Fisher) with a pixel size of 1.11  $\text{\AA}$  pix $^{-1}$ , a total dose rate of 30 e $^{-}$   $\text{\AA}^{-2}$ , and 44 fractions per micrograph. The defocus range was  $-0.5$  to  $-1.5$   $\mu$ m.

For structural characterisation of the intact  $\alpha$ -carboxysome complex, 3  $\mu$ L aliquots of purified sample at a concentration of 3 mg $\cdot$ mL $^{-1}$  were applied to Graphene Oxide coated grids, 300 mesh, 2/2  $\mu$ m hole/spacing, holey carbon grids (EMR). A Leica EM GP Automatic Plunge Freezer (Leica) was used to plunge freeze, blotting for 6 s. Cryo-EM data were collected with a 300 kV Titan Krios TEM with a Falcon 3 direct electron detector (FEI) operated in counting mode. 5429 micrographs were collected using EPU software (Thermo Fisher) with a pixel size of 2.23  $\text{\AA}$  pix $^{-1}$  with a total dose rate of 29.7 e $^{-}$   $\text{\AA}^{-2}$  with 33 frames per micrograph. The defocus range was  $-1.0$  to  $-2.2$   $\mu$ m.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Mass spectrometry data analysis

The raw data files were imported into Progenesis QI for Proteomics v4 (Nonlinear Dynamics, Newcastle upon Tyne UK, a Waters Company). The chromatograms are aligned and normalised prior to label-free quantification. Peptide identification was performed by Mascot (v2.8, Matrix Science, UK) against the Uniprot reference *Cyanobium* sp. PCC 7001 database (UP000003950, 2771 proteins). A precursor mass tolerance of 10 ppm and a fragment ion mass tolerance of 0.01 Da were applied with the dynamic modification of Oxidation (M) and with the static modification of carbamidomethylation (C).

### Cryo-EM data processing

All the cryo-EM data processing steps were carried out in CryoSPARC<sup>64</sup> v.3.1.0.

For the RuBisCO structure, automated particle picking was initially used, leading to a dataset of  $\sim 2,800,000$  particles. 2D classification was employed to select particles that clearly correspond to RuBisCO, leading to a final set of 131,356 particles. 3D refinement was performed with these, with D4 symmetry, converging to a map at 2.87  $\text{\AA}$  resolution. The same set of particles was also refined without symmetry imposed, leading to a second map at 3.79  $\text{\AA}$  resolution.

For intact carboxysomes, 131 particles were manually picked from selected micrographs to generate 2D classes subsequently used for template picking for the entire dataset. A total of 15,545 particles were picked and extracted using a 700x700 pixels box. After multiple rounds of 2D classification 8,701 particles from the best 2D classes were selected and used to generate an initial model. Particles were downsampled to a box size of 168x168 pixels for 3D classifications and reconstructions. A reconstruction of the entire carboxysome was generated in I symmetry. Masked classifications of the shell were carried out with C1 symmetry to give a reconstruction at 19  $\text{\AA}$  resolution. Heterogeneous refinements of the carboxysome shell used for model building were carried out with I symmetry to give reconstructions of  $\sim 18$   $\text{\AA}$ .

### Modelling and co-evolution analysis

Atomic models of the CsoS1A and CsoS1E hexamers, the CsoS4A and CsoS4B pentamers, and the CsoS1D trimer were generated with AlphaFold2 (ColabFold).<sup>64,69</sup> The co-evolution analyses were performed using the RaptorX server,<sup>69</sup> with contact probabilities  $> 0.5$  considered to be significant.

To build the *Cyanobium* sp. PCC 7001 RuBisCO structure, an initial atomic model was built for both CbbL and CbbS with AlphaFold2 (ColabFold),<sup>67,70</sup> and 8 copies of each were placed at their respective location on the EM map. The coordinates for the substrate and Mg ion were added manually, and the termini without visible density were deleted. The model was then subjected to real-space refinement in Phenix.<sup>65</sup>

The difference map was calculated by first generating a volume of the RuBisCO structure, and then subtracting this volume from the C1 reconstruction, in ChimeraX.<sup>66</sup>

To generate the atomic model of the shell, a CsoS4a pentamer was placed in one corner of the map icosahedron, using the orientation reported previously in the structure of the  $\beta$ -carboxysome synthetic shell<sup>20</sup> to determine the outward face. Five copies of the



CsoS1E hexamer were placed around it, again using the  $\beta$ -carboxysome structure to determine the outward face. Additional copies of the CsoS1A hexamers were next placed manually, forming additional continuous layers around until five faces of the carboxysome shell were complete. The model was manually curated in Chimera<sup>71</sup> and used for global energy minimisation refinement in Phenix.

For the internal density, copies of the *Cyanobium* sp. PCC 7001 RuBisCO structure were placed in regions of the map of the different shells, and fitted manually in Chimera. If major clashes were observed between adjacent molecules, that with the less optimal fit to the density was removed.

All structural figures were generated in either PyMol,<sup>72</sup> Chimera, or ChimeraX.