New Phytologist

Research

Assembly mechanism of the β-carboxysome shell mediated by the chaperone CcmS

Jing Li[1]*, Jia-Xin Deng[1]*, Xin Chen[2]*, Bo Li[1], Bo-Rui Li[1], Zhong-Liang Zhu[1], Jiexi Liu[2], Yuxing Chen[1], Hualing Mi[2], Cong-Zhao Zhou[1] and Yong-Liang Jiang[1]

[1]School of Life Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, 230027, China; [2]National Key Laboratory of Plant Molecular Genetics, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Science, 300 Fenglin Road, Shanghai, 200032, China

Authors for correspondence:
Hualing Mi
Email: mihl@cemps.ac.cn

Cong-Zhao Zhou
Email: zcz@ustc.edu.cn

Yong-Liang Jiang
Email: jyl@ustc.edu.cn

Received: *10 January 2025*
Accepted: *27 February 2025*

New Phytologist (2025)
doi: 10.1111/nph.70086

Key words: carboxysome, chaperone, CO_2-concentrating mechanism, cyanobacteria, shell proteins, *Synechocystis* sp. PCC 6803.

Summary

- Carboxysomes are self-assembled bacterial microcompartments (BMCs) that encapsulate the enzymes RuBisCO and carbonic anhydrase into a proteinaceous shell, enhancing the efficiency of photosynthetic carbon fixation. The chaperone CcmS was reported to participate in the assembly of β-carboxysomes; however, the underlying molecular mechanism remains elusive.
- We report the crystal structure of CcmS from *Synechocystis* sp. PCC 6803, revealing a monomer of α/β fold. Moreover, its complex structures with two types of BMC hexamers, CcmK1 homohexamer and CcmK1-CcmK2 heterohexamer, reveal a same pattern of CcmS binding to the featured C-terminal segment of CcmK1.
- Upon binding to CcmS, this C-terminal segment of CcmK1 is folded into an amphipathic α-helix protruding outward that might function as a hinge to crosslink adjacent BMC-H hexamers, thereby facilitating concerted and precise assembly of the β-carboxysome shell. Deletion of the *ccmS* gene or the 8-residue C-terminal coding region of *ccmK1* resulted in the formation of aberrant and fewer carboxysomes, suppressed photosynthetic capacity in *Synechocystis* sp. PCC 6803.
- These findings enable us to propose a putative model for the chaperone-assisted assembly of β-carboxysome shell and provide clues for the design and engineering of efficient carbon fixation machinery.

Introduction

Photosynthesis, one of the most important biochemical reactions on Earth, provides energy and organic materials for almost all living organisms (Hayer-Hartl & Hartl, 2020). The enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), the most abundant protein in nature (Bar-On & Milo, 2019), catalyzes the carboxylation of ribulose 1,5-bisphosphate (RuBP) to produce two molecules of 3-phosphoglycerate (3-PGA) (Miziorko & Lorimer, 1983). However, RuBisCO exhibits a low carboxylation efficiency (Zhao *et al.*, 2024). Additionally, it can recognize O_2 and perform an oxygenase reaction during photorespiration, leading to a significant loss of carbon fixation (Feller *et al.*, 2008; Hagemann & Bauwe, 2016; Walker *et al.*, 2016).

Cyanobacteria and many autotrophs, including algae and C4 plants, have developed a CO_2-concentrating mechanism (CCM) to enhance the carbon fixation efficiency of RuBisCO (Hennacy & Jonikas, 2020). The cyanobacterial CCM comprises inorganic carbon (HCO_3^- and CO_2) transport systems and a core component termed the carboxysome (Kupriyanova *et al.*, 2023). Carboxysomes are well-investigated bacterial microcompartments (BMCs) of *c.* 100–400 nm in diameter that encapsulate the cargo enzymes RuBisCO and carbonic anhydrase (CA) within an icosahedral-like proteinaceous shell (Stewart *et al.*, 2021; Sutter *et al.*, 2022). Carboxysomes are present in two distinct evolutionary lineages: α-carboxysomes, which globally dominate most aquatic habitats (Cabello-Yeves *et al.*, 2022); and β-carboxysomes, which are distributed mainly in freshwater/estuarine cyanobacteria (Rae *et al.*, 2013; Whitehead *et al.*, 2014; Nguyen *et al.*, 2023). In both α- and β-carboxysomes, the 20 facets of the shell are self-assembled by numerous hexameric proteins BMC-H of the Pfam00936 domain and the trimeric proteins BMC-T of tandem Pfam00936 domains (Sommer *et al.*, 2017; Kerfeld *et al.*, 2018; Lee *et al.*, 2019). The 12 vertices are sealed by the pentameric protein BMC-P of the Pfam03319 domain, forming an intact icosahedral-like shell (Tanaka *et al.*, 2008; Kerfeld *et al.*, 2018).

The major building blocks of the β-carboxysome shell are composed of variable BMC-H paralogs CcmKs (Sommer *et al.*, 2017). The model cyanobacterium *Synechocystis* sp. PCC 6803 (hereafter *Syn*6803) contains four BMC-H paralogs, CcmK1-4, and two BMC-T proteins, CcmP and CcmO

*These authors contributed equally to this work.

(Kaneko *et al.*, 1996; Kerfeld *et al.*, 2005; Tanaka *et al.*, 2009; Larsson *et al.*, 2017; Wang *et al.*, 2022). Among them, CcmK1 and CcmK2, whose coding genes are within a single *ccm* operon, are highly conserved shell proteins that constitute the majority of the shell (Tanaka *et al.*, 2009; Sommer *et al.*, 2017; Wang *et al.*, 2022), whereas CcmK3, CcmK4, CcmP, and CcmO are minor shell proteins that decorate the shell. These minor shell proteins were proposed to confer plasticity to the shell for adapting to environmental changes (Rae *et al.*, 2012; Cameron *et al.*, 2013; Garcia-Alles *et al.*, 2019; Wang *et al.*, 2022). The BMC-H protein adopts a conserved α/β fold and forms hexamers with a disk-like shape, characterized by a concave and convex side (Yeates *et al.*, 2010, 2011). The BMC-T protein assembles into a trimer/pseudohexamer resembling the structure of the BMC-H hexamer (Ochoa & Yeates, 2021). Notably, BMC-H proteins can also form chimeric heterohexamers, such as CcmK1-CcmK2 and CcmK3-CcmK4 heterohexamers (Garcia-Alles *et al.*, 2019; Sommer *et al.*, 2019). All BMC-H homo/heterohexamers and BMC-T trimers feature a central pore with distinct polarities, which are presumed to modulate the translocation of substrates and products across the carboxysome shell (Tsai *et al.*, 2007; Samborska & Kimber, 2012; Sutter *et al.*, 2019; Faulkner *et al.*, 2020). Phylogenetic analysis indicated that the major shell proteins CcmK1 and CcmK2 are grouped into the same branch (Sommer *et al.*, 2017), indicating their functional redundancy and complementarity in β-carboxysome assembly. Despite sharing > 90% sequence identity with CcmK2, CcmK1 features an 8-residue C-terminal extension (Kerfeld *et al.*, 2005; Tanaka *et al.*, 2008, 2009). Previous structural studies on C-terminal deletion mutants of *Syn*6803 CcmK1 demonstrated that the C-terminus tends to participate in protein–protein interactions during β-carboxysome assembly (Tanaka *et al.*, 2009).

The assembly of β-carboxysomes was proposed to follow a 'core-first' model, in which the inner cargo interactions initiate the assembly process, while the final morphology is jointly regulated by both shell and cargo components (Rotskoff & Geissler, 2018; Trettel *et al.*, 2024). Specifically, the biogenesis of β-carboxysomes begins with the condensation of RuBisCO and CA, which form the inner core, followed by the recruitment of shell proteins (Wang *et al.*, 2019; Zang *et al.*, 2021; Wang & Hayer-Hartl, 2023). The scaffolding proteins CcmM and CcmN play central roles in mediating the assembly and maturation of β-carboxysomes (Wang *et al.*, 2019; Sun *et al.*, 2021). Atomic force microscopy (AFM) studies of purified CcmK proteins revealed shell patches that might be precursors of the β-carboxysome shell (Mahalik *et al.*, 2016; Garcia-Alles *et al.*, 2017, 2019). Moreover, the structures of individual shell components and recombinant mini-shells have provided detailed information on the interaction patterns among shell oligomers (Cai *et al.*, 2016; Dai *et al.*, 2018; Garcia-Alles *et al.*, 2019; Sommer *et al.*, 2019; Sutter *et al.*, 2019; Tan *et al.*, 2021; Ni *et al.*, 2023). Our recently reported cryo-electron microscopy structure of *Prochlorococcus* intact α-carboxysomes demonstrated the fine assembly pattern of the α-carboxysome shell, which is reinforced by the scaffolding protein CsoS2 (Zhou *et al.*, 2024). Despite extensive structural studies of shell proteins (Kerfeld *et al.*, 2005; Tanaka *et al.*, 2009; Samborska & Kimber, 2012; Garcia-Alles *et al.*, 2017, 2019, 2023; Sommer *et al.*, 2019), the detailed mechanism underlying the assembly of various CcmK homo/hetero-oligomers in the β-carboxysome shell remains elusive.

Our previous study identified a new carboxysomal protein, CcmS, which binds to the shell protein CcmK1 (Chen *et al.*, 2023). Deletion of *ccmS* results in aberrant carboxysome formation and suppressed photosynthetic capacity, leading to a slow-growth phenotype, particularly under CO_2-limited conditions (Chen *et al.*, 2023). The recently reported crystal structure of *Anabaena* sp. PCC 7120 (termed *Ana*7120 for short) CcmS revealed that CcmS functions as a chaperone that specifically binds to the C-terminal extension of CcmK1 (Cheng *et al.*, 2024). These findings suggested that CcmS is required for the coordinated assembly of β-carboxysomes, maintaining proper CCM function. However, the precise role of CcmS in the assembly and functionality of β-carboxysomes remains unclear. In this study, we solved the crystal structures of CcmS and its two complexes with the CcmK1 homohexamer and the CcmK1-CcmK2 heterohexamer. Structural analysis revealed that the chaperone CcmS protects the protruding configuration of the C-terminal segment of CcmK1, which might function as a hinge that cross-links the adjacent CcmK hexamers and further stabilizes the shell. Together with previous reports, we propose a model for the assembly of the β-carboxysome shell mediated by the chaperone CcmS. These findings provide new insights into the biogenesis of β-carboxysomes and may guide the design of efficient carbon fixation machinery.

Materials and Methods

Cloning, plasmids, and strains

The genes encoding CcmS (Pro15-Ala142) with a deletion of the N-terminal 14 residues, CcmK1, CcmK2, CcmK3, CcmK4, CcmL, CcmP, and CcmO were amplified via PCR from the genomic DNA of *Synechocystis* sp. PCC 6803 T. All mutants were generated using a standard PCR-based strategy. The CcmS gene encoding residues Pro15-Ala142 was cloned and inserted into the pET22a expression vector with a C-terminal 6×His tag, while the remaining genes were cloned and inserted into the pCDFDuet expression vector. The sequences of the cyanobacterial strains, plasmids, and proteins used in this study are listed in Supporting Information Table S1.

Protein expression and purification

Both the wild-type (WT) and mutant proteins were overexpressed individually in the *Escherichia coli* strain BL21 (DE3). The *E. coli* cells were cultured in LB media supplemented with the corresponding antibiotics (50 μg ml^{-1} ampicillin or 100 μg ml^{-1} spectinomycin) at 37°C. Protein expression was induced by adding 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) when the optical density at 600 nm (OD_{600}) reached 0.6–0.8. The cells were then incubated at 16°C on a shaker for 20–24 h. The cells were harvested by

centrifugation at 7700 ***g*** for 3 min. The suspension was flash-frozen in liquid nitrogen and stored at −80°C for future use.

The cells expressing CcmS were resuspended in buffer A (20 mM Tris–HCl pH 7.5, 300 mM NaCl) and lysed using ultrasonication for 30 min to facilitate cell disruption. After centrifugation at 17 000 ***g*** for 30 min, the supernatant containing the target protein was loaded onto a Ni-NTA column (GE Healthcare, Uppsala, Sweden) preequilibrated with buffer A. Protein elution was achieved using buffer A supplemented with 0.5 M imidazole. The eluted proteins were further purified by gel filtration using a Superdex 200 pg column (GE Healthcare) preequilibrated with buffer A.

The CcmS-CcmK1 complex was obtained by mixing the suspensions of the *E. coli* cells expressing His_6-tagged CcmS and the cells expressing 1×Flag-tagged CcmK1, followed by resuspension in buffer B (20 mM Tris–HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, and 5% glycerol). The CcmS-CcmK1-CcmK2 complex was obtained by mixing the suspensions of the *E. coli* cells expressing His_6-tagged CcmS and the cells coexpressing CcmK1 and 3×Flag-tagged CcmK2, followed by resuspension in buffer B. Then the mixed cells were lysed by ultrasonication for 30 min. After centrifugation, the supernatant was loaded onto a Ni-NTA column (GE Healthcare), and the column was washed with buffer B. Contaminants were then removed using the wash buffer supplemented with 20 mM imidazole. Protein elution was achieved using buffer B supplemented with 0.5 M imidazole. The eluted proteins were further purified by gel filtration using a Superdex 200 pg column (GE Healthcare) in buffer B. The obtained protein was then concentrated for further use. The CcmK1-CcmK2 proteins were expressed and purified in the same manner as the CcmS-CcmK1-CcmK2 complex.

For the pull-down experiment, suspensions of cells expressing His_6-tagged CcmS and cells expressing shell proteins or mutants were mixed with buffer B. The cells were subsequently lysed by ultrasonication for 30 min to facilitate cell disruption. After centrifugation at 17 000 ***g*** for 30 min, the supernatant containing the target protein was loaded onto a Ni-NTA column (GE Healthcare) preequilibrated with buffer B. Protein elution was achieved using buffer B supplemented with 0.5 M imidazole. Finally, the protein complexes were evaluated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) to assess protein interactions.

Size exclusion chromatography with multiangle light scattering

Size exclusion chromatography with multiangle light scattering (SEC-MALS) was used to determine the molecular weight of CcmS in the solution. The assay was performed using a Superdex 200 10/300 GL column connected to the DAWN HELEOS II light scattering detector (Wyatt Technology, Santa Barbara, CA, USA) and the Optilab T-rEx refractive index detector (Wyatt Technology). The protein samples (100 μl, 1.0 mg ml^{-1}) were injected into and then eluted from the column pre-equilibrated with the buffer 20 mM Tris–HCl pH 8.0, 150 mM NaCl. The results were recorded and processed by ASTRA 7.0.1 software (Wyatt Technology). The final figures were prepared using the GraphPad Prism 8 software.

Mass photometry assays

Mass photometry experiments were conducted using Refeyn TwoMP instruments (Refeyn Ltd, Oxford, UK) to determine the molar masses of CcmS-CcmK1 and CcmS-CcmK1-CcmK2 complexes (Young *et al.*, 2018). Microscope coverslips (Thorlabs, Newton, NJ, USA) were assembled into the flow chamber, and silicone gaskets (Grace Bio-Labs, Bend, OR, USA) were positioned on the glass surface to facilitate sample loading. The gaskets were designed with 3 × 2 wells to accommodate the sample drops, ensuring proper alignment and containment before measurements. Contrast-to-mass calibration was performed by measuring the contrast of bovine serum albumin. This calibration standard was used to establish a reference curve, which was then applied to each sample measurement to calculate the molecular mass corresponding to each histogram distribution during data analysis.

The CcmS-CcmK1 or CcmS-CcmK1-CcmK2 proteins were diluted to a final concentration of 500 nM using the buffer 20 mM Tris–HCl pH 8.0, 150 mM NaCl, and 1 mM EDTA. For each acquisition, 2 μl of the diluted protein sample was added to a well containing 15 μl of freshly prepared working buffer (20 mM Tris–HCl pH 8.0, 150 mM NaCl, and 1 mM EDTA) pre-equilibrated at room temperature. Measurements were performed using the standard image acquisition mode with each recording lasting 60 s (6000 frames). The acquired data were processed and analyzed using Refeyn DiscoverMP software, where each histogram was fitted to a Gaussian distribution to determine the molecular mass (kDa) and normalized counts.

Redox assays of CcmS

The purified CcmS protein and its mutants, in the buffer 20 mM Tris–HCl pH 8.0, 150 mM NaCl, were divided into two equal parts, respectively. One part was treated with 0.5 mM $CuCl_2$, and the other with 0.1 M β-mercaptoethanol, followed by incubation on ice for 30 min. Afterward, denatured polyacrylamide gel electrophoresis (SDS-PAGE) and Native-PAGE analyses were performed to evaluate the formation of intermolecular disulfide bonds in CcmS. Notably, the loading buffer for both SDS-PAGE and Native-PAGE was free of any reducing agents.

Crystallization and structure determination

Crystals of CcmS were grown at 16°C by hanging drop vapor diffusion with 1 μl of 9 mg ml^{-1} protein solution and 1 μl of reservoir solution. The crystals were grown in 0.1 M Tris pH 7.5, 25% PEG 6000, and 0.6 M LiCl. The crystals were soaked in 500 mM KI for 1 min. The crystal structure of CcmS was determined by the single-wavelength anomalous dispersion method. Crystals of CcmS-CcmK1 were grown at 16°C by hanging drop vapor diffusion with 0.1 μl of 14 mg ml^{-1} protein solution and 0.1 μl of reservoir solution. The crystals were grown in 0.1 M MES pH 6.5, 0.2 M Na acetate, and 2 M NaCl. The

crystal structure of the CcmS-CcmK1 complex was determined by a molecular replacement method using the previously solved structure of CcmK1 from *Syn*6803 (PDB: 3BN4) and the crystal structure of CcmS as a model. Crystals of CcmS-CcmK1-CcmK2 were grown at 16°C by sitting drop vapor diffusion with 1 µl of 22 mg ml^{-1} protein solution and 1 µl of reservoir solution. The crystals were grown in 0.1 M Tris pH 8.5 and 0.7 M ammonium tartrate dibasic. The crystal structure of CcmS-CcmK1-CcmK2 was determined by molecular replacement using the structures of *Syn*6803 CcmK1 (PDB: 3BN4), CcmK2 (PDB: 2A1B), and CcmS as search models.

Single-crystal X-ray diffraction measurements were performed on an XtaLAB PRO diffractometer at 1.5406 Å and 100 K using Cu X-rays generated by an MMF007 rotating-anode X-ray (Rigaku, Japan) with a Pilatus 200K detector at the Core Facility Center for Life Sciences, USTC. Data processing and reduction were carried out using HKL2000 (Otwinowski & Minor, 1997) and CryAlisPro (v.39.35c) (Matsumoto *et al.*, 2021). The crystal structures were refined by the maximum likelihood method implemented in REFMAC5 (Murshudov *et al.*, 2011) as part of the CCP4i (Winn *et al.*, 2011) program suite. Iterative model building was performed using the Coot program (Emsley & Cowtan, 2004). The final models were evaluated using MolProbity (Chen *et al.*, 2010). The search and analysis of protein folding types were conducted using SCOPe (Fox *et al.*, 2014; Chandonia *et al.*, 2022). All interface areas were calculated using PDBsum (McDonald & Thornton, 1994; Wallace *et al.*, 1995), and all structural figures were prepared using PyMOL (https://pymol.org/2/). The models of CcmK1 and CcmK3 hexamers are predicted by AlphaFold3 (Abramson *et al.*, 2024). A list of parameters for data collection, processing, structure determination, and refinement is provided in Table S2.

Cultivation of cyanobacteria

Wild-type *Syn*6803 and mutant *Syn*6803 cells were cultured in BG11 media supplemented with 20 mM TES-NaOH pH 8.0 at 30°C under growth light (40 µmol photons m^{-2} s^{-1}). The cells were grown under two different conditions: one with 4% CO_2 in air (referred to as HC) and the other with ambient air (referred to as LC). The BG11 solid medium was supplemented with 1.5% agar. The corresponding antibiotics were added to the BG11 media during the cultivation of the mutants (*ΔccmK1* and *ΔccmK2*: 10 µg ml^{-1} kanamycin; *ccmK1ΔC8* and *ΔccmS*: 1 µg ml^{-1} gentamycin).

Construction of mutant lines

The *ΔccmS*, *ΔccmK1*, and *ccmK1ΔC8* mutant strains were generated as reported previously (Chen *et al.*, 2023). The upstream and downstream regions of sll1028 (*ccmK2*) were amplified by PCR, along with amplification of the cassette encoding a kanamycin resistance gene. These fragments were then ligated into the pMD19T vector via homologous recombination. The pMD19T plasmid was subsequently transformed into WT *Syn*6803 cells, and the transformed cells were spread onto solid BG11 media supplemented with 10 µg ml^{-1} kanamycin. Incubation was conducted in 4% (v/v) CO_2 in air. The deletion of *ccmK2* was validated by PCR.

Transmission electron microscopy

*Syn*6803 cells were centrifuged at 4300 ***g*** for 3 min. The pellet was resuspended in 0.1 M PBS pH 7.2, 5% glutaraldehyde, and 4% paraformaldehyde and incubated at room temperature for 3 h. Thereafter, the cells were transferred to 4°C for 24 h. The fixed samples were washed three times with 0.1 M PBS pH 7.2 and then fixed overnight in 2% OsO_4 at room temperature. The fixed samples were dehydrated in a series of ethanol, followed by three washes with acetone. The samples were then infiltrated with a series of epoxy resins and embedded in epoxy resin. Ultrathin sections were cut using a diamond knife and mounted on copper grids. The sections were stained with 2% uranyl acetate for 10 min, followed by lead citrate staining for 2 min. Imaging was performed using a transmission electron microscope (FEI Tecnai G2 Spirit 120 kV).

CO_2 uptake measurement

CO_2 uptake was measured by using a portable Li-6400 open-flow gas exchange system (LI-COR Biosciences, Lincoln, NE, USA). Thirty microliters of each cell suspension was placed on BG11 solid media. The CO_2 concentration was controlled at 400 µmol mol^{-1}. The values represent the mean ± SE of three independent measurements.

Oxygen exchange

The cells were cultured in BG11 media pH 8.0 bubbled with air at 30°C under growth light (40 µmol photons m^{-2} s^{-1}) until they reached the mid-logarithmic phase. Photosynthetic oxygen evolution was determined in the BG11 medium that contained the cell at *c.* 2.5 µg Chl ml^{-1} with a Clark-type oxygen electrode. The entire monitoring process was performed at 30°C. The values represent the mean ± SE of four independent measurements.

Phylogenetic and conservation analysis

The complete cyanobacterial genomes used for phylogenetic analyses were obtained from the National Center for Biotechnology Information (NCBI) genome database (May 2024), which contains 2064 cyanobacterial species and 3738 genomes. RbcL, CcmK1, CcmK2, and CcmS in these cyanobacteria were identified by alignment with the NCBI nr database using DIAMOND (Buchfink *et al.*, 2021). The CcmK2, CcmK1, and CcmS sequences were aligned with Mafft v.7.515 using the G-INS-i algorithm (–maxiterate 1000), and sequence logos were generated with WebLogo3 (Crooks *et al.*, 2004; Katoh & Standley, 2013). Nonredundant RbcL sequences were clustered at 100% identity using CD-Hit v.4.8.1 (-c 1) (Fu *et al.*, 2012). A maximum likelihood tree was constructed using IQ Tree v.2.2.2.3 after

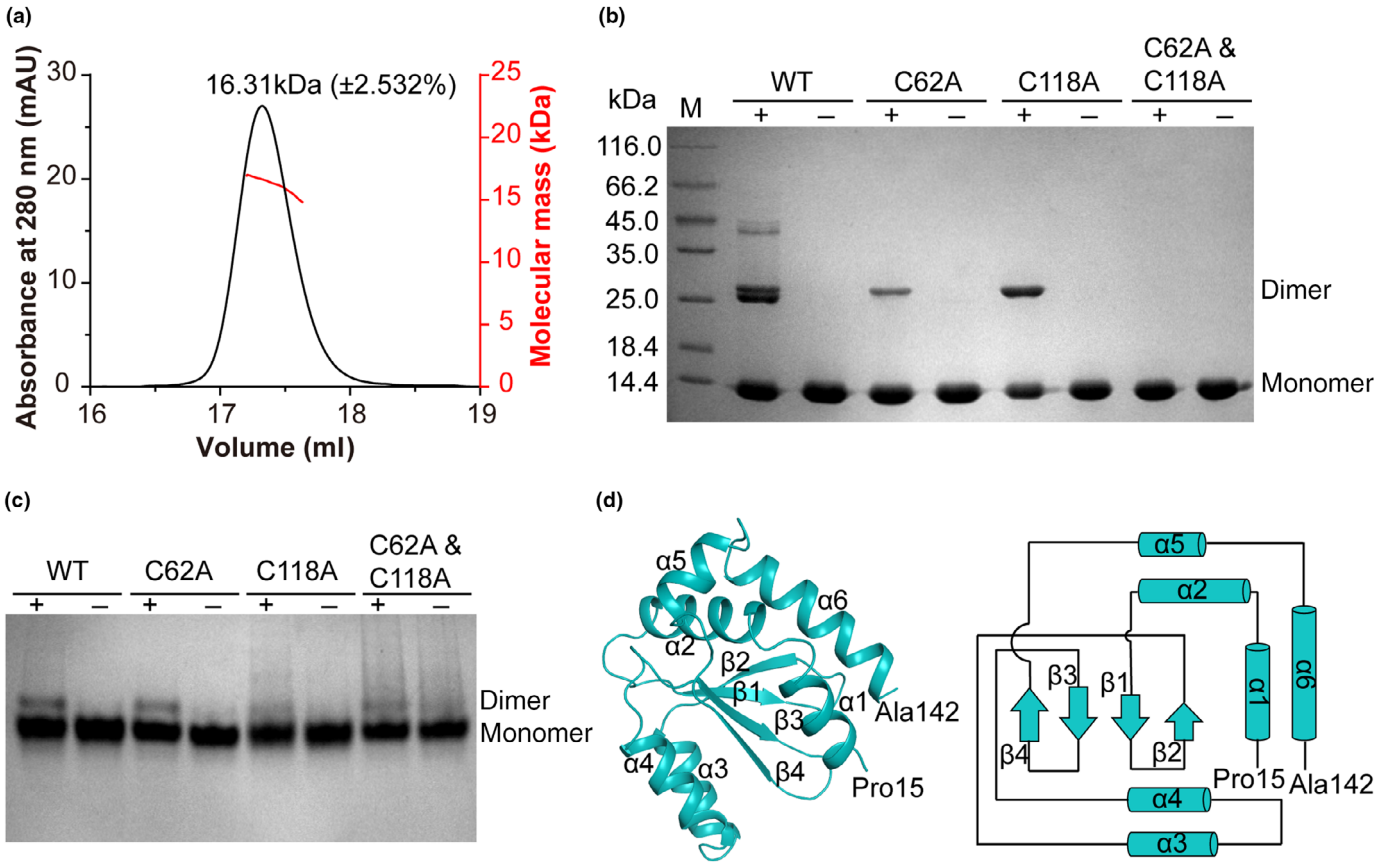

Fig. 1 Biochemical and structural analyses of *Syn*6803 CcmS. (a) Size exclusion chromatography with multiangle light scattering assays of *Syn*6803 CcmS. The absorbance is shown on the left *y*-axis, whereas the molecular weight represented by the jagged short line is shown on the right *y*-axis. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (b) and Native-PAGE (c) of wild-type CcmS and its mutants. CcmS in the presence of $CuCl_2$ and β-mercaptoethanol are shown as '+' and '−', respectively. (d) Cartoon representation (left) and topology diagram (right) of the CcmS structure, which are colored teal. The secondary structural elements and terminal residues are labeled.

alignment with MAFFT (Katoh & Standley, 2013; Minh *et al.*, 2020). The output tree was visualized using iTOL (Letunic & Bork, 2021). The CcmK1 and CcmS proteins of 132 species for the phylogenetic analysis were obtained by CcmK1 sequence-based redundancy reduction of species from the aforementioned evolutionary tree.

Results

Synechocystis CcmS is a monomer of α/β fold

The SEC-MALS technique indicated that CcmS in solution mainly exists as a monomer of *c.*16 kDa, which is comparable to the theoretical mass of 15.9 kDa (Figs 1a, S1a). Moreover, SDS-PAGE and Native-PAGE analyses in the presence of the oxidant Cu^{2+} showed that the majority of CcmS exists as a monomer (Fig. 1b,c), in addition to a minor portion of dimers that were observed in our previous report (Chen *et al.*, 2023). Notably, the dimer is completely dissociated upon the addition of β-mercaptoethanol. Indeed, sequence analysis showed that CcmS contains two cysteine residues, Cys62 and Cys118, which are highly conserved among CcmS homologs (Fig. S1b). Mutations of either Cys62 or Cys118 to alanine led to much less dimer formation compared to the WT protein under oxidative conditions. The double mutation of both cysteine residues nearly completely abolished the dimer formation (Fig. 1b,c), although a very faint dimeric band remained, likely resulting from nonspecific interactions between CcmS monomers.

Using single-wavelength anomalous dispersion, we determined the crystal structure of CcmS at 2.35 Å resolution, with each asymmetric unit containing five CcmS molecules. Crystal packing analysis revealed that the two cystine residues are positioned far away from each other and the maximum contact area among CcmS molecules is only *c.* 450 $Å^2$, which is insufficient to support the formation of a stable dimer. This further confirms that *Syn*6803 CcmS adopts a monomeric structure. Notably, both Cys62 and Cys118 are exposed on the surface of the CcmS structure (Figs 1d, S1b), which might form nonspecific intermolecular disulfide bonds under oxidative conditions.

CcmS adopts a three-layered α/β/α fold composed of a central four-stranded mixed β-sheet (β1–β4) sandwiched by helices α3 and α4 on one side and helices α1-2 and α5-α6 on the other side, and the central β-sheet follows a 2-1-3-4 order (Fig. 1d). The structure of *Syn*6803 CcmS closely resembles the recently

reported structure of *Ana*7120 CcmS (PDB: 8ZLH) (Cheng *et al.*, 2024), with a root mean square deviation (RMSD) of 0.834 Å for the 101 Cα atoms (Fig. S1c). However, *Syn*6803 CcmS exists mainly as a monomer, in contrast to the dimeric structure of *Ana*7120 CcmS, which is stabilized by main-chain hydrogen bonds between the β4 strands of two subunits, as well as salt bridges involving Glu111-Arg16 and Glu111-Lys104 (Cheng *et al.*, 2024). Further sequence analysis showed that, despite sharing a sequence identity of 36%, the residues involved in the dimeric interface of *Ana*7120 CcmS are not conserved in *Syn*6803 CcmS (Fig. S1b). Particularly, the key residue Glu111 for *Ana*7120 CcmS dimerization is replaced by a hydrophobic residue Ile112 in *Syn*6803 CcmS. These variations make the different CcmS: monomer in *Syn*6803 and dimer in *Ana*7120.

CcmS binds to the CcmK1 hexamer by specifically recognizing its C-terminal hinge domain

Our previous study showed that CcmS interacts with the shell protein CcmK1 (Chen *et al.*, 2023). To investigate whether CcmS could also bind to other shell proteins, we performed pull-down assays using His_6-tagged CcmS with various β-carboxysome shell proteins from *Syn*6803. The results showed that CcmS can interact with CcmK1 but not with other shell proteins (Fig. 2a). Deletion of the unique C-terminal eight residues (termed C-tail for short) of CcmK1 completely abolished its interaction with CcmS (Fig. 2b). In addition, fusing the C-tail of CcmK1 to the C-terminus of CcmK2 enabled CcmK2 to bind to CcmS (Fig. 2b). These results suggest that the C-tail of CcmK1 is indispensable for binding to CcmS.

To further elucidate the interaction pattern between CcmS and CcmK1, we purified the CcmS-CcmK1 complex. Gel filtration chromatography and SDS-PAGE analysis revealed that CcmS forms a stable complex with CcmK1 (Fig. S2a). Then, we crystallized the CcmS-CcmK1 complex and solved its structure at 2.95 Å resolution. The structure revealed that two CcmS molecules bind to subunits A and D of one CcmK1 hexamer, forming a '2 + 6' binding pattern (Fig. 2c). This complex possesses a theoretical molecular weight of 112 kDa, which is comparable to that of the calculated mass of 107 kDa (Fig. S2b). CcmS binding to the CcmK1 hexamer did not induce notable conformational changes in either the CcmS or CcmK1 hexamer, with RMSDs of 0.593 Å for the 113 Cα atoms in CcmS and 1.403 Å for the 512 Cα atoms in the CcmK1 hexamer, respectively. CcmS binds to the C-terminal segment (residues Glu95-Arg111) of CcmK1, which is clearly resolved as an α-helix (α3) followed by a C-tail loop in the density map, whereas these residues are devoid of density in the other four CcmK1 subunits (Fig. 2c). Overall, the binding of CcmS stabilizes the C-terminal segment of CcmK1, forming an individual domain composed of α3 and the C-tail (termed the hinge domain) that protrudes outward from the core structure.

In the complex structure, each CcmS binds specifically to the hinge domain of one CcmK1 subunit (Fig. 2c). Two distinct interfaces were identified for CcmS binding to the CcmK1 hexamer, yielding a total buried interface area of *c.* 2000 $Å^2$ (Fig. 2c,d). The first interface (Interface I) involves the insertion of α3 and the C-tail of CcmK1 into the groove of CcmS formed by the central β-sheet and helices α2-3, contributing to *c.* 1000 $Å^2$ of the interface area (Fig. 2d). The α3 of CcmK1 aligns parallel to the α3 of CcmS, forming extensive hydrophobic interactions. Notably, the C-tail of CcmK1 binds to the α2 and β2 of CcmS via numerous polar interactions, including several pairs of hydrogen bonds and salt bridges, such as Arg110-Gln41, Ile108-Trp42, Arg110-Glu46, Asn105-Ser61, Arg101-Asp52, and Arg101-His54 (Fig. 2d). In particular, Arg110 of CcmK1 makes a group of interactions with CcmS, including a cation-π interaction with Trp42 of CcmS (Fig. 2d). Mutation of Arg110 to alanine in CcmK1 almost completely abolished its ability to bind to CcmS (Fig. S2c). Similarly, the triple mutation of residues Gln41, Trp42, and Glu46 to alanine in CcmS also abolished its CcmK1-binding capability (Fig. S2c). These results confirmed the critical role of these residues in mediating the interactions between CcmK1 and CcmS. The second interface of *c.* 1000 $Å^2$ (Interface II) involves the formation of an extended β-sheet formed by β5 of CcmK1 and the central β-sheet of CcmS (Fig. 2d). Upon CcmS binding, the His82-Val88 residues of CcmK1 undergo drastic structural rearrangement, transitioning from a short α-helix to a β-strand (β5). Consequently, the β5 of CcmK1 aligns antiparallel to the β4 of CcmS, forming an extended five-stranded β-sheet. Main-chain hydrogen bonds between β5 of CcmK1 and β4 of CcmS were found to stabilize Interface II (Fig. 2c,d). Notably, Arg77 of CcmS forms two hydrogen bonds with Glu83 of the adjacent CcmK1 subunit (Fig. 2d), further stabilizing the interface.

Upon CcmS binding, the hinge domain of CcmK1 undergoes conformational changes, and flips outward from the core structure, resulting in a reduction in the interface area between the CcmK1 subunits from 1000 to 680 $Å^2$ (Fig. S3a). Compared to the CcmK1 hexamer, the CcmS-bound CcmK1 hexamer adopts a flatter conformation, which deviates *c.* 10° along the central axis (Fig. S3b). However, despite these changes, the central pore remains similar in size, *c.* 5.1 Å in the CcmS-bound CcmK1 hexamer, which is comparable to the 4.8 Å observed in the CcmS-free CcmK1 hexamer (Kinney *et al.*, 2011).

Compared to the previously reported complex structure of *Ana*7120 CcmS bound to the C-terminal 15-residue peptide of CcmK1 (termed $CcmK1^{C15}$) (Cheng *et al.*, 2024), *Syn*6803 CcmS adopts a similar pattern binding to the CcmK1 C-tail, emphasizing the essential role of this C-tail in facilitating CcmK1-CcmS interaction (Fig. S3c). However, the C-tail of *Syn*6803 CcmK1 consists of eight residues, which is shorter than the 12-residue C-tail found in *Ana*7120 CcmK1 (Fig. S4). In the *Ana*7120 CcmS-$CcmK1^{C15}$ complex structure, the $CcmK1^{C15}$ peptide interacts with two CcmS dimers. One interaction mode resembles that observed in the *Syn*6803 CcmS-CcmK1 complex, whereas the additional C-terminus (Arg108-Pro114) of *Ana*7120 CcmK1 also interacts with the neighboring CcmS dimer (Cheng *et al.*, 2024). Given the CcmS subunit in the two complex structures aligned, one CcmK1 subunit in our complex partially overlaps with the other subunit of CcmS dimer in the *Ana*7120

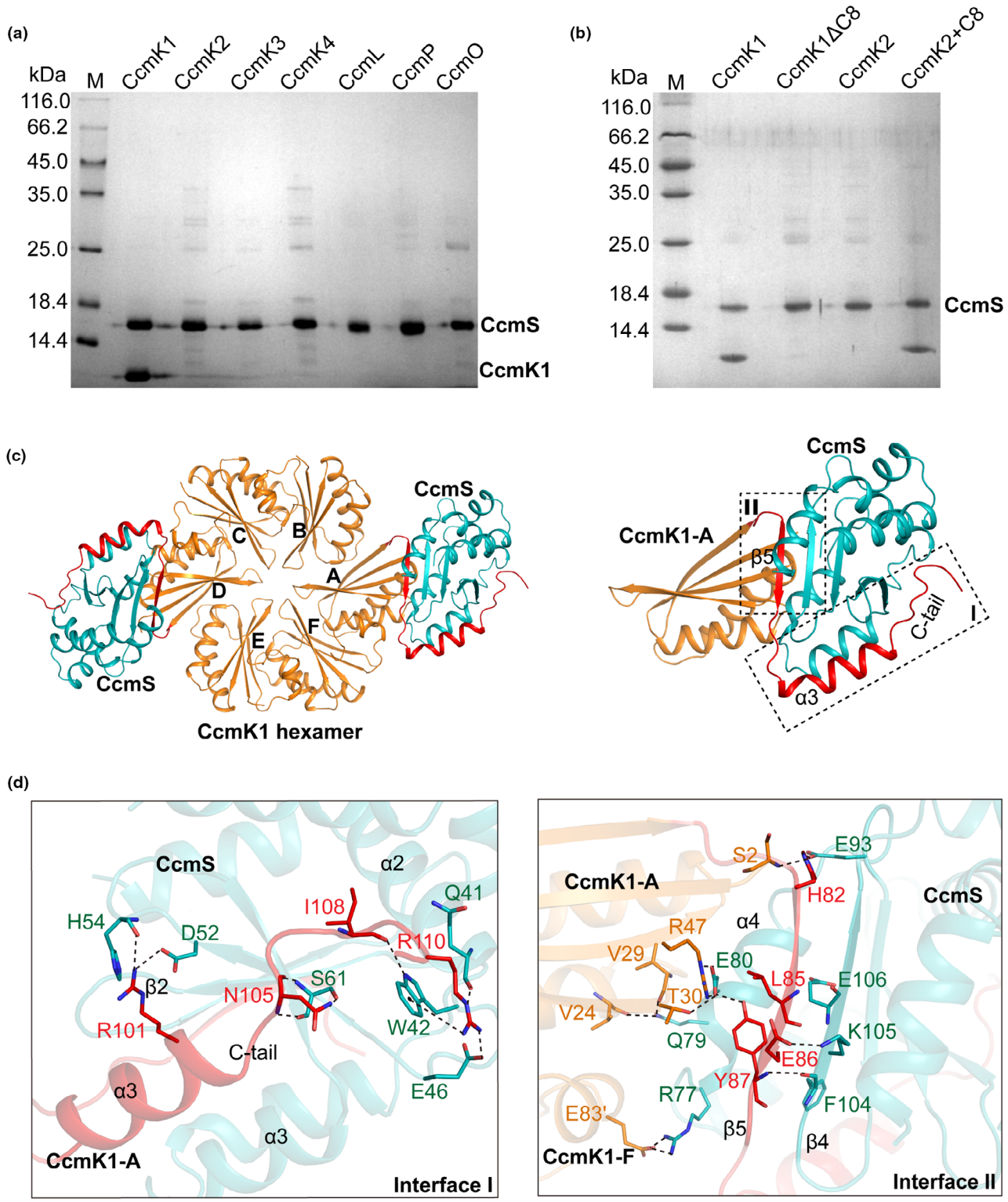

Fig. 2 CcmS specifically binds to the hinge domain of CcmK1. (a) Pull-down assays of CcmS against all types of shell proteins of *Syn*6803. BMC-H proteins: CcmK1-CcmK4; BMC-T proteins: CcmP and CcmO; BMC-P protein: CcmL. The CcmS protein was fused with a His_6 tag. The protein marker was loaded in lane 1. BMC, bacterial microcompartment. (b) Wild-type and mutant CcmK1 and CcmK2 were pulled down by His-tagged CcmS. CcmK1ΔC8: CcmK1 with eight C-terminal residues deleted. CcmK2+C8: CcmK2 fused with the eight C-terminal residues of CcmK1. (c) Cartoon representation of the CcmS-CcmK1 complex. The overall structure is shown on the left. The CcmS and CcmK1 subunits that bind to CcmS are highlighted on the right. The CcmS subunits and CcmK1 hexamer are colored teal and orange, respectively. Two hinge domains (β5, α3, and the C-tail) of CcmS-bound CcmK1 subunits are highlighted in red. Two distinct interfaces (labeled I and II) are outlined by black boxes. (d) The interfaces between CcmS and CcmK1. Interfaces I and II correspond to (c). The interacting residues are shown as sticks. The polar interactions are indicated by dashed lines.

complex (Fig. S3c), suggesting that *Syn*6803 adopts a CcmS-CcmK1 binding mode distinct from that in *Ana*7120.

CcmS also stabilizes the hinge domain of CcmK1 in the CcmK1-CcmK2 heterohexamer

Previous studies have shown that coexpression of four CcmK paralogs (CcmK1–4) in *E. coli* results in the formation of two predominant heterohexameric complexes, CcmK1-CcmK2 and CcmK3-CcmK4, which also suggests the plasticity of the carboxysome shell (Rae *et al.*, 2012; Garcia-Alles *et al.*, 2019; Sommer *et al.*, 2019). To analyze the expression profiles of CcmK1 and CcmK2, whose coding genes are located in the same *ccm* operon, we cloned and coexpressed the *ccmK2-ccmK1* operon in *E. coli*. Gel filtration chromatography and denatured polyacrylamide gel electrophoresis analysis revealed that a major fraction of the purified proteins formed the CcmK1-CcmK2 heterocomplex, in addition to the minor fraction of the CcmK1 and CcmK2 homohexamers (Figs 3a, S5a). Further pull-down assays revealed that the CcmK1-CcmK2 heterocomplex could also interact with CcmS (Fig. 3b). We therefore purified the ternary complex that has a calculated molecular weight of 109 kDa (Fig. S5b), comparable to that of the CcmS-CcmK1 complex. Then we solved the crystal structure of CcmS-CcmK1-CcmK2 at 2.5 Å resolution by molecular replacement. In the complex structure, four CcmK1 subunits and two CcmK2 subunits form a 4 : 2 heterohexamer, in which subunits C and F of the CcmK1 hexamer are replaced by two CcmK2 subunits, resulting in a twofold symmetric assembly pattern (Fig. 3c). Similar to the CcmS-CcmK1 complex, two CcmS molecules adopt the same pattern of binding to two CcmK1 subunits by fixing the hinge domain of CcmK1 (Figs 2c, 3c).

The structure of the CcmK1-CcmK2 heterohexamer closely resembles the previously reported structures of CcmK1 and CcmK2 hexamers (Kerfeld *et al.*, 2005; Tanaka *et al.*, 2008). Compared to the CcmK1 or CcmK2 hexamer, the structure of the CcmK1-CcmK2 heterohexamer adopts a flatter conformation, which deviates only *c.* 4° along the central axis (Fig. S5c). α4 at the C-terminus of CcmK2 pairs with the α3 of CcmK1 via hydrophobic interactions (Fig. 3c), which is also observed in the structures of the CcmK2 hexamers from *Synechococcus elongatus* PCC 7942 (PDB: 4OX7) (Cai *et al.*, 2015) and *Halothece* sp. PCC 7418 (PDB: 6OWG) (Sutter *et al.*, 2019) (Fig. S6). Although these various CcmK hexamers share a similar assembly pattern, they differ from each other in the polarities of the central pore. The area around the central pore of the CcmK1 hexamer is generally hydrophobic, which is mainly attributed to the residue Leu11, which points toward the center of the central pore (Fig. S7). In CcmK2, the counterpart residue is replaced by Arg11, which results in the formation of a highly positively charged central pore, potentially conferring the transport capability of negatively charged metabolites (Faulkner *et al.*, 2020). In the CcmK1-CcmK2 heterohexamer, the presence of Arg11 residues from two CcmK2 subunits also makes the central pore somewhat positively charged (Fig. S7). Notably, an additional density was observed at the central pore of the CcmK1-CcmK2 heterohexamer, which could be fitted with a molecule of tartrate ion that is most likely incorporated from the crystallization buffer (Fig. S7). The binding of tartrate ions to the central pore of the CcmK1-CcmK2 heterohexamer further supports the notion that the pores execute the transport of substrates and products into and out of the carboxysome (Faulkner *et al.*, 2020).

The hinge domain of CcmK1 is necessary for the formation of regular β-carboxysomes in *Syn*6803

To better understand the role of CcmK1 and CcmS in carboxysome assembly and functionality, we generated a series of *Syn*6803 genetic mutants with specific alterations in the CcmS and CcmK1/K2 proteins. Deletion of either *ccmK1* or *ccmK2* prohibited carboxysome biogenesis (Fig. S8), causing the cyanobacterial cells to only grow under high CO_2 (4%) conditions (HC) and therefore to exhibit a high-CO_2-requiring phenotype (Fig. 4a), similar to previous studies on CcmK mutants (Rae *et al.*, 2012; Cameron *et al.*, 2013). Moreover, the growth of the *ccmS* deletion mutant (*ΔccmS*) or C-tail truncation mutant of *ccmK1* (*ccmK1ΔC8*) was similar to that of the WT under HC conditions but was significantly slower under low CO_2 conditions (air, LC) (Fig. 4a). Compared to those in the WT, the rates of net photosynthetic CO_2 uptake in the *ΔccmS* and *ccmK1ΔC8* mutants were significantly reduced by *c.* 62.6% and 56.8%, respectively (Fig. 4b). Moreover, the rates of photosynthetic O_2 evolution in the *ΔccmS* and *ccmK1ΔC8* mutants decreased by 65.6% and 61.5%, respectively (Fig. 4c).

We further detected the ultrastructures of β-carboxysomes in the WT and *Syn*6803 mutant strains using transmission electron microscopy. Under LC conditions, the carboxysomes became significantly larger in *ΔccmS* than in the WT, and the overall architecture of the carboxysomes was somewhat abnormal (Fig. 4d), which also resulted in a decrease in photosynthetic capacity (Fig. 4b). The diameter of the carboxysomes in the *ΔccmS* strain was 359 ± 102 nm, which was *c.* twofold greater than that in the WT strain (215 ± 36 nm, Fig. 4e). Similarly, larger carboxysomes were also found in the *ΔccmS* strain under HC conditions (Fig. 4d,e). By contrast, unlike the *ΔccmS* mutant, the *ccmK1ΔC8* mutant could barely form carboxysomes under HC conditions (Fig. 4d). Even under LC conditions, only aberrant carboxysomes with irregular shapes and smaller sizes were formed in the *ccmK1ΔC8* mutant (Fig. 4d,e). Moreover, under either HC or LC conditions, the *ΔccmS* and *ccmK1ΔC8* mutants had fewer carboxysomes in a cell (approximately one per cell) than the *c.* 2 carboxysomes in a WT cell (Fig. 4f). These results clearly suggest that the C-tail of CcmK1 is necessary for the proper assembly of regular carboxysomes in *Syn*6803.

Discussion

The carboxysome shell gains more plasticity in the presence of the CcmK1-CcmK2 heterohexamer

The carboxysome shells contain variable BMC-H paralogs, each with distinctly conserved residues surrounding the pore, which

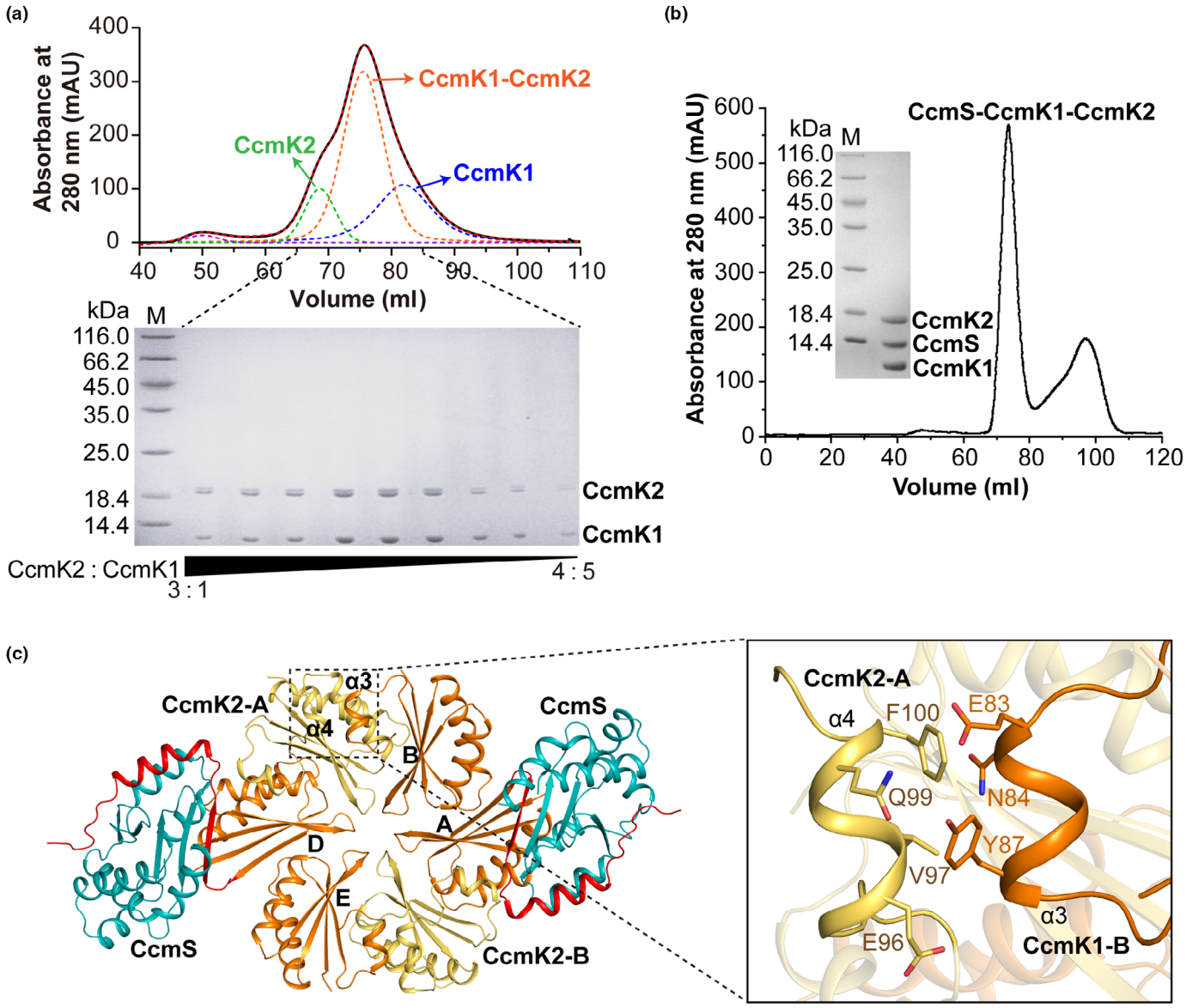

Fig. 3 Structure of the CcmS-CcmK1-CcmK2 complex and interactions between CcmK1 and CcmK2. (a) Gel filtration chromatography and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of the expression products of the *ccmK2-ccmK1* operon. CcmK1 and CcmK2 were fused with His_6-tags. The acquired chromatographic data were fitted to the log-normal peak equation, and the calculated peak parameters were reported by Peakfit (Stingl & Luider, 2021). The acquired peak data (black solid line) and the generated peak model (red dashed line) fit well together. The peak model is fitted by three log-normal peaks distinguished by different colored dashed lines. The peak positions correspond to the CcmK2 hexamer (green), CcmK1-CcmK2 heterohexamer (orange), and CcmK1 hexamer (blue). The purified protein was visualized by Coomassie blue-stained SDS-PAGE. The ratio of CcmK2 to CcmK1 proteins in different protein samples was calculated using grayscale integration and is shown below the SDS-PAGE gel. (b) Gel filtration chromatography and SDS-PAGE analysis of the CcmS-CcmK1-CcmK2 complex. (c) Cartoon representation of the CcmS-CcmK1-CcmK2 complex and the interfaces between CcmK1 and CcmK2. The CcmS, CcmK1, and CcmK2 subunits are colored teal, orange, and yellow, respectively. The two hinge domains are highlighted in red. The interface between the C-terminus of CcmK1 and CcmK2 is outlined by black boxes, and the detailed interactions are shown in the insets. The interacting residues are shown as sticks.

are assumed to be associated with specific metabolites (Kerfeld *et al.*, 2005; Rae *et al.*, 2013; Sommer *et al.*, 2017; Melnicki *et al.*, 2021). Previous structural studies and molecular dynamics simulations suggested that the central pores of CcmKs can serve as channels for metabolite entry and exit (Kerfeld *et al.*, 2005; Tanaka *et al.*, 2008; Faulkner *et al.*, 2020; Raza *et al.*, 2024). In addition, CcmK3 and CcmK4 were found to form heterohexamers at a 1 : 2 stoichiometry, which potentially alters the permeability properties of metabolite channels in carboxysome shells (Sommer *et al.*, 2019). In this study, we found that the deletion of either *ccmK1* or *ccmK2* could barely form carboxysomes in *Syn*6803 (Fig. S8), suggesting that both CcmK1 and CcmK2 are indispensable for the proper assembly of carboxysomes in some β-cyanobacterial strains. Therefore, despite

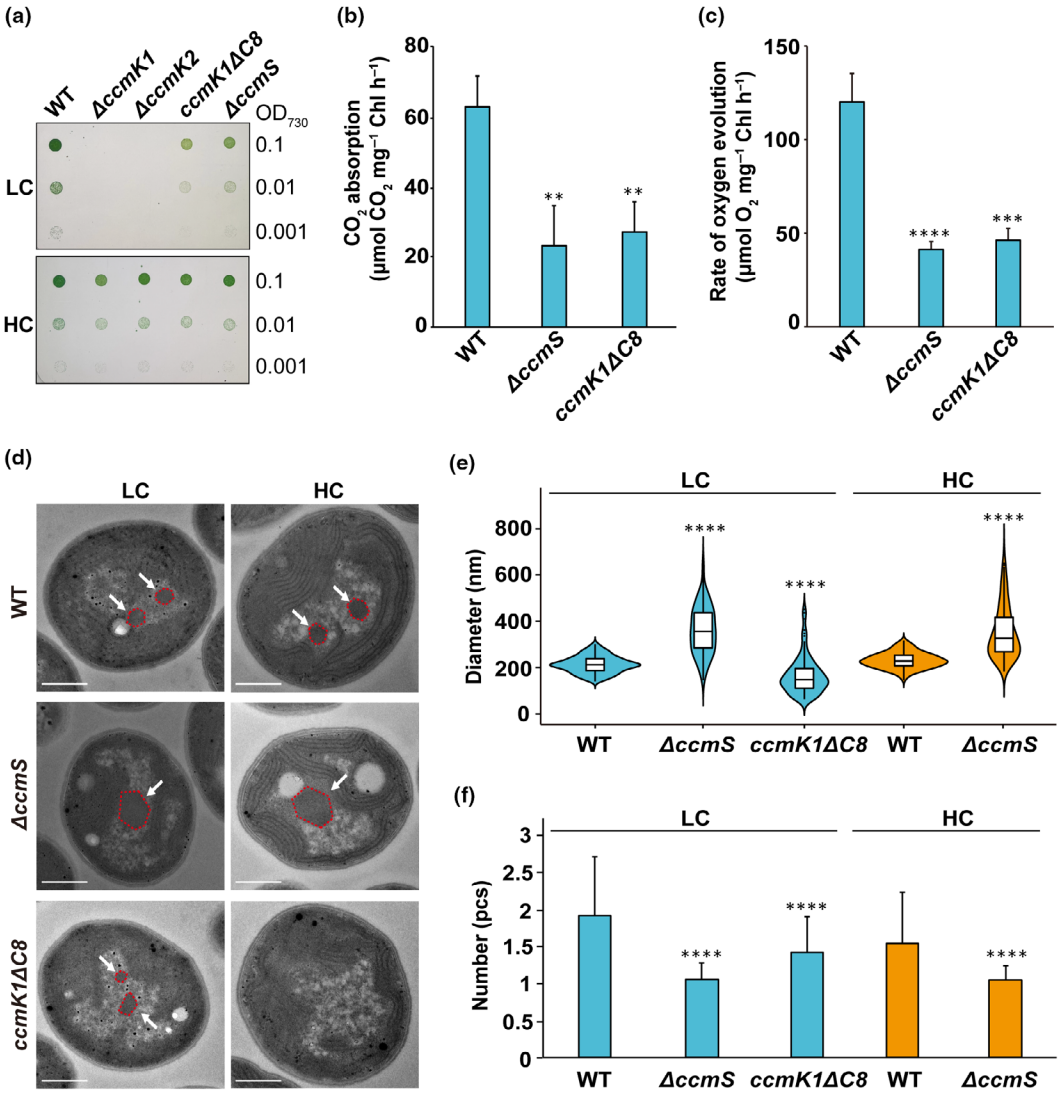

Fig. 4 Changes in the growth phenotype and carboxysomes of *Syn*6803 wild-type (WT) and mutants. (a) Growth phenotypes of the WT and mutants with different carbon sources. *ΔccmK1*: deletion of CcmK1; *ΔccmK2*: deletion of CcmK2; *ΔccmS*: deletion of CcmS; *ccmK1ΔC8*: deletion of the eight residues in the C-terminus of CcmK1. The cells were cultured to the logarithmic growth phase in BG-11 medium (pH 8.0) bubbled with 4% (v/v) CO_2 in air. The cells were adjusted to OD_{730} = 0.1, 0.01, or 0.001. Afterward, 2.5 μl of each cell suspension was placed on solid agar plates and incubated at 40 μmol photons m^{-2} s^{-1} under 4% CO_2 in air (HC) or ambient air (LC). (b) The net photosynthetic rate (i.e. the rate of CO_2 uptake) of the WT and mutants. The CO_2 concentration was controlled at 400 μmol mol^{-1}. Values represent the mean ± SE of three independent measurements. (c) Maximal rate of photosynthetic oxygen evolution. Photosynthetic oxygen evolution was measured using a Clark-type oxygen electrode. Values represent the mean ± SE of four independent measurements. (d) Transmission electron micrographs of ultrathin sections of the WT and mutated strains grown under different conditions. HC: cultured under 4% CO_2 in air; LC: cultured under ambient air. Carboxysomes are indicated by white arrows and outlined by red dashed lines. Bars, 500 nm (white line). The diameter (e) and number (f) of carboxysomes in the WT and mutant strains under HC (colored orange) and LC (colored teal) conditions. Transmission electron micrographs of carboxysome-visible cells with a maximum cross-sectional diameter not lower than 1.5 μm were used to measure the number of carboxysomes and the maximum cross-sectional diameter of the largest carboxysomes. The number of carboxysomes used was as follows: WT (LC): 324; *ΔccmS* (LC): 99; *ccmK1ΔC8* (LC): 95; WT (HC): 214; and *ΔccmS* (HC): 153. The number of cells used to count carboxysomes was as follows: WT (LC): 145; *ΔccmS* (LC): 94; *ccmK1ΔC8* (LC): 69; WT (HC): 185; and *ΔccmS* (HC): 146. In the violin plot of (e), whiskers represent the 10th and 90th percentiles, and boxes represent the 25th and 75th percentiles. The horizontal line in the box represents the median. The outliers are shown as individual points. In the histogram of (f), the error bars represent the mean ± SD. Asterisks indicate significant differences between the mutants and the WT (*t*-test: **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$).

sharing a sequence identity of over 90%, CcmK1 and CcmK2 are not functionally redundant in carboxysomes. However, biochemical and structural analyses revealed that CcmK1 and CcmK2 could form stable heterohexamers at a 2 : 1 stoichiometry (Fig. 3a,c). Structural analysis revealed that the CcmK1 and CcmK2 homo/heterohexamers have distinct structural features at the central pore. The CcmK1 hexamer features a highly hydrophobic central pore, which might not favor metabolite translocation. By contrast, the central pore of the CcmK2 hexamer is substantially positively charged, which is complementary to

negatively charged metabolites, including RuBP and 3-PGA. The physical properties of the central pore of the CcmK1-CcmK2 heterohexamer are notably different from those of the CcmK1 and CcmK2 homohexamers (Fig. S7). Accordingly, the CcmK1-CcmK2 heterohexamer increases the permeability toward various metabolites. Moreover, in addition to the major shell proteins CcmK1 and CcmK2, the β-cyanobacterial genome also contains varying numbers of additional CcmK paralogs (CcmK3-CcmK6) in satellite loci (Sommer *et al.*, 2017), which increases the plasticity of carboxysome shells, enabling β-cyanobacteria to inhabit a variety of dynamic habitats.

Phylogenetic analysis of 637 representative cyanobacterial species revealed that both CcmK1 and CcmK2 exist in most species that contain β-carboxysomes (Fig. S9). CcmK1 differs from CcmK2 in the presence of an extended C-tail (Fig. S4), and a lower abundancy of 50% in *Syn*6803 (Wang *et al.*, 2022). Our structural analysis revealed that the chaperone CcmS could specifically recognize this C-tail of CcmK1 (Fig. 2c,d), indicating that CcmS is a CcmK1-specific chaperone. Indeed, phylogenetic analysis revealed that most CcmK1-containing species also encode a CcmS, indicating a strong correlation in function and evolution between CcmS and CcmK1 (Fig. S9). As demonstrated by our *in vivo* genetic experiments, both CcmK1 and CcmK2, as well as the chaperone CcmS, are necessary for the assembly of regular carboxysomes (Fig. 4a,d). Notably, our previous studies indicated that the intracellular abundance of *Syn*6803 CcmS is *c.* 1 : 20 to that of CcmK1 (Chen *et al.*, 2023), indicating that CcmS is an efficient chaperone to protect the C-tail of CcmK1. The presence of both homo- and heterohexamers of CcmK2 and CcmK1 enhances the plasticity of the carboxysome shell enabling fine-tuned pore permeability for the efficient translocation of various metabolites.

Co-evolution of CcmK1 hinge domain and its chaperone CcmS

Our results together with previous findings demonstrated that the chaperone CcmS binds to the CcmK1 hexamer or the CcmK1-CcmK2 heterohexamer via specifically recognizing the C-tail of CcmK1 (Chen *et al.*, 2023; Cheng *et al.*, 2024). Sequence analysis showed that the C-tails of CcmK1 homologs are rich in arginine residues (such as Arg101, Arg110, and Arg111 in *Syn*6803 CcmK1; Arg101, Arg108, Arg112, and Arg113 in *Ana*7120 CcmK1), which participate in interacting with CcmS (Fig. S4). Phylogenetic analysis of the C-tails showed that the CcmK1 homologs could be clearly divided into two groups: one group including *Syn*6803 has a shorter C-tail and the other group represented by *Ana*7120 possesses a longer C-tail (Fig. S10). In the *Syn*6803 group, a CcmS monomer interacts with the shorter C-tail of CcmK1 hexamer; differently, in the *Ana*7120 group, the longer C-tail of CcmK1 is protected by two CcmS dimers. The residues in CcmS binding to CcmK1 are also exclusively conserved in their respective groups (Figs S1b, S10). It suggested that CcmK1 and CcmS in two distinct evolutionary groups are co-evolved, respectively (Fig. S10).

The CcmS-CcmK1 structure showed that CcmS interacted only with the hinge domain but not with the core of the CcmK1 hexamer. Given that each CcmK1 subunit could bind to one CcmS molecule, we propose that one CcmK1 hexamer could maximally bind to six CcmS molecules, forming a '6 + 6' complex. Under physiological conditions, CcmS and CcmK1 might form different complexes at varying stoichiometries. Our CcmS-CcmK1 structure at a 2 : 6 stoichiometry most likely represents a relatively stable intermediate of the carboxysome shell. The C-tail of CcmK1 is usually unfolded in the absence of CcmS binding, as seen in our complex structures CcmS-CcmK1 and CcmS-CcmK1-CcmK2 (Figs 2c, 3c), as well as in the CcmK1 structure (PDB: 3BN4) (Tanaka *et al.*, 2008). Upon CcmS binding, the helix α3 and C-tail of CcmK1 become folded into a protruded hinge domain (Figs 2c, 3c). Notably, sequence analysis showed that the α3 in the hinge domain is a typical amphipathic helix (Fig. S11a), which implied that it is prone to interact with each other. Molecular docking by HDOCK showed that two amphipathic α3 helices could form a coiled-coil structure in an antiparallel manner (Yan *et al.*, 2020) (Fig. S11b). Therefore, the α3 together with the C-tail of CcmK1 might function as a hinge to crosslink adjacent CcmK1 hexamers and CcmK1-CcmK2 heterohexamers by forming coiled-coil structures during carboxysome biogenesis (Fig. S11c). A previous AFM also observed the adjacent CcmK4 hexamers are bridged (Garcia-Alles *et al.*, 2017), most likely via the extended C-tails.

Superposition of the CcmS-CcmK1 structure onto the structure of a synthetic intact β-carboxysome shell (PDB: 6OWG) (Sutter *et al.*, 2019) revealed that CcmS sits upon two adjacent CcmK1 hexamers without any clashes (Fig. S11d). Notably, the α3 helices from two subunits of adjacent CcmK1 hexamers align with each other and form a coiled-coil structure, which is consistent with the molecular docking results (Fig. S11b). By contrast, superposition of the CcmS-CcmK1 structure onto the CcmK1 crystal structure, which adopts a flat sheet configuration in the crystal lattice (PDB: 3BN4) (Tanaka *et al.*, 2008), showed that CcmS partially overlaps with adjacent CcmK1 hexamers (Fig. S11e). Based on these observations, we propose that CcmS facilitates the formation of curved shell patches during β-carboxysome assembly, which might be mediated by the coiled-coil structure of adjacent CcmK1 hexamers at the outer surface of the shell. Therefore, CcmS may function as a regulator that helps to introduce curvature to the shell patches and thereby contributing to the proper shape and size of β-carboxysome.

A proposed model of CcmS-assisted assembly of β-carboxysome shell

The β-carboxysomes generally adopt a 'core-first' assembly pattern: the inner cargo interactions initiate the assembly process, and the final morphology is jointly regulated by both the shell and cargo components (Rotskoff & Geissler, 2018; Trettel *et al.*, 2024). Together with previous studies (Rotskoff & Geissler, 2018; Chen *et al.*, 2023; Trettel *et al.*, 2024), we propose a putative model for the assembly of β-carboxysomes assisted by the chaperone CcmS (Fig. 5). Initially, the condensation of the enzymes RuBisCOs and CAs by the scaffolding proteins CcmM and CcmN results in the formation of the inner core of

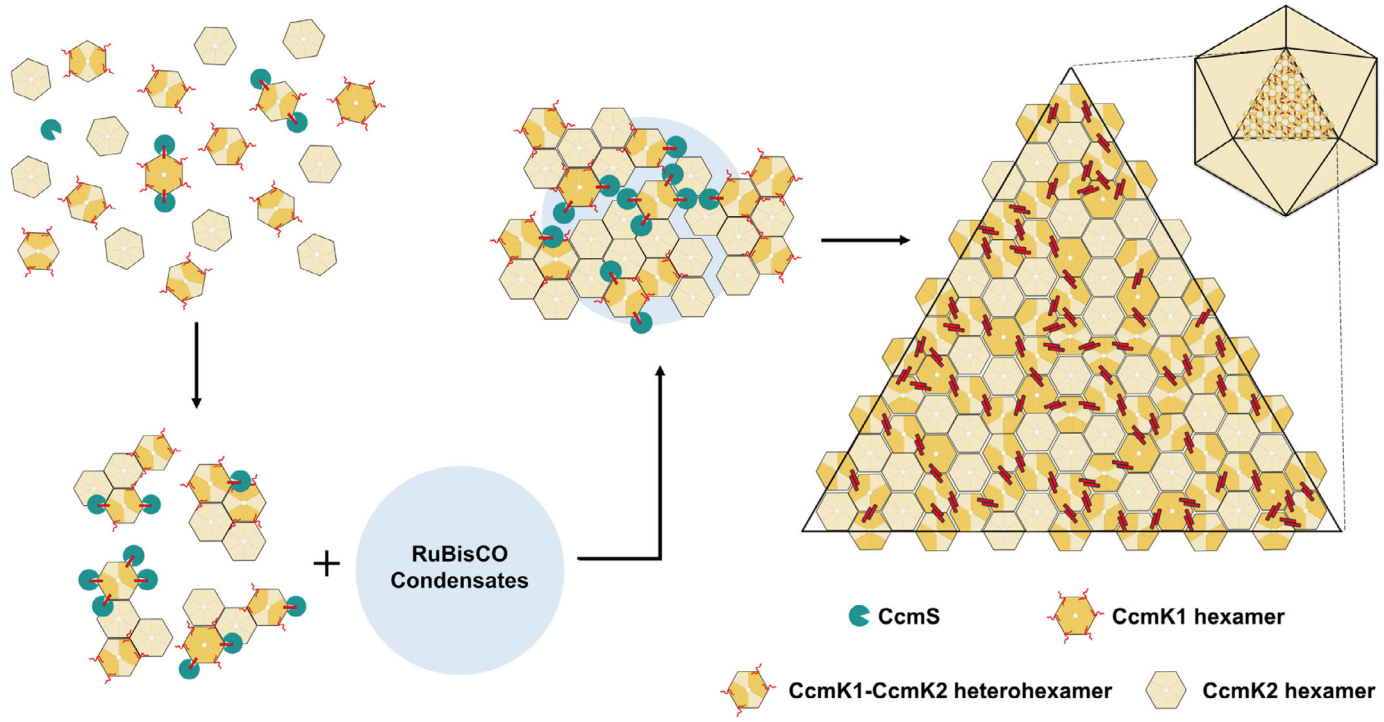

Fig. 5 A proposed model for the assembly of the β-carboxysome shell assisted by the chaperone CcmS in *Syn*6803. The CcmS and shell proteins are shown as the schemes. The hinge domains of CcmK1 are represented by red lines or sticks. Various shell hexamers, including the CcmK1 hexamers, the CcmK2 hexamers, and CcmK1-CcmK2 heterohexamers, are the major building blocks of the shell. In the absence of CcmS, the hinge domain of CcmK1 most likely is flexible, denoted by red lines. The binding of CcmS stabilizes the hinge domain of CcmK1 (denoted by sticks) that protrudes outward from the core structure of CcmK1, forming a coiled-coil structure that promotes the accumulation of hexamers to form shell patches. Eventually, various shell patches form the 20 shell facets, which are enclosed by the BMC-P proteins to form the intact shell.

carboxysomes (Wang *et al.*, 2019; Sun *et al.*, 2021; Zang *et al.*, 2021; Wang & Hayer-Hartl, 2023). Afterward, CcmK1, CcmK2, and other shell proteins self-assemble to form the shell patches that are subsequently recruited to the surface of the core, completing the encapsulation into an intact carboxysome (Cameron *et al.*, 2013; Aussignargues *et al.*, 2015; Trettel *et al.*, 2024). Assembly of the β-carboxysome shell is a complicated and fine-tuning process that requires the coordinated actions of various shell proteins and the chaperone CcmS. During β-carboxysome biogenesis, CcmS stabilizes the hinge domains of CcmK1, which protrude outward and may form coiled-coil structures between the adjacent CcmK homo/heterohexamers. Formation of these coiled-coil structures between CcmK1 hexamers at the outer surface of the shell might facilitate the formation of curved shell patches. Therefore, we propose that CcmS may function as a regulator, not only stabilizing the adjacent hexamers, but also introducing local curvature of the shell and thereby contributing to the proper shape and size for β-carboxysome. Indeed, deletion of CcmS in *Syn*6803 results in the formation of larger and structurally heterogeneous β-carboxysomes (Fig. 4d,e). The CcmS-mediated assembly of β-carboxysome differs from the assembly mechanism of α-carboxysome, where the scaffolding protein CsoS2 binds to the inner shell surface and is proposed to regulate the size of α-carboxysome (Li *et al.*, 2024; Zhou *et al.*, 2024). Upon the accumulation of different BMC-H and BMC-T proteins into patches surrounding the core, the intact carboxysome shell is eventually enclosed by the pentameric BMC-P at the vertices, accompanied by the turnover of CcmS (Fig. 5). In addition to forming homohexamers, the major shell proteins CcmK1 and CcmK2 can form heterohexamers, increasing the adaptability of shell permeability. Given the shared interface between the CcmK1 and CcmK2 subunits (Tanaka *et al.*, 2009; Garcia-Alles *et al.*, 2017, 2019; Sommer *et al.*, 2017), CcmK1-CcmK2 heterohexamers may exist in various stoichiometries to finely tune pore permeability for efficient substrate translocation under physiological conditions (Garcia-Alles *et al.*, 2019). Notably, the carboxysome assembly usually occurs under the oxidative microenvironment (Chen *et al.*, 2013), and interactions among components including CcmM-RuBisCO (Wang *et al.*, 2019) and CcmM-CcmN (Sun *et al.*, 2021) are usually subject to redox regulations. In addition, the HCO_3^- transporter SbtA is subject to redox regulation through its partner SbtB, which contains a redox-sensing structural motif capable of forming disulfide bonds in response to the circadian rhythm (Selim *et al.*, 2023). These findings highlight the redox regulation is commonly observed in cyanobacterial CCM, which finely tunes its functionality and helps cyanobacteria adapt to varying growth conditions and diurnal fluctuations. We therefore propose that CcmS may also be subject to redox regulation during β-carboxysome assembly and

maturation, although the fine regulatory mechanism remains to be elucidated.

Taken together, our findings elucidate the role of the chaperone CcmS in the regular assembly of β-carboxysomes, thus maintaining CCM function and normal cell growth. We propose a multistep fine regulatory process for β-carboxysome biogenesis, which provides a new avenue for the synthetic design of efficient carbon fixation machinery.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (http://www.nsfc.gov.cn; grant nos. 32171198 and 32241025), the Anhui Provincial Key Research and Development Project (http://kjt.ah.gov.cn; grant no. 2022l07020034), the Chaohu Lake Biological Resource Investigation and Research Project (grant no. 2020-340181-77-01-037328), and the USTC Research Funds of the Double First-Class Initiative (YD9100002023). Y-LJ thanks the Youth Innovation Promotion Association of the Chinese Academy of Sciences for their support (Membership No. 2020452).

Competing interests

None declared.

Author contributions

Y-LJ, C-ZZ, H-LM and YC conceived, designed, and supervised the project. Y-LJ, C-ZZ, H-LM, JL and J-XD analyzed the data and wrote the manuscript. JL and J-XD performed the molecular cloning, protein expression, purification, and crystallization. JL, J-XD, BL, Z-LZ and Y-LJ performed the X-ray data collection, structure determination, and model building. JL, B-RL and J-XD performed the biochemical assays. JL, BL, XC and J-XL performed the physiological experiments. All authors discussed the data and read the manuscript. JL, J-XD and XC contributed equally to this work.

ORCID

Xin Chen https://orcid.org/0009-0005-4956-7555
Yuxing Chen https://orcid.org/0000-0002-7560-1922
Jia-Xin Deng https://orcid.org/0009-0000-0621-9124
Yong-Liang Jiang https://orcid.org/0000-0001-5441-4936
Bo-Rui Li https://orcid.org/0009-0009-7142-9932
Bo Li https://orcid.org/0009-0008-7957-2744
Jing Li https://orcid.org/0009-0000-0078-9660
Hualing Mi https://orcid.org/0000-0003-1021-8372
Cong-Zhao Zhou https://orcid.org/0000-0002-6881-7151
Zhong-Liang Zhu https://orcid.org/0000-0002-3500-0727

Data availability

The accession numbers for the crystal structures reported in this paper are as follows: PDB: 9IUR for CcmS; 9IV3 for the CcmS-CcmK1 complex; and 9IV7 for the CcmS-CcmK1-CcmK2 complex.

References

Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J *et al.* 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**: 493–500.

Aussignargues C, Paasch BC, Gonzalez-Esquer R, Erbilgin O, Kerfeld CA. 2015. Bacterial microcompartment assembly: the key role of encapsulation peptides. *Communicative & Integrative Biology* **8**: e1039755.

Bar-On YM, Milo R. 2019. The global mass and average rate of rubisco. *Proceedings of the National Academy of Sciences, USA* **116**: 4738–4743.

Buchfink B, Reuter K, Drost H-G. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nature Methods* **18**: 366–368.

Cabello-Yeves PJ, Scanlan DJ, Callieri C, Picazo A, Schallenberg L, Huber P, Roda-Garcia JJ, Bartosiewicz M, Belykh OI, Tikhonova IV *et al.* 2022. α-cyanobacteria possessing form IA RuBisCO globally dominate aquatic habitats. *The ISME Journal* **16**: 2421–2432.

Cai F, Bernstein SL, Wilson SC, Kerfeld CA. 2016. Production and characterization of synthetic carboxysome shells with incorporated luminal proteins. *Plant Physiology* **170**: 1868–1877.

Cai F, Sutter M, Bernstein SL, Kinney JN, Kerfeld CA. 2015. Engineering bacterial microcompartment shells: chimeric shell proteins and chimeric carboxysome shells. *ACS Synthetic Biology* **4**: 444–453.

Cameron JC, Wilson SC, Bernstein SL, Kerfeld CA. 2013. Biogenesis of a bacterial organelle: the carboxysome assembly pathway. *Cell* **155**: 1131–1140.

Chandonia J-M, Guan L, Lin S, Yu C, Fox NK, Brenner SE. 2022. SCOPe: improvements to the structural classification of proteins – extended database to facilitate variant interpretation and machine learning. *Nucleic Acids Research* **50**: D553–D559.

Chen AH, Robinson-Mosher A, Savage DF, Silver PA, Polka JK. 2013. The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo. *PLoS ONE* **8**: e76127.

Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica. Section D, Biological Crystallography* **66**: 12–21.

Chen X, Zheng F, Wang P, Mi H. 2023. Novel protein CcmS is required for stabilization of the assembly of β-carboxysome in *Synechocystis* sp. strain PCC 6803. *New Phytologist* **239**: 1266–1280.

Cheng J, Li C-Y, Meng M, Li J-X, Liu S-J, Cao H-Y, Wang N, Zhang Y-Z, Liu L-N. 2024. Molecular interactions of the chaperone CcmS and carboxysome shell protein CcmK1 that mediate β-carboxysome assembly. *Plant Physiology* **196**: kiae438.

Crooks GE, Hon G, Chandonia J-M, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Research* **14**: 1188–1190.

Dai W, Chen M, Myers C, Ludtke SJ, Pettitt BM, King JA, Schmid MF, Chiu W. 2018. Visualizing individual RuBisCO and its assembly into carboxysomes in marine cyanobacteria by cryo-electron tomography. *Journal of Molecular Biology* **430**: 4156–4167.

Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallographica. Section D, Biological Crystallography* **60**: 2126–2132.

Faulkner M, Szabó I, Weetman SL, Sicard F, Huber RG, Bond PJ, Rosta E, Liu L-N. 2020. Molecular simulations unravel the molecular principles that mediate selective permeability of carboxysome shell protein. *Scientific Reports* **10**: 17501.

Feller U, Anders I, Mae T. 2008. Rubiscolytics: fate of Rubisco after its enzymatic function in a cell is terminated. *Journal of Experimental Botany* **59**: 1615–1624.

Fox NK, Brenner SE, Chandonia J-M. 2014. SCOPe: Structural classification of proteins–extended, integrating SCOP and ASTRAL data and classification of new structures. *Nucleic Acids Research* **42**: D304–D309.

Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**: 3150–3152.

Garcia-Alles LF, Fuentes-Cabrera M, Truan G, Reguera D. 2023. Inferring assembly-curving trends of bacterial micro-compartment shell hexamers from crystal structure arrangements. *PLoS Computational Biology* **19**: e1011038.

Garcia-Alles LF, Lesniewska E, Root K, Aubry N, Pocholle N, Mendoza CI, Bourillot E, Barylyuk K, Pompon D, Zenobi R *et al.* 2017. Spontaneous non-canonical assembly of CcmK hexameric components from β-carboxysome shells of cyanobacteria. *PLoS ONE* **12**: e0185109.

Garcia-Alles LF, Root K, Maveyraud L, Aubry N, Lesniewska E, Mourey L, Zenobi R, Truan G. 2019. Occurrence and stability of hetero-hexamer associations formed by β-carboxysome CcmK shell components. *PLoS ONE* **14**: e0223877.

Hagemann M, Bauwe H. 2016. Photorespiration and the potential to improve photosynthesis. *Current Opinion in Chemical Biology* **35**: 109–116.

Hayer-Hartl M, Hartl FU. 2020. Chaperone machineries of Rubisco – the most abundant enzyme. *Trends in Biochemical Sciences* **45**: 748–763.

Hennacy JH, Jonikas MC. 2020. Prospects for engineering biophysical CO_2 concentrating mechanisms into land plants to enhance yields. *Annual Review of Plant Biology* **71**: 461–485.

Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, Miyajima N, Hirosawa M, Sugiura M, Sasamoto S *et al.* 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions (supplement). *DNA Research* **3**: 185–209.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software v.7: improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.

Kerfeld CA, Aussignargues C, Zarzycki J, Cai F, Sutter M. 2018. Bacterial microcompartments. *Nature Reviews Microbiology* **16**: 277–290.

Kerfeld CA, Sawaya MR, Tanaka S, Nguyen CV, Phillips M, Beeby M, Yeates TO. 2005. Protein structures forming the shell of primitive bacterial organelles. *Science* **309**: 936–938.

Kinney JN, Axen SD, Kerfeld CA. 2011. Comparative analysis of carboxysome shell proteins. *Photosynthesis Research* **109**: 21–32.

Kupriyanova EV, Pronina NA, Los DA. 2023. Adapting from low to high: an update to CO_2-concentrating mechanisms of cyanobacteria and microalgae. *Plants* **12**: 1569.

Larsson AM, Hasse D, Valegård K, Andersson I. 2017. Crystal structures of β-carboxysome shell protein CcmP: ligand binding correlates with the closed or open central pore. *Journal of Experimental Botany* **68**: 3857–3867.

Lee MJ, Palmer DJ, Warren MJ. 2019. Biotechnological advances in bacterial microcompartment technology. *Trends in Biotechnology* **37**: 325–336.

Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v.5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research* **49**: W293–W296.

Li T, Chen T, Chang P, Ge X, Chriscoli V, Dykes GF, Wang Q, Liu L-N. 2024. Uncovering the roles of the scaffolding protein CsoS2 in mediating the assembly and shape of the α-carboxysome shell. *MBio* **15**: e0135824.

Mahalik JP, Brown KA, Cheng X, Fuentes-Cabrera M. 2016. Theoretical study of the initial stages of self-assembly of a carboxysome's facet. *ACS Nano* **10**: 5751–5758.

Matsumoto T, Yamano A, Sato T, Ferrara JD, White FJ, Meyer M. 2021. 'What is this?' A structure analysis tool for rapid and automated solution of small molecule structures. *Journal of Chemical Crystallography* **51**: 438–450.

McDonald IK, Thornton JM. 1994. Satisfying hydrogen bonding potential in proteins. *Journal of Molecular Biology* **238**: 777–793.

Melnicki MR, Sutter M, Kerfeld CA. 2021. Evolutionary relationships among shell proteins of carboxysomes and metabolosomes. *Current Opinion in Microbiology* **63**: 1–9.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**: 1530–1534.

Miziorko HM, Lorimer GH. 1983. Ribulose-1,5-bisphosphate carboxylase-oxygenase. *Annual Review of Biochemistry* **52**: 507–535.

Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica. Section D, Biological Crystallography* **67**: 355–367.

Nguyen ND, Pulsford SB, Hee WY, Rae BD, Rourke LM, Price GD, Long BM. 2023. Towards engineering a hybrid carboxysome. *Photosynthesis Research* **156**: 265–277.

Ni T, Jiang Q, Ng PC, Shen J, Dou H, Zhu Y, Radecke J, Dykes GF, Huang F, Liu L-N *et al.* 2023. Intrinsically disordered CsoS2 acts as a general molecular thread for α-carboxysome shell assembly. *Nature Communications* **14**: 5512.

Ochoa JM, Yeates TO. 2021. Recent structural insights into bacterial microcompartment shells. *Current Opinion in Microbiology* **62**: 51–60.

Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**: 307–326.

Rae BD, Long BM, Badger MR, Price GD. 2012. Structural determinants of the outer shell of β-carboxysomes in *Synechococcus elongatus* PCC 7942: Roles for CcmK2, K3-K4, CcmO, and CcmL. *PLoS ONE* **7**: e43871.

Rae BD, Long BM, Badger MR, Price GD. 2013. Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO_2 fixation in cyanobacteria and some proteobacteria. *Microbiology and Molecular Biology Reviews* **77**: 357–379.

Raza S, Sarkar D, Chan LJG, Mae J, Sutter M, Petzold CJ, Kerfeld CA, Ralston CY, Gupta S, Vermaas JV. 2024. Comparative pore structure and dynamics for bacterial microcompartment shell protein assemblies in sheets or shells. *ACS Omega* **9**: 35503–35514.

Rotskoff GM, Geissler PL. 2018. Robust nonequilibrium pathways to microcompartment assembly. *Proceedings of the National Academy of Sciences, USA* **115**: 6341–6346.

Samborska B, Kimber MS. 2012. A dodecameric CcmK2 structure suggests β-carboxysomal shell facets have a double-layered organization. *Structure* **20**: 1353–1362.

Selim KA, Haffner M, Mantovani O, Albrecht R, Zhu H, Hagemann M, Forchhammer K, Hartmann MD. 2023. Carbon signaling protein SbtB possesses atypical redox-regulated apyrase activity to facilitate regulation of bicarbonate transporter SbtA. *Proceedings of the National Academy of Sciences, USA* **120**: e2205882120.

Sommer M, Cai F, Melnicki M, Kerfeld CA. 2017. β-Carboxysome bioinformatics: identification and evolution of new bacterial microcompartment protein gene classes and core locus constraints. *Journal of Experimental Botany* **68**: 3841–3855.

Sommer M, Sutter M, Gupta S, Kirst H, Turmo A, Lechno-Yossef S, Burton RL, Saechao C, Sloan NB, Cheng X *et al.* 2019. Heterohexamers formed by CcmK3 and CcmK4 increase the complexity of beta carboxysome shells. *Plant Physiology* **179**: 156–167.

Stewart AM, Stewart KL, Yeates TO, Bobik TA. 2021. Advances in the world of bacterial microcompartments. *Trends in Biochemical Sciences* **46**: 406–416.

Stingl C, Luider TM. 2021. Applying log-normal peak fitting to parallel reaction monitoring data analysis. *Journal of Proteome Research* **20**: 4186–4192.

Sun H, Cui N, Han S-J, Chen Z-P, Xia L-Y, Chen Y, Jiang Y-L, Zhou C-Z. 2021. Complex structure reveals CcmM and CcmN form a heterotrimeric adaptor in β-carboxysome. *Protein Science* **30**: 1566–1576.

Sutter M, Kerfeld CA, Scott KM. 2022. A typical carboxysome loci: JEEPs or Junk? *Frontiers in Microbiology* **13**: 872708.

Sutter M, Laughlin TG, Sloan NB, Serwas D, Davies KM, Kerfeld CA. 2019. Structure of a synthetic β-carboxysome shell. *Plant Physiology* **181**: 1050–1058.

Tan YQ, Ali S, Xue B, Teo WZ, Ling LH, Go MK, Lv H, Robinson RC, Narita A, Yew WS. 2021. Structure of a minimal α-carboxysome-derived shell and its utility in enzyme stabilization. *Biomacromolecules* **22**: 4095–4109.

Tanaka S, Kerfeld CA, Sawaya MR, Cai F, Heinhorst S, Cannon GC, Yeates TO. 2008. Atomic-level models of the bacterial carboxysome shell. *Science* **319**: 1083–1086.

Tanaka S, Sawaya MR, Phillips M, Yeates TO. 2009. Insights from multiple structures of the shell proteins from the β-carboxysome. *Protein Science* **18**: 108–120.

Trettel DS, Pacheco SL, Laskie AK, Gonzalez-Esquer CR. 2024. Modeling bacterial microcompartment architectures for enhanced cyanobacterial carbon fixation. *Frontiers in Plant Science* **15**: 1346759.

Tsai Y, Sawaya MR, Cannon GC, Cai F, Williams EB, Heinhorst S, Kerfeld CA, Yeates TO. 2007. Structural analysis of CsoS1A and the protein shell of the *Halothiobacillus neapolitanus* carboxysome. *PLoS Biology* **5**: e144.

Walker BJ, VanLoocke A, Bernacchi CJ, Ort DR. 2016. The costs of photorespiration to food production now and in the future. *Annual Review of Plant Biology* **67**: 107–129.

Wallace AC, Laskowski RA, Thornton JM. 1995. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering* **8**: 127–134.

Wang H, Hayer-Hartl M. 2023. Phase separation of Rubisco by the folded SSUL domains of CcmM in beta-carboxysome biogenesis. *Methods in Molecular Biology* **2563**: 269–296.

Wang H, Yan X, Aigner H, Bracher A, Nguyen ND, Hee WY, Long BM, Price GD, Hartl FU, Hayer-Hartl M. 2019. Rubisco condensate formation by CcmM in β-carboxysome biogenesis. *Nature* **566**: 131–135.

Wang J, Huang X, Ge H, Wang Y, Chen W, Zheng L, Huang C, Yang H, Li L, Sui N *et al.* 2022. The quantitative proteome atlas of a model cyanobacterium. *Journal of Genetics and Genomics* **49**: 96–108.

Whitehead L, Long BM, Price GD, Badger MR. 2014. Comparing the *in vivo* function of α-carboxysomes and β-carboxysomes in two model cyanobacteria. *Plant Physiology* **165**: 398–411.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A *et al.* 2011. Overview of the CCP4 suite and current developments. *Acta Crystallographica. Section D, Biological Crystallography* **67**: 235–242.

Yan Y, Tao H, He J, Huang S-Y. 2020. The HDOCK server for integrated protein-protein docking. *Nature Protocols* **15**: 1829–1852.

Yeates TO, Crowley CS, Tanaka S. 2010. Bacterial microcompartment organelles: protein shell structure and evolution. *Annual Review of Biophysics* **39**: 185–205.

Yeates TO, Thompson MC, Bobik TA. 2011. The protein shells of bacterial microcompartment organelles. *Current Opinion in Structural Biology* **21**: 223–231.

Young G, Hundt N, Cole D, Fineberg A, Andrecka J, Tyler A, Olerinyova A, Ansari A, Marklund EG, Collier MP *et al.* 2018. Quantitative mass imaging of single biological macromolecules. *Science* **360**: 423–427.

Zang K, Wang H, Hartl FU, Hayer-Hartl M. 2021. Scaffolding protein CcmM directs multiprotein phase separation in β-carboxysome biogenesis. *Nature Structural & Molecular Biology* **28**: 909–922.

Zhao L, Cai Z, Li Y, Zhang Y. 2024. Engineering Rubisco to enhance CO_2 utilization. *Synthetic and Systems Biotechnology* **9**: 55–68.

Zhou R-Q, Jiang Y-L, Li H, Hou P, Kong W-W, Deng J-X, Chen Y, Zhou C-Z, Zeng Q. 2024. Structure and assembly of the α-carboxysome in the marine cyanobacterium *Prochlorococcus. Nature Plants* **10**: 661–672.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Biochemical and structural analyses of CcmS.

Fig. S2 Biochemical analysis of CcmS-CcmK1 complex.

Fig. S3 Structural analysis of CcmS-CcmK1 complex.

Fig. S4 Multiple sequence alignment of CcmK1 and CcmK2 homologs.

Fig. S5 Biochemical and structural analyses of CcmS-CcmK1-CcmK2 complex.

Fig. S6 Comparison of the C-terminal segments of hexameric shell proteins in β-carboxysomes.

Fig. S7 Polarities of the central pores in different hexameric shell proteins.

Fig. S8 Transmission electron micrographs of ultrathin sections of *ΔccmK1* and *ΔccmK2* strains grown under high CO_2 (4%) conditions.

Fig. S9 Unrooted maximum likelihood phylogeny of 637 RbcL sequences.

Fig. S10 Phylogenetic analysis of CcmS proteins and CcmK1 C-tails.

Fig. S11 Biochemical and functional analyses of the hinge domain of CcmK1.

Table S1 The cyanobacterial strains, plasmids, and protein sequences used in this study.

Table S2 Crystal parameters, data collection, and structure refinement.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

Disclaimer: The New Phytologist Foundation remains neutral with regard to jurisdictional claims in maps and in any institutional affiliations.