# Advances in the bacterial organelles for $CO_2$ fixation

Lu-Ning Liu[1,2,*]

[1] Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB, United Kingdom
[2] College of Marine Life Sciences, and Frontiers Science Center for Deep Ocean Multispheres and Earth System, Ocean University of China, 266003 Qingdao, China
*Corresponding author: Liu, L.-N. (luning.liu@liverpool.ac.uk)

**Carboxysomes are a family of bacterial microcompartments (BMCs) present in all cyanobacteria and some proteobacteria, which encapsulate the primary $CO_2$-fixing enzyme, Rubisco, within a virus-like polyhedral protein shell. Carboxysomes provide significantly elevated levels of $CO_2$ around Rubisco to maximize carboxylation and reduce wasteful photorespiration, thus functioning as the central $CO_2$-fixation organelle of bacterial $CO_2$-concentration mechanisms. Their intriguing architectural features allow carboxysomes to make a vast contribution to carbon assimilation on a global scale. In this review, we discuss recent research progress that provides new insights into the mechanisms of how carboxysomes are assembled and functionally maintained in bacteria and recent advances in synthetic biology to repurpose the metabolic module in diverse applications.**

**Keywords:** Bacterial microcompartment, Carboxysome, $CO_2$ fixation, $CO_2$-concentrating mechanisms, Organelle, Self-assembly, Synthetic biology

All life in the biosphere is based on carbon. It plays pivotal roles in regulating the atmospheric temperature, sustaining our food supply, and providing a major source of energy for our global economy. Much evidence has indicated that the imbalanced global carbon cycle caused by human activities has a drastic impact on the climate and ecosystems.

Autotrophs make significant contributions to the carbon cycle by fixing $CO_2$ to produce biomass. To date, six natural pathways for converting atmospheric $CO_2$ to organic matter have been identified in autotrophs (Box 1) [1], among which the Calvin-Benson-Bassham (CBB) cycle is the dominant $CO_2$-fixation pathway, accounting for ~95% of the carbon fixed on Earth [2]. The CBB cycle, employed by plants, algae, cyanobacteria, and proteobacteria, utilizes the $CO_2$-fixing enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). Despite the biological significance and abundance on a global scale [3], Rubisco is surprisingly an inefficient enzyme, and the catalytical reactions of Rubisco are essentially the limiting steps in photosynthetic $CO_2$ fixation. This is because of its slow catalytic rate and poor capability in discriminating between the competing substrates, $CO_2$ and $O_2$ (Box 1). The fixation of $O_2$ in photoautotrophs leads to a process termed photorespiration, which wastes energy and decreases sugar synthesis. To overcome the inherent limitations of Rubisco, many organisms, including cyanobacteria, algae, and $C_4$ plants, have evolved $CO_2$-concentrating mechanisms (CCM) for elevating the $CO_2$ level near Rubisco in cells to maximize carboxylation and competitively inhibit oxygenation [4].

## Bacterial $CO_2$-concentrating mechanisms

The cyanobacterial CCM system is typically composed of three bicarbonate transporters (BicA, SbtA, BCT1) in the plasma membrane, two $CO_2$-uptake complexes (NDH-$I_3$, NDH-$I_4$) in thylakoid membranes, as well as the $CO_2$-fixing organelles distributed in the cytoplasm, known as the carboxysomes (Fig. 1) [5]. The bicarbonate transporters pump external $HCO_3^-$ into the cell and $CO_2$-uptake complexes convert intracellular $CO_2$ to $HCO_3^-$ that is less membrane permeable than $CO_2$, collectively generating an accumulated $HCO_3^-$ pool in the cytosol and reducing carbon loss from the cell. Carboxysomes are a group of bacterial microcompartments (BMCs) and function as nanoscale "factories" that sequester Rubisco from the cytoplasm where oxygen is massively produced via the light reactions of photosynthesis [6]. The accumulated $HCO_3^-$ in the cytoplasm then enters the carboxysome and is converted back to $CO_2$ by carboxysomal carbonic anhydrase (CA) in close proximity to Rubisco. The resulting high levels of $CO_2$ within the carboxysome maximizes Rubisco carboxylation and reduces photorespiration, therefore increasing overall $CO_2$ fixation (Fig. 1) [7]. The produced three-carbon sugar 3-phosphoglyceric acid (3-PGA) diffuses outward into the cytoplasm and is utilized in the regeneration of ribulose-1,5-bisphosphate (RuBP) and synthesis of biomolecules. The full establishment of a functional CCM system with the physiological coordination of carboxysomes, bicarbonate

transporters, and $CO_2$-uptake complexes enables bacterial cells to adapt to diverse $CO_2$-limited aquatic habitats, such as freshwater, marine, and alkaline lakes [8].

**Box 1. Natural $CO_2$ fixation pathways and Rubisco**

To date, six $CO_2$ fixation pathways employed in nature have been discovered, including the Calvin-Benson-Bassham cycle (CBB cycle), the reductive tricarboxylic acid (TCA) cycle, the 3-hydroxypropionate cycle, the reductive acetyl-CoA pathway (Wood-Ljungdahl pathway), the 3-hydroxypropionate-4-hydroxybutyrate cycle, and the dicarboxylate/4-hydroxybutyrate cycle. The dominant $CO_2$-fixation pathway is the CBB cycle, which utilizes the $CO_2$-fixing enzyme, ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). Almost all carbon in the food we consume is derived from Rubisco's activity.

Rubisco can be classified into three evolutionary lineages, form I, II, and III [112]. The most abundant is form I Rubisco, a hetero-hexadecameric complex consisting of eight large and eight small subunits ($L_8S_8$), denoted as $CbbL_8S_8$ in α-carboxysomes or $RbcL_8S_8$ in β-carboxysomes. The functional Rubisco forms a cylinder-like shape, with a diameter of ~110 Å and a height of 100 Å (Fig. 2). The core of Rubisco consists of a tetramer of antiparallel dimers of the large subunit ($L_8$), capped by 4 small subunits at the top and 4 at the bottom [49]. Each RbcL dimer contains two catalytic sites. RbcS does not play a direct role in catalysis, but it may be necessary to stabilize the active sites of Rubisco and regulate Rubisco activities [113]. Recently, a new clade of form I Rubisco has been identified, which forms an octamer as a functional unit in the absence of RbcS [114].

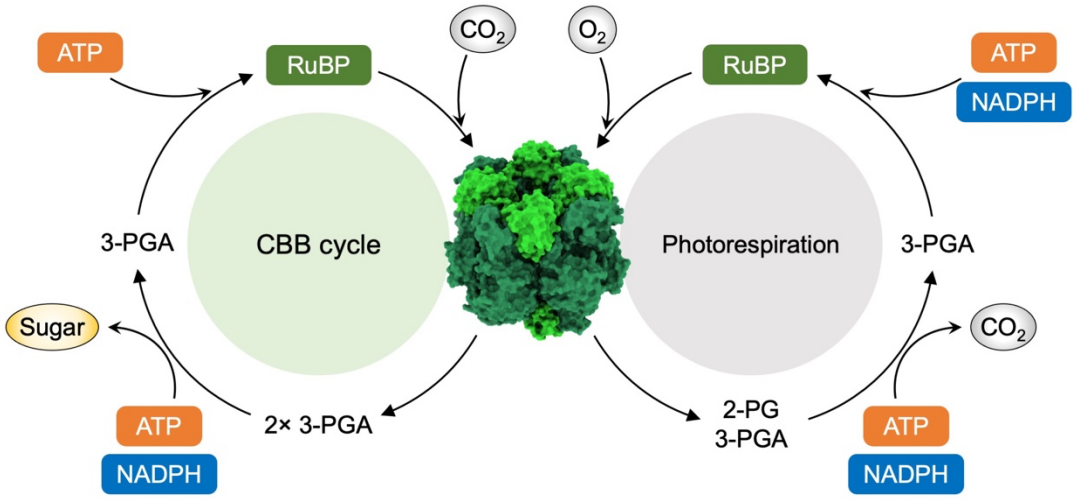


**Figure I. Catalytic reactions of Rubisco: carboxylation and oxygenation.** Regardless of its essential role in carbon fixation, Rubisco is a slow enzyme and its active site has a poor capacity of distinguishing its competing substrates: $CO_2$ and $O_2$, which are similar in size and electrostatic potential. Binding with $CO_2$ promotes the carboxylase activity of Rubisco to convert ribulose-1,5-bisphosphate (RuBP) and $CO_2$ into two molecules of the three-carbon sugar 3-phosphoglyceric acid (3-PGA) in the CBB cycle. Binding of $O_2$ to Rubisco drives the oxygenase activity of Rubisco, which produces 3-PGA and 2-phosphoglycolate (2-PG). Photorespiration converts 2-PG to 3-PGA, in which the rate of 3-PGA production is reduced and more energy is consumed compared with RuBP carboxylation. Thus, this process has been assumed to be a "wasteful" process.

Recent studies have advanced our understanding of the structures and functional mechanisms of the CCM components metabolically coupled with carboxysomes. Structural analysis of BicA revealed that this low-affinity $HCO_3^-$ transporter comprises an N-terminal transmembrane (TM) domain ($BicA^{TM}$), containing 14 TM helices that form the "7+7" fold inverted-topology repeats with a putative $HCO_3^-/Na^+$–binding pocket, and a C-terminal Sulphate Transporter and Anti-Sigma factor antagonist (STAS) domain ($BicA^{STAS}$) composed of 5 β-strands and 5 α-helices [9]. The functional BicA forms a dimeric structure, mediated by dimerization of $BicA^{STAS}$ domains [9]. The high-affinity $HCO_3^-$ transporter, SbtA, forms interactions with its regulatory protein SbtB [10]. Recent cryo-electron microscopy (cryo-EM) structures of the SbtA–SbtB complex showed that SbtA forms a functional trimer with $HCO_3^-/Na^+$ binding sites; binding of the SbtB T-loop to the core and scaffold domains of SbtA induces the allosteric inhibition of SbtA [11]. Both BicA and SbtA undergo $HCO_3^-$ transport via an "elevator mechanism", which involves the movement of the transport domain containing the substrate-binding sites to slide up and down the scaffold domain between an inward-facing state and an

outward-facing state [9,11]. The cyanobacterial $CO_2$-converting complex NDH-$I_3$ consists of a $CO_2$-uptake CupA/S module on the cytoplasmic side and functions as a primarily directional CA in $CO_2$ conversion [12]. A putative $CO_2$ channel spanning the TM subunit NdhF3 to the $Zn^{2+}$-binding site of CupA was proposed to allow $CO_2$ transport from the luminal side across the thylakoid membrane to the CupA $CO_2$-hydration site. Moreover, new inorganic carbon uptake systems, namely DabA1/DabA2/DabB1/DabB2, have been identified in diverse prokaryotes including carbon-fixing chemolithoautotrophs and achaea, but appear to be absent from cyanobacteria [13-15].

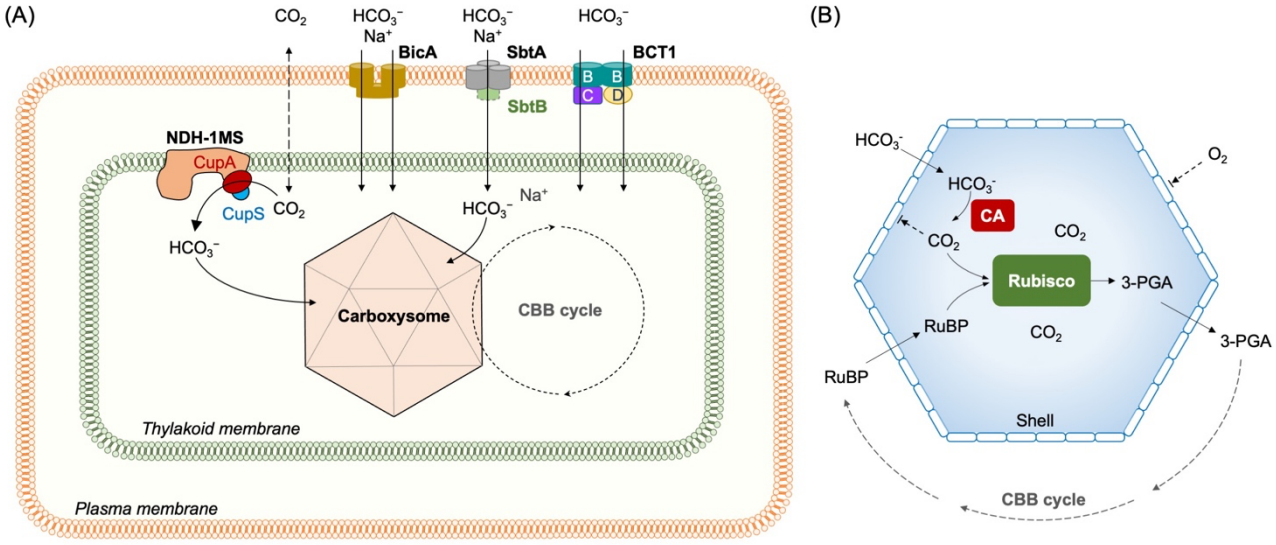


**Fig. 1. Schematic models of effective cyanobacterial $CO_2$-concentrating mechanisms and the carbon fixation pathway in the carboxysome.** (A) The cyanobacterial CCM consists of three bicarbonate transporters BicA, SbtA, and BCT1 in the plasma membrane, the $CO_2$-uptake complex NDH-1MS in thylakoid membranes, and carboxysomes in the cytoplasm. NDH-1MS contains an inducible form (NDH-$I_3$) and a constitutive form (NDH-$I_4$). The bicarbonate ion is accumulated within the cell by active bicarbonate transporters and NDH-1MS. Carboxysomes encapsulate the primary $CO_2$-fixing enzyme Rubisco and carbonic anhydrase (CA) within a semi-permeable icosahedral protein shell. The accumulated bicarbonate passes across the protein shell and is converted to $CO_2$ by CA to foster Rubisco $CO_2$ fixation, a key reaction in the Calvin-Benson-Bassham (CBB) cycle. (B) The carboxysome shell serves as a physical barrier for controlling the flux of specific metabolites in and out of the carboxysome. The shell permits the passage of cytosolic $HCO_3^-$ and ribulose-1,5-bisphosphate (RuBP) into the carboxysome. The shell may also function as a barrier to $O_2$ and $CO_2$, precluding $O_2$ influx and leakage of $CO_2$ from the carboxysomal lumen to the cytoplasm. CA in the carboxysome dehydrates $HCO_3^-$ to $CO_2$ and provides high levels of $CO_2$ near Rubisco, maximizing the carboxylation of RuBP, a process that adds $CO_2$ to generate 3-phosphoglycerate (3-PGA). 3-PGA is then transported across the shell and is metabolized via the CBB cycle.

## Carboxysomes – the $CO_2$-fixing bacterial microcompartments

Organelle formation and compartmentalization provide the structural basis for the physiological optimization and regulation of metabolic functions and efficiency in cells [16-18]. Over the past decade, numerous studies have characterized the specific organelles widespread in the bacterial kingdom, BMCs, which play important roles in $CO_2$ fixation, pathogenesis, and microbial ecology [6,19-22]. BMCs are membrane-free organelles composed purely of proteins and form a polyhedral structure with a diameter ranging from tens to hundreds of nanometers. The BMC consists of a virus-like shell, which is constructed by hundreds of shell proteins in the forms of hexamers (BMC-H), trimers (BMC-T), and pentamers (BMC-P) (Box 2, see recent reviews [23,24]), and an encapsulated enzymatic core that catalyzes a series of metabolic reactions (Fig. 2A, 2B). According to the distinct enzymes encapsulated and their functions, BMCs can be categorized into anabolic BMCs (carboxysomes) and catabolic BMCs (metabolosomes) (Box 2) [25,26].

Carboxysomes are anabolic BMCs responsible for $CO_2$ fixation in all identified cyanobacteria, many chemoautotrophs, and some purple photoautotrophs. Carboxysomes can be categorized into α-carboxysomes and β-carboxysomes, differing in the phylogenetic subclasses of Rubisco enclosed (Box 1) and their protein composition [27]. The α-carboxysomes possess form IA Rubisco and are found in some anoxygenic phototrophs and a wide range of α-cyanobacteria, including marine α-cyanobacteria that dominate oceanic ecosystems. The β-carboxysomes encapsulate plant-like form IB Rubisco and exist in a variety of β-cyanobacteria that live in diverse habitats with fluctuating environmental conditions, including freshwater β-

cyanobacteria, such as the model organisms *Synechococcus elongatus* PCC 7942 (*Synechococcus* 7942) and *Synechocystis* sp. PCC 6803 as well as some filamentous and bloom-forming genera.

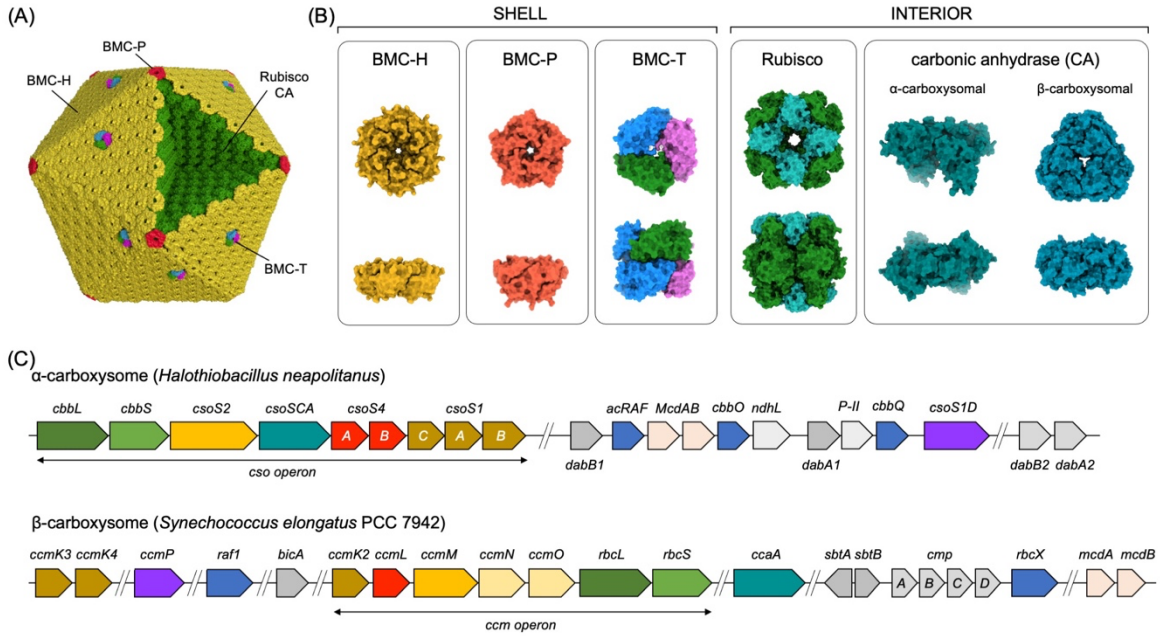


**Fig. 2. The structure, composition, and modularity of the carboxysome.** (A) A schematic model of the carboxysome structure, based on a confined icosahedral symmetry. (B) Protein components and structures of the carboxysomes. The carboxysome shell is composed of hexamers (BMC-H), trimers (BMC-T), and pentamers (BMC-P). The BMC-H proteins form hexagonal disks and tile the major area of the shell facets. The BMC-P proteins cap the vertices of fully assembled shells to confine the polyhedral geometry of the carboxysome. The BMC-T proteins form trimeric pseudo-hexamers, appearing as hexamers, which make up a smaller fraction of the shells. The interior enzymes within the carboxysome are Rubisco and CA. (C) The genomic organizations of gene loci expressing carboxysome proteins, ancillary factors, and CCM-related components, including cargo enzymes (green and teal), structural shell proteins (yellow, brown, red, purple), ancillary proteins (blue, peach), bicarbonate transporters and $CO_2$-uptake complexes (grey). The genetic loci of *Halothiobacillus neapolitanus* and *Synechococcus* 7942 are selected as representative loci of α- and β-carboxysomes, respectively.

**Box 2. Diverse BMC families**

BMCs can be categorized into anabolic and catabolic BMCs, based on their encapsulated enzymes and metabolic functions. The first-discovered BMCs are the carboxysomes, which are anabolic BMCs responsible for $CO_2$ fixation. Catabolic BMCs, also called metabolosomes, degrade diverse carbon substrates in heterotrophs. To date, experimentally characterized metabolosomes include propanediol utilization (PDU), ethanolamine utilization (EUT), glycyl radical enzyme-associated microcompartments (GRM), choline utilization (GRM2), fucose and rhamnose utilization (GRM5 and PVM), and 1-amino-2-propanol utilization (RMM) metabolosomes.

The BMC shell is built by hundreds of homologous shell proteins that belong to three main protein paralogs in the forms of hexamers (BMC-H), trimers (BMC-T), and pentamers (BMC-P). The BMC-H protein polypeptides assemble to form hexagonal disks and tile the majority of shell facets. The BMC-T proteins contain two tandem BMC-H domains and form trimeric pseudo-hexamers, appearing as hexamers, which make up a smaller fraction of the shells. The BMC-P proteins cap the vertices of the shell to maintain the polyhedral geometry of the whole BMC architecture [115,116]. In the last 15 years, large numbers of BMC proteins have been structurally and functionally characterized. An online MCPdb database has been recently developed to collect and view the structures of various BMC proteins [117]. Despite the architectural similarity, there is no bioinformatic or structural evidence in support of any evolutionary relationship between BMC shells and virus capsids [118].

The current models of BMC structures are largely based on a confined icosahedral symmetry, which is corroborated by the structures of recombinant mini-shells that are 20-40 nm in diameter [72,119]. However, increasing evidence has indicated the structural variability and plasticity of native BMCs [25,34,35,43], suggesting the adaptive nature of BMC structures and functions to cope with specific environmental niches.

Why the two subgroups of carboxysomes exist in different bacterial organisms that live in ecologically distinctive niches is still not fully understood (for details see reviews [27-29]). Increasing evidence has delineated the independent evolution of α-carboxysomes and β-carboxysomes [30]. Analysis of the BMC phylogeny, organizational composition, and assembly pathways have proposed that the α-carboxysomes of α-cyanobacteria and chemoautotrophs share a common ancestry and were likely originated from proteobacteria through horizontal gene transfer [31]. It has been hypothesized that the α-carboxysomes were evolved by invasion of another form of BMC possessing specific recruitment mechanisms, which triggered encapsulation of Rubisco and CA into the heterologous BMC shell [27]. This hypothesis was supported by the fact that some bacterial species contain both the Rubisco genes and BMC operons. In contrast, based on phylogenetic analysis of the sequences of BMC-P shell proteins, β-carboxysomes appear to have more closed evolutionary relationships to metabolosomes than to α-carboxysomes [29]. Despite the evolutionary divergence, the extraordinary genetic and structural features of carboxysomes lay the foundation for their functions in carbon assimilation and regulation [32], providing prokaryotes with a competitive growth advantage in natural habitats with varying temperatures, external pH, photo fluxes, inorganic carbon and water availability [6,27].

## Gene organization and protein composition

The genetic composition and organization of carboxysomes vary among bacterial species (Fig. 2C). Typically, the α-carboxysomes are composed of 8-11 polypeptides. The genes encoding α-carboxysome proteins are mainly clustered in a single *cso* operon, including *cbbL/S* encoding Rubisco large/small subunits, *csoSCA* that encodes CA, as well as shell and shell-associated proteins (*csoS2A/B*, *csoS4A/B*, *csoS1A/B/C/D/E*). The *cbbQ* and *cbbO* genes that encode Rubisco activases (Rca) are close to the *cso* operon. In contrast, the genes that encode β-carboxysome proteins are clustered in several genomic locations. For example, the *ccm* operon of *Synechococcus* 7942 involves *rbcL/S* encoding Rubisco large/small subunits and *ccmK2/L/M/N/O* encoding shell proteins. It is worth noting that association of *rbcLS and ccmO* with the *ccmKLMN* locus occurs in only 60% of cyanobacteria [33]. Other carboxysome-related genes located outside the *ccm* operon include *ccaA* (encoding CA), *ccmK3/K4* (encoding minor shell proteins), as well as *rbcX* and *raf1* that encode Rubisco chaperones. These carboxysome-related operons are nested within conserved gene clusters, termed superloci, which typically encode CCM-related components, such as bicarbonate transporters and other proteins ancillary to carboxysome assembly and function (Fig. 2C). The dispersal of carboxysome genes in the genome presumably provides a means for independent regulation and increasing the plasticity of their expression in response to environmental changes.

The stoichiometric composition is key for the proper assembly and structures of BMCs in both native and recombinant forms [34]. Recently, the protein stoichiometry of native β-carboxysomes from *Synechococcus* 7942 has been examined using single-molecule fluorescence microscopy [35]. Rubisco is a highly abundant component of the β-carboxysome, with ~850 copies per β-carboxysome in cells grown under moderate light, forming a densely packed Rubisco core within the polyhedral organelle [36]. CcmM, a linker protein that binds Rubisco to shell proteins via interacting with the recruitment protein CcmN [37,38], has ~700 copies per β-carboxysome. The shell hexamers CcmK2/K3/K4 have ~1100 copies in total [39,40]. CcmK3, CcmK4, CcmL, CcmN, CcaA, and RbcX are minor components in the β-carboxysome [35]. Given the notion that BMCs possess a typical icosahedral shape, there should be 12 BMC-P proteins confining the vertices of a fully assembled shell. However, it has been revealed that the copy numbers of CcmL per β-carboxysome vary depending on environmental conditions [35]. The vertices without pentamer capping have also been seen in engineered BMC shells [41,42]. This dynamic vertex organization of β-carboxysomes provides a possible structural mechanism for regulating shell permeability. The gene encoding the BMC-T protein CcmP is distant from the *ccmKLMN* locus. The abundance of CcmP has not been determined by single-molecule fluorescence imaging and mass spectrometry [35,36]. How CcmP proteins are organized in the β-carboxysome structure and whether it has similar dynamic behavior as its BMC-T structural ortholog in other BMCs [43] to serve as the potential modulators of shell permeability are unclear.

Current knowledge of the protein stoichiometric composition of α-carboxysomes, prevalent in oceanic cyanobacteria and autotrophic proteobacteria, is based on protein electrophoresis profiles, the size and geometry of α-carboxysomes, and low-resolution cryo-electron microscopy [44-47]. The α-carboxysomes in the chemoautotroph *Halothiobacillus* (*H.*) *neapolitanus* encapsulate up to 270 copies of Rubisco composed of CbbL and CbbS. Other major components of α-carboxysomes involve the CsoS1 shell protein paralogs and the linker protein CsoS2. The *H. neapolitanus* α-carboxysome contains 585 CsoS1 paralogs and 329 CsoS2 proteins; the α-carboxysome from the marine α-cyanobacterium *Prochlorococcus marinus* MED4 comprises 538 CsoS1 paralogs and 163 CsoS2 proteins [45,47]. The BMC-T protein CsoS1D, the structural ortholog of CcmP in β-carboxysomes, appears as a minor shell component (12 copies) in the α-carboxysome, with similar

content as the BMC-P proteins CsoS4. The precise stoichiometry of individual α-carboxysome components remains to be further evaluated, for example by exploiting quantitative mass spectrometry.

## Rubisco – the key $CO_2$-fixation enzyme within the carboxysome

***Rubisco assembly.*** The assembly of the major cargo Rubisco (Box 1) is fundamental for the construction of the overall carboxysome architecture [48]. Rubisco assembly represents a paradigm of chaperone-assisted processes. Folding of newly-synthesized RbcL requires chaperonin proteins GroEL and GroES [49], and the following assembly of folded RbcL monomers into oligomers is triggered by binding with the chaperones, Rubisco assembly factor 1 (Raf1) and RbcX. In *Synechococcus* 7942, Raf1 mediates the formation of the RbcL dimer and RbcL dimer-dimer interactions, thus critical for the assembly of Rubisco holoenzymes [48]. The Δ*raf1* mutant showed a reduction in Rubisco content and impaired formation of canonical carboxysomes. Consequently, a few small carboxysome-like structures with disordered Rubisco packing and many Rubisco-containing assemblies in the absence of shell proteins were observed, indicating the involvement of Raf1 in Rubisco assembly and carboxysome biogenesis [48]. In contrast, carboxysomes in the RbcX-deficient *Synechococcus* 7942 strain exhibited an increased size and reduced copy numbers per cell [50]. Collectively, these results suggested that Raf1 may function in an antagonistic way with RbcX in Rubisco assembly in *Synechococcus* 7942.

***Rubisco activation.*** While Rubisco catalyzes the carboxylation or oxygenation of RuBP, early binding of RuBP at the active sites could result in inactive Rubisco. The Rubisco activase (Rca) has evolved to remove inhibitors from Rubisco active sites to restore its activity, a process that is driven by ATP hydrolysis. The genes encoding the Rca, such as CbbX [51] and CbbQO, are present in α-carboxysome superloci. CbbQ and CbbO form a hetero-oligomeric complex that docks on the C-terminus of the Rubisco large subunit CbbL [52]. In the formed complex, CbbQ functions as a motor and exists as a hexameric ring with the AAA+ ATPase domains; CbbO acts as an adapter to bridge CbbQ and inactive Rubisco through interactions with its von Willebrand factor A domain [52-54]. It has been shown that both CbbQ and CbbO are incorporated within the native and engineered α-carboxysomes as the structural components [53,55]. Their function in activation of Rubisco and the strong association of CbbQ with the α-carboxysome shell [53] lead to the assumption that CbbQ and CbbO associate with the inner surface of the α-carboxysome shell and the cargo Rubisco. How CbbQO conduct their roles and dynamically interact with Rubisco within the α-carboxysome remains unclear. For the form IB Rubisco from many β-cyanobacteria, the Rca-like protein was revealed to have dual functions [56]. It is positioned on the Rubisco catalytic site for metabolic repair and pulls the N-terminal tail of RbcL into the central pore of the hexameric Rca-like protein. The specific interactions with Rubisco may lead to recruitment of the Rca-like protein to the β-carboxysome structure. It was further suggested that the β-cyanobacterial Rca-like protein, with a C-terminal small-subunit-like (SSUL) domain, might be incorporated in β-carboxysomes and mediate Rubisco condensation, following a similar way as CcmM does. However, it does not function as a canonical Rca to activate Rubisco [57].

***Rubisco nucleation.*** Rubisco enzymes are densely packed to form the core of the carboxysome, mediated by specific linker proteins that bear structurally disordered domains. It is now apparent that Rubisco nucleation is essential in mediating carboxysome assembly and structure. For the β-carboxysomes, nucleation of Rubisco is driven by CcmM that binds with another linker protein, CcmN [38,58-61]. CcmM comprises an N-terminal CA domain and 3–5 Rubisco SSUL domains, connected by flexible linkers. The SSUL modules link the Rubisco holoenzyme through multivalent interactions with the interface between two RbcL dimers, without replacing RbcS [60,61]. The SSUL-mediated Rubisco packaging results in formation of a Rubisco condensate via liquid-liquid phase separation [61]. For the α-carboxysomes, the intrinsically disordered protein CsoS2 functions as the linker protein to bind Rubisco and shell proteins [62,63]. The N-terminal motif of CsoS2 forms multivalent interactions with Rubisco, over the similar regions where CcmM binds, inducing phase separation eventually into a liquid-like Rubisco condensate [64].

## Carbonic anhydrase for bicarbonate conversion

The bicarbonate transporters and $CO_2$-uptake complexes ensure accumulation of high levels of $HCO_3^-$ within the cyanobacterial cytoplasm, which drives $HCO_3^-$ diffusion into the carboxysome. Hence, minimal CA activity in the cytoplasm and encapsulation of CA near Rubisco within the carboxysome are required. Expressing exogenous CA in the cytoplasm impedes CCM function and results in cyanobacterial cell growth only under high $CO_2$ [65]. Within the carboxysome, CA catalyzes the conversion of $HCO_3^-$ to $CO_2$, resulting in accumulation of $CO_2$ around Rubisco to enhance carboxylation (Fig. 3). The α-carboxysomes contain a

dimeric β-CA, namely CsoSCA (Fig. 2) [66]. However, CsoSCA differs from canonical β-CA, because it lacks significant sequence similarity to β-CA and a pair of active sites that a typical β-CA has. Each CsoSCA consists of an N-terminal domain presumably responsible for protein-protein interactions, a catalytic domain, and a C-terminal domain. CsoSCA likely associates with the inner surface of the shell, and it can be incorporated into empty α-carboxysome shells but is not required for shell formation [62]. In β-carboxysomes, the role of β-CA is fulfilled by two proteins, CcaA and CcmM. CcaA possesses a hexameric (trimer-of-dimers) organization and is likely encased within the β-carboxysomes through interactions with CcmM [67]. The cyanobacterial cells without CcaA required high $CO_2$ levels to grow, demonstrating the role of β-CA in $CO_2$ fixation [68]. Interestingly, the *ccaA* gene is only present in two-thirds of reported β-cyanobacterial genomes. In a subset of species that lack *ccaA*, the CA function could be replaced by CcmM, whose N-terminal domain serves as an active γ-CA [69]. Therefore, CcmM has dual functions in mediating β-carboxysome formation and performing the CA activity.

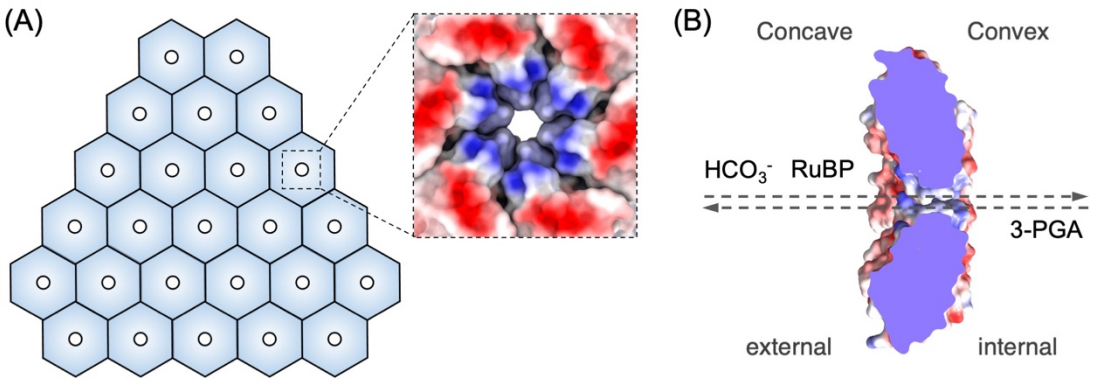


**Fig. 3. Selective permeability of carboxysome shell proteins.** (A) The central pores of shell proteins are proposed to mediate the permeability of metabolites to pass inward and outward across the shell. The pores possess positive electrostatic surface potentials (blue), enabling transport of negatively charged $HCO_3^-$ rather than uncharged molecules $CO_2$ and $O_2$. The negatively charged surface area of shell proteins is shown in red. (B) A cross-section representation illustrates the transport of Rubisco substrates and products through the pore of a carboxysome shell protein.

## Shell organization and permeability

The substrates, products, and enzymatic cofactors involved in the encapsulated catalytic pathways should pass inward and outward across the BMC shell. Both α-carboxysomes and β-carboxysomes possess a single-layer protein shell, as evident by electron microscopic observations [36,70,71]. The single-layer shell organization is further supported by recent studies on recombinant BMC shells and shell protein assemblies [43,72-74]. The carboxysome shell was speculated to function as a semi-permeable barrier to mediate transport of Rubisco's substrates and products, reduce $CO_2$ leakage into the cytosol and exclude the competing substrate $O_2$ (Fig. 1) [75]. This feature allows the generation of a specific microenvironment within the carboxysome to favor Rubisco carboxylation and alleviate oxygenation. Impairing shell formation could render cells incapable of $CO_2$ fixation in ambient air and cell growth requires high $CO_2$ levels. The $CO_2$-enriched environment within the carboxysome has been utilized in the repurposed nanoreactor based on the α-carboxysome shell to promote the catalytic activities of oxygen-sensitive hydrogenases for hydrogen production [62].

How the shell structure governs its selective permeability remains obscure. All shell proteins bear a central pore with distinctive electrostatic properties on the concave and convex sides. Structural analysis of the recombinant shell assemblies revealed that shell proteins display the same orientation in shell facets, with their concave side facing outward [72,73]. Since shell proteins are closely packed in the shell facets, the central pores of shell proteins are presumed to provide portals for metabolite transport in and out of BMCs (Fig. 3) [76-79]. Recent computational simulations have suggested that the positively charged pores of shell hexamers (CcmK2 and CcmK4 in β-carboxysomes, CsoS1A in α-carboxysomes) allow diffusion of negatively charged $HCO_3^-$ rather than uncharged $CO_2$ and $O_2$ [76,80]. Mutation on the key pore residues could affect carboxysomal $CO_2$ fixation and cell growth [76]. However, the pores of these major shell proteins are relatively small for the passage of large metabolites RuBP and 3-PGA. A conformational change of the pore occurs when binding with 3-PGA [76]. An alternative channel for large negatively charged metabolites to pass across the shell could be the pores of the BMC-T proteins, which are larger than those of the BMC-H and BMC-P proteins [81,82]. The stacked conformations of the β-carboxysomal CcmP and the α-carboxysomal CsoS1D were proposed to enable molecular transport following an "airlock" mechanism [24,81,83].

Despite a polyhedral organization composed of shell facets and vertices, the carboxysome shell possesses a marked degree of structural flexibility, which may provide a means for modulating shell permeability [25,78]. It was reported that CcmK3 and CcmK4 could form hetero-hexamers in a pH-dependent manner, likely leading to the opening and closure of the β-carboxysome shell [84,85]. Likewise, dynamic protein assembly and "capping" of shell proteins, as observed in the recombinant shell facets and mini-shells, represent other possible mechanisms for tuning shell permeability [42,73,74,84]. The carboxysome shell has also been suggested to be permeable to protons and reducing agents, to allow similar pH conditions in the carboxysome and the cytoplasm [86] and generation of an oxidizing internal environment [87]. Protons and the redox environment within the carboxysomes appear essential for Rubisco condensation, dehydration of $HCO_3^-$, CA activity and carboxysome function [69,88].

## Carboxysome biogenesis and positioning

Carboxysome biogenesis, structure, and function are actively regulated by the metabolism of their native hosts. The biogenesis pathways of α- and β-carboxysomes likely differ (Fig. 4). Recent studies documented that *de novo* assembly of β-carboxysomes exploits the "inside-out" model [87,89]. Rubisco enzymes first condense, mediated by CcmM, to form a liquid-like Rubisco core [60,61]. The recruitment protein CcmN interacts with CcmM by its N-terminus and shell proteins through its C-terminal peptide [38], inducing encapsulation of shell proteins surrounding the Rubisco matrix and eventually formation of an intact β-carboxysome. For the α-carboxysomes, it has been shown that Rubisco condensation is driven by CsoS2, which forms multivalent interactions with Rubisco on its N-terminus and binds shell proteins on its C-terminus [62-64]. Electron microscopy has visualized "partial" carboxysome structures, with varying amounts of Rubisco proteins surrounded incompletely by shell proteins [70,90] and native-like empty shells reconstructed in *Escherichia coli* (*E. coli*) [62], implying that Rubisco core assembly and shell encapsulation in α-carboxysome biogenesis may exploit a simultaneous and/or independent manner [6].

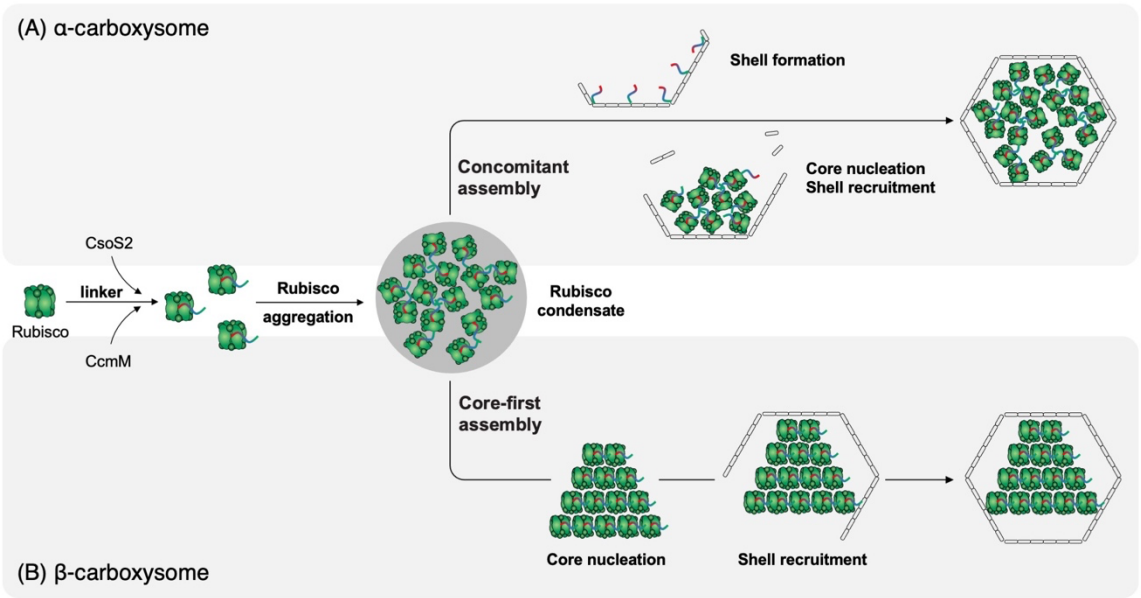


**Fig. 4. Assembly pathways of carboxysomes.** Rubisco binds specific linker proteins that are intrinsically disordered to facilitate carboxysome cargo assembly and formation of liquid-like Rubisco condensates, through a process termed liquid-liquid phase separation. The assembly of α- and β-carboxysomes is proposed to utilize different pathways. (A) α-carboxysomes undertake "concomitant" assembly, in which the shell assembles concurrently with aggregation of Rubisco. (B) β-carboxysomes employ "core-first" (or inside-out) assembly, in which Rubisco nucleation results in the formation of a densely packed enzyme core, followed by shell recruitment and encapsulation, to generate a complete carboxysome structure.

Spatial positioning of carboxysomes in cells is crucial for cell metabolism and growth in specific environments. β-carboxysomes exhibit a linear arrangement and even distribution in cyanobacterial cells, which is key for retaining carboxysome function and inheritance during cell division [91]. The specific subcellular localization of carboxysomes is determined by McdA and McdB proteins that are functionally similar to bacterial DNA partitioning systems [92-94]. When lacking McdAB, the carboxysome structures are prone to aggregate at the cell poles, where carboxysome biogenesis and degradation take place [89,95]. The McdAB systems exist among β-cyanobacteria that possess β-carboxysomes [96]. Recently, a McdAB-like

system has been identified in chemoautotrophic cells to position α-carboxysomes [97]. Whether similar α-carboxysome-positioning mechanisms exist in α-cyanobacteria remains unknown. In addition, varying light conditions during cell growth, including light intensity and light-dark cycles, has been illustrated to modulate the biosynthesis, structure, and distribution of carboxysomes [35,98-100].

## Repurposing carboxysomes for improving carbon fixation and new functions

The intrinsic self-assembly and modular characteristics of carboxysomes and their significance in enhancing $CO_2$ fixation make them a promising engineering objective. To date, great efforts have been made to engineer α- and β-carboxysome structures in various heterologous organisms. The α-carboxysome-like structures have been reconstituted in *E. coli* and a Gram-positive bacterium by expressing the *cso* operon from *H. neapolitanus* [101-103]. Moreover, expressing the multiple operons encoding β-carboxysome proteins has resulted in formation of β-carboxysome-like structures in *E. coli*, with detectable Rubisco activity [104].

Other strategies have been exploited in parallel, including producing functional carboxysome structures with reduced compositions, to unravel BMC assembly principles and inform strategies for repurposing BMCs in synthetic biology [72]. To simplify the β-carboxysome composition, a single peptide CcmC was synthesized to replace endogenous CcmM, CcmN, and CcaA in *Synechococcus* 7942, leading to formation of β-carboxysome-like structures [105]. By expressing four shell proteins (*ccmK1*, *ccmK2*, *ccmL*, *ccmO*) from the cyanobacterium *Halothece* sp. PCC 7418, β-carboxysome mini-shells with a diameter of 20-25 nm have been reconstituted in *E. coli* [106,107]. Expressing six α-carboxysome shell proteins and CsoS2 in *E. coli* has led to the production of native-like α-carboxysome shells, which were further utilized as a nanoreactor to enhance the $H_2$-production capacities of encapsulated oxygen-sensitive hydrogenases, taking advantage of the $O_2$-limited internal environment created by the carboxysome shell [62].

The role of carboxysomes in enhancing carbon fixation and the feasibility of engineering carboxysome structures in heterologous hosts have inspired increasing efforts to introduce carboxysomes into crop plants for improving photosynthesis and crop yields [4,108,109]. A milestone work was expressing cyanobacterial form IB Rubisco along with the short-form CcmM to replace endogenous Rubisco in the chloroplasts of the model crop tobacco [110]. Recent work has made a step forward to produce α-carboxysome-like structures in tobacco chloroplasts, by expressing form IA Rubisco together with a shell protein CsoS1A and the linker protein CsoS2 from *Cyanobium marinum* PCC 7001 [111].

Progress has also been made recently for installing functional CCM pathways in *E. coli*. Expression of the genes encoding bacterial CCM components, including the inorganic carbon transporters DabAB and α-carboxysome proteins, allowed the establishment of a functional CCM network in an *E. coli* mutant that requires Rubisco activity for growth [103]. The introduced CCM enables the engineered *E. coli* strain to grow mixotrophically at ambient $CO_2$ levels. All these studies pave the way for reprogramming the metabolism of industrially important organisms for enhanced carbon assimilation.

## Concluding remarks

Systematic studies on carboxysome biosynthesis, structure, and function offer mechanistic insights into the building principles and evolutionary diversity of bacterial organelles for $CO_2$ fixation and microbial ecology. Advanced knowledge about the self-assembly and modularity of carboxysomes and the development of synthetic biology techniques provide a framework for sustainably engineering these carbon fixation factories or pathways in other organisms, and will foster further biotechnological advances in new contexts, e.g. crop food production, bioenergy production, and development of new carrier nanomaterials for the delivery of compounds and drug molecules. Future efforts are expected to decipher the molecular mechanisms and diversity of carboxysome assembly and functional regulation, and to develop synthetic biology approaches to promote design and reprogramming of carboxysome-based "factories" for specific applications (Outstanding questions).

### Outstanding questions

- What are the stoichiometric composition and variations of carboxysome proteins in the physiological context?
- What are the molecular mechanisms and regulation of shell permeability?
- What are the encapsulation mechanisms of cargo enzymes?
- What governs the diversity of structures and biogenesis pathways of carboxysomes and metabolosomes in an evolutionary context?

- What is the physiological coordination of carboxysomes with ancillary proteins and other CCM components to fulfil metabolic activities and regulation?
- What are the key factors that should be considered and manipulated in the design and sustainable engineering of functional $CO_2$-fixing modules based on the carboxysome structures?

**Glossary**

**Bacterial microcompartments**: proteinaceous megadalton-complexes that encapsulate metabolic pathways within the subcellular "micro-factories" using a virus-like protein shell to enhance enzymatic functions.

**Carboxysome**: a type of bacterial microcompartment that sequesters Rubisco and carbonic anhydrase (CA) to increase the rate of carbon fixation.

**Metabolosomes**: the bacterial microcompartments that encase pathway enzymes for the catabolism of various metabolites, including choline, ethanolamine, and 1,2-propanediol.

**Carboxylation**: a chemical process that is catalyzed by Rubisco to fix $CO_2$ to the five-carbon compound ribulose-1,5-bisphosphate (RuBP) and the splitting of the resulting six-carbon compound into two molecules of the three-carbon compound 3-phosphoglycerate (3-PGA).

**Oxygenation**: a chemical process in which Rubisco adds $O_2$ to RuBP to produce 3-PGA and 2-phosphoglycolate (2-PG). The latter is subsequently recycled along the following metabolic pathway known as photorespiration.

**Organelles**: the specialized structures that compartmentalize specific enzymes, pathways, and functions within a cell to improve and regulate metabolic performance.

**Photorespiration**: a process including the oxygenation of RuBP by Rubisco and release of $CO_2$ from organic compounds. It is assumed to be a wasteful process, as it causes a net loss of $CO_2$ and consumes energy (ATP and NADPH) produced in photosynthesis. However, photorespiration plays a protective role in regulating photosynthesis and is especially vital under stress conditions.

**Chemoautotrophs**: organisms that derive energy from the oxidation of inorganic compounds.

**Anoxygenic phototrophs**: organisms that can grow using energy from solar light and perform photosynthesis without evolving oxygen.

**Acknowledgments**

We acknowledge that we could not include and cite many other important papers in this review due to space constraints. This work was supported by the Royal Society (URF\R\180030, RGF\EA\181061, RGF\EA\180233), the Biotechnology and Biological Sciences Research Council Grant (BB/V009729/1, BB/M024202/1, BB/R003890/1), the National Natural Science Foundation of China (32070109), and the Leverhulme Trust (RPG-2021-286).

**References**

1. Ducat, D.C. and Silver, P.A. (2012) Improving carbon fixation pathways. Curr Opin Chem Biol 16 (3-4), 337-44.
2. Field, C.B. et al. (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. Science 281 (5374), 237-240.
3. Bar-On, Y.M. and Milo, R. (2019) The global mass and average rate of rubisco. Proceedings of the National Academy of Sciences 116 (10), 4738-4743.
4. Hennacy, J.H. and Jonikas, M.C. (2020) Prospects for Engineering Biophysical $CO_2$ Concentrating Mechanisms into Land Plants to Enhance Yields. Annu Rev Plant Biol 71, 461-485.
5. Sui, N. et al. (2020) Photosynthesis in Phytoplankton: Insights from the Newly Discovered Biological Inorganic Carbon Pumps. Molecular Plant 13 (7), 949-951.
6. Kerfeld, C.A. et al. (2018) Bacterial microcompartments. Nat Rev Microbiol 16 (5), 277-290.
7. Price, G.D. et al. (2008) Advances in understanding the cyanobacterial CO2-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants. J Exp Bot 59 (7), 1441-1461.
8. Hagemann, M. et al. (2016) Evolution of photorespiration from cyanobacteria to land plants, considering protein phylogenies and acquisition of carbon concentrating mechanisms. J Exp Bot 67 (10), 2963-2976.

9. Wang, C. et al. (2019) Structural mechanism of the active bicarbonate transporter from cyanobacteria. Nature Plants 5 (11), 1184-1193.
10. Selim, K.A. et al. (2018) PII-like signaling protein SbtB links cAMP sensing with cyanobacterial inorganic carbon response. Proc Natl Acad Sci U S A 115 (21), E4861-E4869.
11. Fang, S. et al. (2021) Molecular mechanism underlying transport and allosteric inhibition of bicarbonate transporter SbtA. Proc Natl Acad Sci U S A 118 (22), e2101632118.
12. Schuller, J.M. et al. (2020) Redox-coupled proton pumping drives carbon concentration in the photosynthetic complex I. Nature Communications 11 (1), 494.
13. Desmarais, J.J. et al. (2019) DABs are inorganic carbon pumps found throughout prokaryotic phyla. Nat Microbiol 4 (12), 2204–2215.
14. Mangiapia, M. et al. (2017) Proteomic and Mutant Analysis of the $CO_2$ Concentrating Mechanism of Hydrothermal Vent Chemolithoautotroph *Thiomicrospira crunogena*. J Bacteriol 199 (7), e00871-16.
15. Scott, K.M. et al. (2019) Diversity in $CO_2$-Concentrating Mechanisms among Chemolithoautotrophs from the Genera *Hydrogenovibrio*, *Thiomicrorhabdus*, and *Thiomicrospira*, Ubiquitous in Sulfidic Habitats Worldwide. Appl Environ Microbiol 85 (3), e02096-18.
16. Greening, C. and Lithgow, T. (2020) Formation and function of bacterial organelles. Nature Reviews Microbiology 18 (12), 677-689.
17. Gabaldon, T. and Pittis, A.A. (2015) Origin and evolution of metabolic sub-cellular compartmentalization in eukaryotes. Biochimie 119, 262-268.
18. Mullineaux, C.W. and Liu, L.N. (2020) Membrane dynamics in phototrophic bacteria. Annual Review of Microbiology 74, 633-654.
19. Axen, S.D. et al. (2014) A taxonomy of bacterial microcompartment loci constructed by a novel scoring method. PLoS Comput Biol 10 (10), e1003898.
20. Sutter, M. et al. (2021) A catalog of the diversity and ubiquity of bacterial microcompartments. Nat Commun 12 (1), 3809.
21. Kerfeld, C.A. and Erbilgin, O. (2015) Bacterial microcompartments and the modular construction of microbial metabolism. Trends Microbiol 23 (1), 22-34.
22. Bobik, T.A. et al. (2015) Bacterial microcompartments: widespread prokaryotic organelles for isolation and optimization of metabolic pathways. Mol Microbiol 98 (2), 193-207.
23. Kerfeld, C.A. and Sutter, M. (2020) Engineered bacterial microcompartments: apps for programming metabolism. Curr Opin Biotechnol 65, 225-232.
24. Ochoa, J.M. and Yeates, T.O. (2021) Recent structural insights into bacterial microcompartment shells. Curr Opin Microbiol 62, 51-60.
25. Liu, L.N. et al. (2021) Protein stoichiometry, structural plasticity and regulation of bacterial microcompartments. Curr Opin Microbiol 63, 133-141.
26. Liu, L.N. (2021) Bacterial metabolosomes: new insights into their structure and bioengineering. Microbial Biotechnology 14 (1), 88-93.
27. Rae, B.D. et al. (2013) Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate $CO_2$ fixation in cyanobacteria and some proteobacteria. Microbiol Mol Biol Rev 77 (3), 357-379.
28. Badger, M.R. et al. (2002) Evolution and diversity of $CO_2$ concentrating mechanisms in cyanobacteria. Funct Plant Biol 29 (3), 161-173.
29. Kerfeld, C.A. and Melnicki, M.R. (2016) Assembly, function and evolution of cyanobacterial carboxysomes. Curr Opin Plant Biol 31, 66-75.
30. Melnicki, M.R. et al. (2021) Evolutionary relationships among shell proteins of carboxysomes and metabolosomes. Curr Opin Microbiol 63, 1-9.
31. Gupta, R.S. (2012) Origin and spread of photosynthesis based upon conserved sequence features in key bacteriochlorophyll biosynthesis proteins. Mol Biol Evol 29 (11), 3397-412.
32. Giessen, T.W. and Silver, P.A. (2017) Engineering carbon fixation with artificial protein organelles. Curr Opin Biotechnol 46, 42-50.
33. Sommer, M. et al. (2017) β-Carboxysome bioinformatics: identification and evolution of new bacterial microcompartment protein gene classes and core locus constraints. J Exp Bot 68 (14), 3841-3855.
34. Yang, M. et al. (2020) Decoding the stoichiometric composition and organisation of bacterial metabolosomes. Nat Commun 11 (1), 1976.
35. Sun, Y. et al. (2019) Single-organelle quantification reveals the stoichiometric and structural variability of carboxysomes dependent on the environment. Plant Cell 31 (7), 1648-1664.
36. Faulkner, M. et al. (2017) Direct characterization of the native structure and mechanics of cyanobacterial carboxysomes. Nanoscale 9 (30), 10662-10673.
37. Sun, H. et al. (2021) Complex structure reveals CcmM and CcmN form a heterotrimeric adaptor in beta-carboxysome. Protein Sci 30 (8), 1566-1576.

38. Kinney, J.N. et al. (2012) Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly. J Biol Chem 287 (21), 17729-17736.
39. Long, B.M. et al. (2011) Over-expression of the beta-carboxysomal CcmM protein in Synechococcus PCC7942 reveals a tight co-regulation of carboxysomal carbonic anhydrase (CcaA) and M58 content. Photosynth Res 109 (1-3), 33-45.
40. Rae, B.D. et al. (2012) Structural determinants of the outer shell of beta-carboxysomes in *Synechococcus elongatus* PCC 7942: roles for CcmK2, K3-K4, CcmO, and CcmL. PLoS One 7 (8), e43871.
41. Sutter, M. et al. (2019) Structural Characterization of a Synthetic Tandem-Domain Bacterial Microcompartment Shell Protein Capable of Forming Icosahedral Shell Assemblies. ACS Synth Biol 8 (4), 668-674.
42. Hagen, A. et al. (2018) Programmed loading and rapid purification of engineered bacterial microcompartment shells. Nat Commun 9 (1), 2881.
43. Greber, B.J. et al. (2019) The Plasticity of Molecular Interactions Governs Bacterial Microcompartment Shell Assembly. Structure 27 (5), 749-763 e4.
44. Cannon, G.C. and Shively, J.M. (1983) Characterization of a homogenous preparation of carboxysomes from *Thiobacillus neapolitanus*. Archives of Microbiology 134 (1), 52-59.
45. Heinhorst, S. et al. (2006) Carboxysomes and Carboxysome-like Inclusions. In Complex Intracellular Structures in Prokaryotes (Shively, J.M. ed), pp. 141-165, Springer Berlin Heidelberg.
46. Iancu, C.V. et al. (2007) The structure of isolated *Synechococcus* strain WH8102 carboxysomes as revealed by electron cryotomography. J Mol Biol 372 (3), 764-773.
47. Roberts, E.W. et al. (2012) Isolation and characterization of the *Prochlorococcus* carboxysome reveal the presence of the novel shell protein CsoS1D. J Bacteriol 194 (4), 787-795.
48. Huang, F. et al. (2020) Rubisco accumulation factor 1 (Raf1) plays essential roles in mediating Rubisco assembly and carboxysome biogenesis. Proc Natl Acad Sci USA 117 (29), 17418-17428.
49. Hayer-Hartl, M. and Hartl, F.U. (2020) Chaperone Machineries of Rubisco - The Most Abundant Enzyme. Trends Biochem Sci 45 (9), 748-763.
50. Huang, F. et al. (2019) Roles of RbcX in carboxysome biosynthesis in the cyanobacterium *Synechococcus elongatus* PCC7942. Plant Physiol 179 (1), 184-194.
51. Mueller-Cajar, O. et al. (2011) Structure and function of the AAA+ protein CbbX, a red-type Rubisco activase. Nature 479 (7372), 194-199.
52. Tsai, Y.C. et al. (2015) Identification and characterization of multiple rubisco activases in chemoautotrophic bacteria. Nat Commun 6, 8883.
53. Sutter, M. et al. (2015) Structural characterization of a newly identified component of α-carboxysomes: The AAA+ domain protein CsoCbbQ. Sci Rep 5, 16243.
54. Tsai, Y.C. et al. (2020) Insights into the mechanism and regulation of the CbbQO-type Rubisco activase, a MoxR AAA+ ATPase. Proc Natl Acad Sci U S A 117 (1), 381-387.
55. Chen, T. et al. (2021) Incorporation of Functional Rubisco Activases into Engineered Carboxysomes to Enhance Carbon Fixation. ACS Synth Biol, https://doi.org/10.1021/acssynbio.1c00311.
56. Flecken, M. et al. (2020) Dual Functions of a Rubisco Activase in Metabolic Repair and Recruitment to Carboxysomes. Cell 183 (2), 457-473 e20.
57. Lechno-Yossef, S. et al. (2020) Cyanobacterial carboxysomes contain an unique rubisco-activase-like protein. New Phytol 225 (2), 793-806.
58. Long, B.M. et al. (2007) Analysis of carboxysomes from *Synechococcus* PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. J Biol Chem 282 (40), 29323-29335.
59. Long, B.M. et al. (2010) Functional cyanobacterial beta-carboxysomes have an absolute requirement for both long and short forms of the CcmM protein. Plant Physiol 153 (1), 285-293.
60. Ryan, P. et al. (2019) The small RbcS-like domains of the beta-carboxysome structural protein CcmM bind RubisCO at a site distinct from that binding the RbcS subunit. J Biol Chem 294 (8), 2593-2603.
61. Wang, H. et al. (2019) Rubisco condensate formation by CcmM in beta-carboxysome biogenesis. Nature 566 (7742), 131-135.
62. Li, T. et al. (2020) Reprogramming bacterial protein organelles as a nanoreactor for hydrogen production. Nature Communications 11, 5448.
63. Cai, F. et al. (2015) Advances in understanding carboxysome assembly in *Prochlorococcus* and *Synechococcus* implicate CsoS2 as a critical component. Life (Basel) 5 (2), 1141-1171.
64. Oltrogge, L.M. et al. (2020) Multivalent interactions between CsoS2 and Rubisco mediate alpha-carboxysome formation. Nat Struct Mol Biol 27 (3), 281-287.
65. Price, G.D. and Badger, M.R. (1989) Expression of Human Carbonic Anhydrase in the Cyanobacterium *Synechococcus* PCC7942 Creates a High $CO_2$-Requiring Phenotype: Evidence for a Central Role for Carboxysomes in the $CO_2$ Concentrating Mechanism. Plant Physiol 91 (2), 505-513.
66. Sawaya, M.R. et al. (2006) The structure of beta-carbonic anhydrase from the carboxysomal shell reveals a distinct subclass with one active site for the price of two. J Biol Chem 281 (11), 7546-7555.

67. McGurn, L.D. et al. (2016) The structure, kinetics and interactions of the beta-carboxysomal beta-carbonic anhydrase, CcaA. Biochem J 473 (24), 4559-4572.
68. So, A.K.-C. et al. (2002) Characterization of the C-terminal extension of carboxysomal carbonic anhydrase from *Synechocystis* sp. PCC6803. Functional Plant Biology 29 (3), 183-194.
69. Pena, K.L. et al. (2010) Structural basis of the oxidative activation of the carboxysomal gamma-carbonic anhydrase, CcmM. Proc Natl Acad Sci U S A 107 (6), 2455-2460.
70. Dai, W. et al. (2018) Visualizing Individual RuBisCO and Its Assembly into Carboxysomes in Marine Cyanobacteria by Cryo-Electron Tomography. J Mol Biol 430 (21), 4156-4167.
71. Schmid, M.F. et al. (2006) Structure of *Halothiobacillus neapolitanus* carboxysomes by cryo-electron tomography. J Mol Biol 364 (3), 526-535.
72. Sutter, M. et al. (2017) Assembly principles and structure of a 6.5-MDa bacterial microcompartment shell. Science 356 (6344), 1293-1297.
73. Sutter, M. et al. (2016) Visualization of bacterial microcompartment facet assembly using high-speed atomic force microscopy. Nano Letters 16 (3), 1590-1595.
74. Faulkner, M. et al. (2019) Self-assembly stability and variability of bacterial microcompartment shell proteins in response to the environmental change. Nanoscale Research Letters 14, 54.
75. Dou, Z. et al. (2008) $CO_2$ fixation kinetics of *Halothiobacillus neapolitanus* mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for $CO_2$. J Biol Chem 283 (16), 10377-10384.
76. Faulkner, M. et al. (2020) Molecular simulations unravel the molecular principles that mediate selective permeability of carboxysome shell protein. Scientific Reports 10 (1), 17501.
77. Park, J. et al. (2017) Molecular dynamics simulations of selective metabolite transport across the propanediol bacterial microcompartment shell. J Phys Chem B 121 (34), 8149-8154.
78. Bobik, T.A. and Stewart, A.M. (2021) Selective molecular transport across the protein shells of bacterial microcompartments. Curr Opin Microbiol 62, 76-83.
79. Chowdhury, C. et al. (2015) Selective molecular transport through the protein shell of a bacterial microcompartment organelle. Proc Natl Acad Sci U S A 112 (10), 2990-2995.
80. Mahinthichaichan, P. et al. (2018) Selective Permeability of Carboxysome Shell Pores to Anionic Molecules. J Phys Chem B 122 (39), 9110-9118.
81. Klein, M.G. et al. (2009) Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport. J Mol Biol 392 (2), 319-33.
82. Cai, F. et al. (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 (22), 16055-16063.
83. Larsson, A.M. et al. (2017) Crystal structures of β-carboxysome shell protein CcmP: ligand binding correlates with the closed or open central pore. J Exp Bot 68 (14), 3857-3867.
84. Sommer, M. et al. (2019) Heterohexamers formed by CcmK3 and CcmK4 increase the complexity of beta carboxysome Shells. Plant Physiol 179 (1), 156-167.
85. Garcia-Alles, L.F. et al. (2019) Occurrence and stability of hetero-hexamer associations formed by beta-carboxysome CcmK shell components. PLoS One 14 (10), e0223877.
86. Menon, B.B. et al. (2010) The carboxysome shell is permeable to protons. J Bacteriol 192 (22), 5881-6.
87. Chen, A.H. et al. (2013) The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo. PLoS One 8 (9), e76127.
88. Long, B.M. et al. (2021) Rubisco proton production can drive the elevation of $CO_2$ within condensates and carboxysomes. Proc Natl Acad Sci U S A 118 (18).
89. Cameron, J.C. et al. (2013) Biogenesis of a bacterial organelle: the carboxysome assembly pathway. Cell 155 (5), 1131-40.
90. Iancu, C.V. et al. (2010) Organization, structure, and assembly of alpha-carboxysomes determined by electron cryotomography of intact cells. J Mol Biol 396 (1), 105-117.
91. Savage, D.F. et al. (2010) Spatially ordered dynamics of the bacterial carbon fixation machinery. Science 327 (5970), 1258-61.
92. MacCready, J.S. and Vecchiarelli, A.G. (2021) Positioning the Model Bacterial Organelle, the Carboxysome. mBio 12 (3), e02519-19.
93. Schumacher, M.A. et al. (2019) Structures of maintenance of carboxysome distribution Walker-box McdA and McdB adaptor homologs. Nucleic Acids Res 47 (11), 5950-5962.
94. MacCready, J.S. et al. (2018) Protein gradients on the nucleoid position the carbon-fixing organelles of cyanobacteria. Elife 7, e39723.
95. Hill, N.C. et al. (2020) Life cycle of a cyanobacterial carboxysome. Sci Adv 6 (19), eaba1269.
96. MacCready, J.S. et al. (2020) Origin and Evolution of Carboxysome Positioning Systems in Cyanobacteria. Mol Biol Evol 37 (5), 1434-1451.
97. MacCready, J.S. et al. (2021) The McdAB system positions α-carboxysomes in proteobacteria. Mol Microbiol 16 (1), 277-297.

98. Sun, Y. et al. (2020) Diurnal regulation of *in vivo* localization and $CO_2$-fixing activity of carboxysomes in *Synechococcus elongatus* PCC 7942. Life (Basel) 10 (9), 169.
99. Sun, Y. et al. (2016) Light modulates the biosynthesis and organization of cyanobacterial carbon fixation machinery through photosynthetic electron flow. Plant Physiol 171 (1), 530-541.
100. Barchewitz, T. et al. (2019) Non-canonical localization of RubisCO under high-light conditions in the toxic cyanobacterium *Microcystis aeruginosa* PCC7806. Environ Microbiol 21 (12), 4836-4851.
101. Bonacci, W. et al. (2012) Modularity of a carbon-fixing protein organelle. Proc Natl Acad Sci U S A 109 (2), 478-83.
102. Baumgart, M. et al. (2017) Heterologous expression of the *Halothiobacillus neapolitanus* carboxysomal gene cluster in *Corynebacterium glutamicum*. J Biotechnol 258, 126-135.
103. Flamholz, A.I. et al. (2020) Functional reconstitution of a bacterial $CO_2$ concentrating mechanism in *Escherichia coli*. Elife 9, e59882.
104. Fang, Y. et al. (2018) Engineering and modulating functional cyanobacterial $CO_2$-fixing organelles. Front Plant Sci 9, 739.
105. Gonzalez-Esquer, C.R. et al. (2015) Streamlined construction of the cyanobacterial $CO_2$-fixing organelle via protein domain fusions for use in plant synthetic biology. Plant Cell 27 (9), 2637-44.
106. Cai, F. et al. (2016) Production and Characterization of Synthetic Carboxysome Shells with Incorporated Luminal Proteins. Plant Physiol 170 (3), 1868-1877.
107. Sutter, M. et al. (2019) Structure of a Synthetic beta-Carboxysome Shell. Plant Physiol 181 (3), 1050-1058.
108. Hanson, M.R. et al. (2016) Towards engineering carboxysomes into C3 plants. Plant J 87 (1), 38-50.
109. McGrath, J.M. and Long, S.P. (2014) Can the cyanobacterial carbon-concentrating mechanism increase photosynthesis in crop species? A theoretical analysis. Plant Physiol 164 (4), 2247-2261.
110. Lin, M.T. et al. (2014) β-Carboxysomal proteins assemble into highly organized structures in *Nicotiana* chloroplasts. Plant J 79 (1), 1-12.
111. Long, B.M. et al. (2018) Carboxysome encapsulation of the $CO_2$-fixing enzyme Rubisco in tobacco chloroplasts. Nature Communications 9 (1), 3570.
112. Bracher, A. et al. (2017) Biogenesis and metabolic maintenance of Rubisco. Annu Rev Plant Biol 68, 29-60.
113. Spreitzer, R.J. (2003) Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. Arch Biochem Biophys 414 (2), 141-149.
114. Banda, D.M. et al. (2020) Novel bacterial clade reveals origin of form I Rubisco. Nat Plants 6 (9), 1158-1166.
115. Sutter, M. et al. (2013) Two new high-resolution crystal structures of carboxysome pentamer proteins reveal high structural conservation of CcmL orthologs among distantly related cyanobacterial species. Photosynth Res 118 (1-2), 9-16.
116. Wheatley, N.M. et al. (2013) Bacterial microcompartment shells of diverse functional types possess pentameric vertex proteins. Protein Science 22 (5), 660-665.
117. Ochoa, J.M. et al. (2021) MCPdb: The bacterial microcompartment database. PLoS One 16 (3), e0248269.
118. Kerfeld, C.A. et al. (2010) Bacterial microcompartments. Annu Rev Microbiol 64, 391-408.
119. Kalnins, G. et al. (2020) Encapsulation mechanisms and structural studies of GRM2 bacterial microcompartment particles. Nat Commun 11 (1), 388.