Protein-based organelles in bacteria: carboxysomes and related microcompartments

Todd O. Yeates[*‡], *Cheryl A. Kerfeld*[§||], *Sabine Heinhorst*[¶], *Gordon C. Cannon*[¶] *and Jessup M. Shively*[¶#]

Abstract | Many bacteria contain intracellular microcompartments with outer shells that are composed of thousands of protein subunits and interiors that are filled with functionally related enzymes. These microcompartments serve as organelles by sequestering specific metabolic pathways in bacterial cells. The carboxysome, a prototypical bacterial microcompartment that is found in cyanobacteria and some chemoautotrophs, encapsulates ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) and carbonic anhydrase, and thereby enhances carbon fixation by elevating the levels of CO_2 in the vicinity of RuBisCO. Evolutionarily related, but functionally distinct, microcompartments are present in diverse bacteria. Although bacterial microcompartments were first observed more than 40 years ago, a detailed understanding of how they function is only now beginning to emerge.

Bacterial microcompartment
A large, polyhedral, proteinaceous structure that functions as an organelle by encapsulating specific enzymes inside a protein shell that is reminiscent of a viral capsid.

*UCLA Department of Chemistry and Biochemistry, ‡UCLA-DOE Institute of Genomics and Proteomics, §US Department of Energy — Joint Genome Institute ||Department of Plant and Microbial Biology, University of California, Berkeley, California 94720, USA. ¶Department of Chemistry and Biochemistry, The University of Southern Mississippi, Hattiesburg, Mississippi 39406, USA. #Department of Genetics and Biochemistry, Clemson University, Clemson, South Carolina 29634, USA. Correspondence to T.O.Y. e-mail: yeates@mbi.ucla.edu
doi:10.1038/nrmicro1913
Published online 4 August 2008

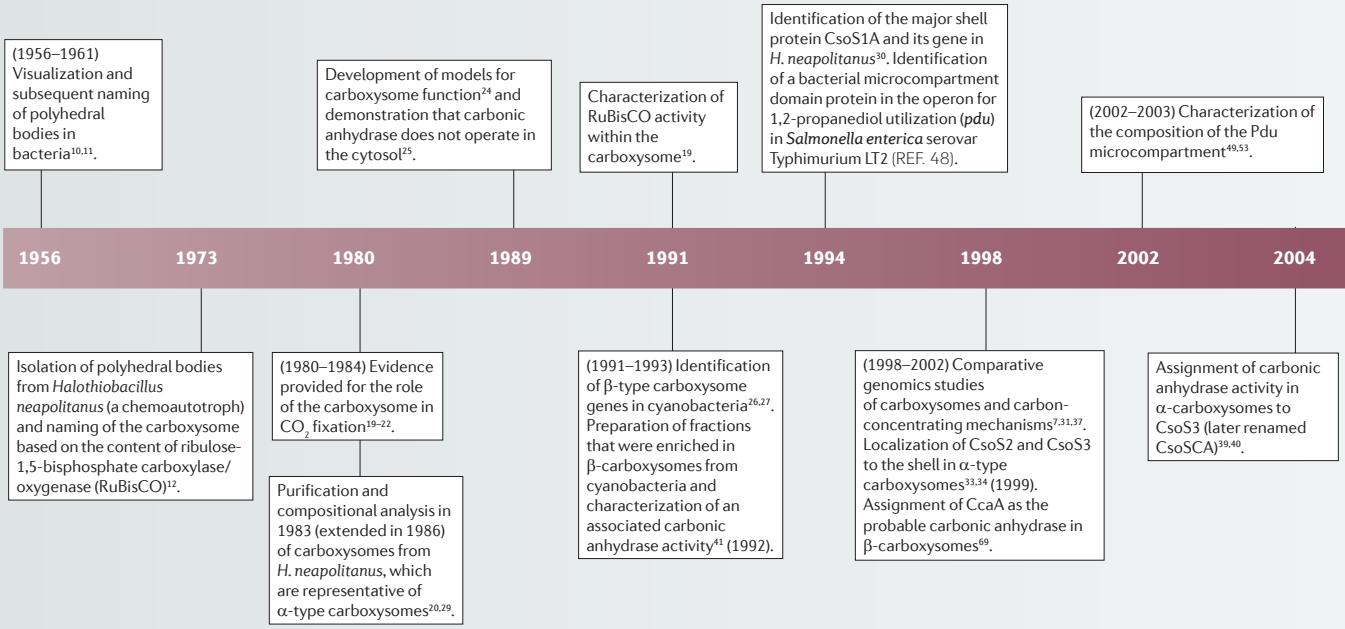
Current research in microbiology is painting a picture of bacterial cells with complex interiors[1]. Imaging studies continue to reveal a myriad of subcellular structures that have diverse architectures and functions (reviewed in REF. 2), and these direct observations have been complemented by protein-sequence and structure analyses. Such analyses have demonstrated, for example, the surprising existence within microorganisms of proteins that are homologous to those that organize the cytoskeletons of eukaryotic cells[3–5]. Various lines of investigation therefore indicate high levels of internal organization in bacteria. Arguably the highest possible level of subcellular organization is the spatial segregation of one region of a cell. Eukaryotic cells achieve this level of organization with their membrane-bound organelles, including mitochondria, chloroplasts, lysosomes and Golgi bodies. These organelles sequester specific proteins and metabolic pathways inside of — and control the flow of molecules in and out of — a defined, subcellular compartment. It is now clear that many bacteria contain proteinaceous microcompartments that function like organelles by sequestering specific enzymes within the bacterial cell (reviewed in REFS 6–8).

Sequestering or colocalizing multiple enzymes from a metabolic pathway can substantially increase the efficiency of that pathway; the product of one enzyme reaction is delivered at high concentration as a substrate for the next enzyme[9]. Furthermore, if colocalized enzymes are surrounded by a barrier that has specific transport or permeability properties, then additional advantages of control and specificity can be gained. These principles are observed in eukaryotic organelles and in subcellular structures that are referred to as bacterial microcompartments. The first bacterial microcompartment to be characterized in detail was the carboxysome. Here, we briefly summarize research on the carboxysome, a bacterial microcompartment that enhances CO_2 fixation, and evolutionarily related microcompartments, and highlight recent structural studies and their mechanistic implications.

A brief history of carboxysome biology

The occurrence of inclusion bodies with a polygonal appearance in prokaryotes was first reported in 1956 in the cyanobacterium *Phormidium uncinatum*[10] (TIMELINE). Over the next 20 years, these inclusions, which were named polyhedral bodies in 1961, were found in various other cyanobacteria (for example, *Nostoc punctiforme*, *Synechococcus elongatus* (previously known as *Anacystis nidulans*), *Anabaena cylindrica* and *Symploca muscorum*) and chemoautotrophic bacteria (for example, *Halothiobacillus neapolitanus*, *Acidithiobacillus ferrooxidans*, *Nitrobacter winogradskyi* and *Nitrococcus mobilis*)[11] (FIG. 1). Polyhedral

Timeline | **Highlights in the discovery and characterization of carboxysomes**



Carboxysome
A polyhedral bacterial microcompartment that enhances carbon fixation by encapsulating the ribulose-1,5-bisphosphate carboxylase/oxygenase and carbonic anhydrase enzymes.

bodies were first isolated in 1973 from the chemoautotrophic bacterium *H. neapolitanus* by differential and sucrose step-gradient centrifugation, following rupture in a French pressure cell. These bodies were surrounded by a proteinaceous envelope (the shell) and contained the enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO); to reflect this finding, the name carboxysome was proposed[12]. Polyhedral bodies or carboxysomes have since been visualized in numerous cyanobacteria and other chemoautotrophic bacteria[13–15]. It is now widely accepted that although carboxysomes can probably be found in all cyanobacteria, they are present in only a limited number of chemoautotrophs: namely, some of the sulphur and nitrifying bacteria (for example, species of *Thiobacillus*, *Halothiobacillus*, *Thiomonas*, *Acidithiobacillus*, *Nitrobacter* and *Nitrosomonas*)[6,16].

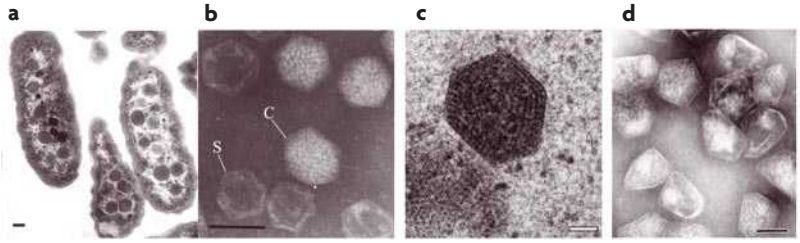
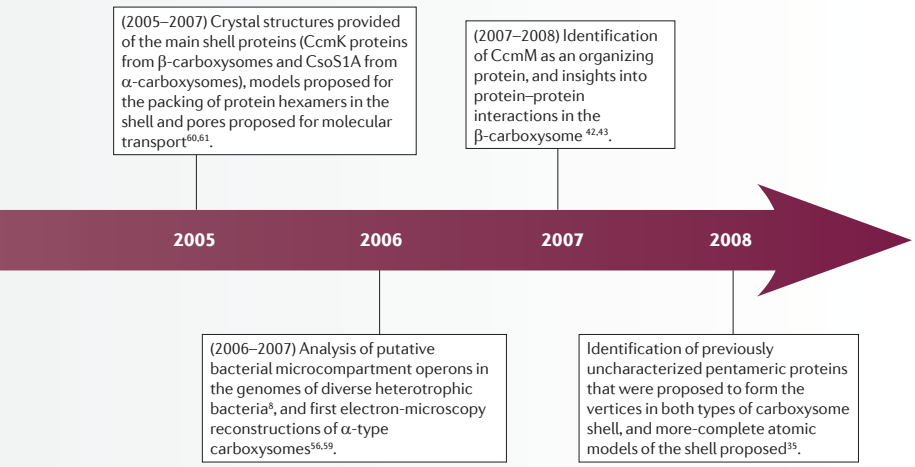


Figure 1 | **Carboxysomes and the Pdu microcompartment. a** | Transmission electron micrograph of thin-sectioned *Halothiobacillus neapolitanus* cells that reveals carboxysomes of the α-type. **b** | Purified carboxysomes from *H. neapolitanus* (C, carboxysome; S, shell). **c** | β-type carboxysomes from *Synechocystis* sp. PCC 6803 are shown in thin section. **d** | Isolated *pdu* microcompartments from *Salmonella enterica* serovar Typhimurium LT2 that were visualized by negative stain. The scale bars represent 100 nm (**a**, **b** and **d**) and 50 nm (**c**). Panels **a** and **b** courtesy of G.C. Cannon and S. Heinhorst, University of Southern Mississippi, Hattiesburg, USA . Panel **c** reproduced, with permission, from REF. 35 © (2008) American Association for the Advancement of Science. Panel **d** courtesy of T. Bobik, Iowa State University, Ames, USA.

The function of the carboxysome has been the subject of much discussion[13–15,17,18]. Two hypotheses were initially proposed: carboxysomes are either storage bodies, analogous to various other prokaryotic inclusions, or are simple microcompartments or organelles for the fixation of CO_2. The observation that the RuBisCO of intact, stable, purified carboxysomes was active in CO_2 fixation and had the same kinetic and inhibitor binding properties as the isolated cytoplasmic enzyme supported the hypothesis that carboxysomes are dedicated organelles for CO_2 fixation[19,20]. Furthermore, when cultures of cyanobacteria and chemoautotrophic bacteria were transferred to conditions in which CO_2 was limiting, RuBisCO production increased. Most of this RuBisCO was found within carboxysomes, which were present in increased numbers, and this provided additional evidence for the organelle hypothesis[21,22]. Subsequent experiments correlated mutations that caused deficient CO_2 fixation with defective carboxysome formation[23], and mathematical models supported the role of the carboxysome as an organelle for CO_2 fixation[24]. A key element was verified by an experiment which showed that ectopic expression of carbonic anhydrase activity in the cyanobacterial cytosol prevented CO_2 fixation[25], which was consistent with the requirement for carbonic anhydrase to be localized with RuBisCO in the carboxysome.



components of the carboxysome shell in their respective organisms[31]. Additional genes in *H. neapolitanus* that are homologous to *csoS1A* were subsequently identified both upstream (*csoS1C*) and downstream (*csoS1B*) of *csoS1A*[30]. Four additional open reading frames (ORFs) were subsequently found between the RuBisCO *cbbLS* and *csoS1C*[6,32]. All of the genes were oriented in the same direction and had a putative promoter upstream of *cbbL*, which indicated that these genes were organized as an operon[6,16,32] (FIG. 2a). The two ORFs that are immediately downstream of *cbbS* (*csoS2* and *csoS3*) were shown to encode proteins that are associated with the shell[33], whereas the presumptive gene products of the two remaining ORFs, which are homologous to each other, were not detected in significant amounts in preparations of carboxysomes[34]. These two genes were originally named *orfA* and *orfB*, but were renamed *csoS4A* and *csoS4B* after the structural characterization of this protein family[35]. Identification of the components of the carboxysome set the stage for subsequent structural and functional studies.

Carboxysome function

Oceanic microorganisms contribute substantially to the global carbon cycle[36]. Because most of the CO_2 fixation in these microorganisms (including all cyanobacteria) seems to occur within carboxysomes, understanding how carboxysomes enhance this process is of considerable interest. Carboxysomes can be divided into two types that differ in their component proteins and the organization of their respective genes (reviewed in REF. 37). Carboxysomes of the α-type are found in α-cyanobacteria (for example, *Prochlorococcus* species and certain *Synechococcus* species, such as *Synechococcus* sp. WH 8102), whereas carboxysomes of the β-type are found in β-cyanobacteria (for example, *S. elongatus* PCC 7942 and *Synechocystis* sp. PCC 6803). Organisms with α-carboxysomes contain type 1A RubBisCO and their carboxysome genes are arranged in a single operon, whereas organisms with β-carboxysomes contain type 1B RuBisCO and their carboxysome genes are typically arranged in multiple gene clusters. Not all of the genes that are present in one type of carboxysome can be identified in the genomes of organisms that produce the other type. For example, homologues of the *ccmM* and *ccmN* genes are present in β-carboxysomes but not α-carboxysomes, whereas homologues of the *csoS2* and *csoS3* genes are found in α-carboxysomes but not β-carboxysomes. Therefore, there could be functional differences between the two types that are not yet fully understood, and in fact our current understanding of carboxysomes is based on information from both types. Studies on the composition of carboxysomes have been based mainly on α-carboxysomes from the chemoautotroph *H. neapolitanus* (reviewed in REF. 16).

Carboxysomes carry out the final stages of the carbon-concentrating mechanism (CCM), through which autotrophic prokaryotes accumulate inorganic carbon to enhance CO_2 fixation (BOX 1). The first part of the CCM involves transmembrane pumps, which actively concentrate bicarbonate inside the cell

A few years later, a group of genes (*ccmKLMNO*) from an operon that is upstream of the genes for the large and small subunits of RuBisCO in *Synechococcus elongatus* PCC 7942 were individually inactivated by insertion mutagenesis. In all cases, the mutants either produced aberrant carboxysomes or lacked them completely, and growth of the mutants required increased CO_2 concentrations[7,26,27]. Similarly, inactivation of *csoS1A*, which encodes one of the carboxysome shell proteins of *H. neapolitanus*, resulted in mutants that had fewer carboxysomes than wild-type cells and required increased levels of CO_2 for growth[28]. These results substantiated the proposed role of carboxysomes in CO_2 fixation.

The purification of *H. neapolitanus* carboxysomes to apparent homogeneity was reported in 1983 (REF. 20), which allowed the subsequent correlation of carboxysome proteins with their genes. This and an additional study indicated that the carboxysome consists of only RuBisCO and a set of 7–11 other proteins that either form, or are attached to, the shell[29]. The main protein constituent of the shell (approximately 13% of the total carboxysome protein) was isolated and sequenced[30]. The gene that encodes this protein was subsequently cloned and named *csoS1A*, and was confirmed to encode a 98-amino-acid 9.9 kDa protein (FIG. 2a). The CsoS1A shell protein from *H. neapolitanus* shares significant sequence homology with the proteins that are encoded by the *ccmK* and *ccmO* genes of cyanobacteria (for example, *S. elongatus* PCC 7942), and were previously linked to carboxysome function by genetic methods, which indicated that those proteins were

CO_2 fixation
The process by which carbon in the biosphere is converted from an inorganic form (for example, CO_2) into organic molecules.

Ribulose-1,5-bisphosphate carboxylase/oxygenase
The enzyme that fixes carbon by combining CO_2 with the five-carbon compound ribulose-1,5-bisphosphate to form two molecules of the three-carbon compound phosphoglycerate.

Carbonic anhydrase
An enzyme that catalyses the interconversion of bicarbonate and CO_2.

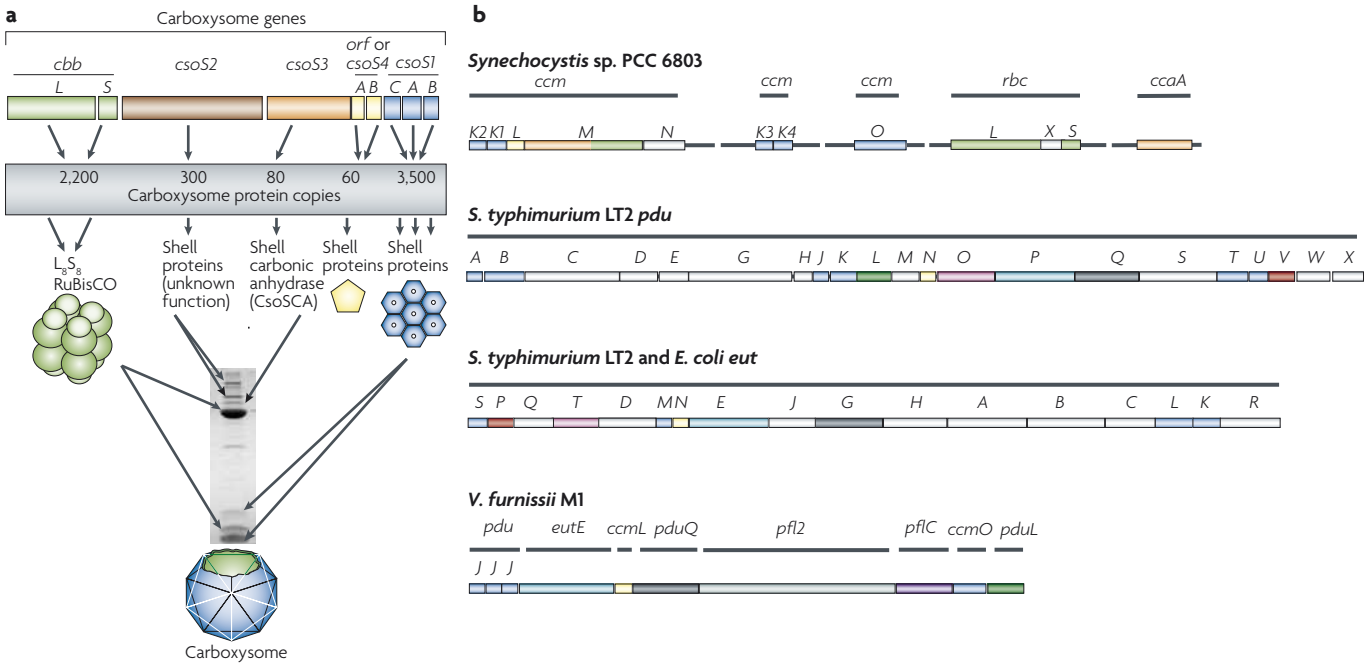
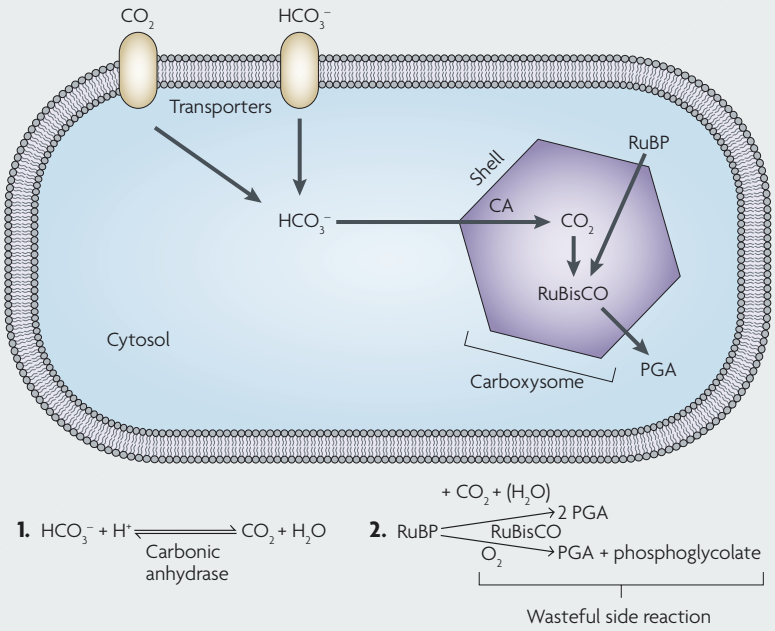


Figure 2 | **The genes and genomic organization of carboxysomes and related bacterial microcompartments.** **a** | Overview of the composition of α-type carboxysomes based on genetic and biochemical studies of *Halothiobacillus neapolitanus*. Paralogous genes are coloured similarly. The number of protein copies in a single carboxysome was estimated based on a combination of experimental data[16] and inferences from protein crystal structures[35,61]. The assembly features of some of the proteins are indicated. The SDS gel in the centre illustrates the proteins that have been identified in purified carboxysome preparations. The colouring of the carboxysome reflects the main constituents of the interior (ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO)) and the shell (CsoS1A, CsoS1B and CsoS1C). **b** | Bacterial gene clusters from other representative microcompartments. *Synechocystis* sp. PCC 6803 contains a β-type carboxysome, whereas the other three organisms (*Salmonella enterica* serovar Typhimurium LT2 (*S. typhimurium* LT2), *Escherichia coli* and *Vibrio furnissii* M1) contain various non-carboxysome microcompartments. Genes that share sequence homology are coloured in two parts similarly. CcaA is coloured orange to reflect its status as a β-type carbonic anhydrase. CcmM is coloured in two parts to reflect sequence similarity that originates from the γ-type carbonic anhydrase at its amino-terminal end and the presence of repeated domains in its carboxy-terminal end that are similar to the RuBisCO small subunit[26]; four such repeats are present in *Synechocystis* sp. PCC 6803 CcmM. Genes in white have no sequence homologues within the bacterial microcompartment clusters that are shown. The lengths of the segments are proportional to the number of amino acids in the full-length gene products, although intergenic regions in the *Synechocystis* sp. PCC 6803 genome are not to scale. Established or putative functions for the *pdu*-specific gene products are: B_{12}-dependent diol dehydratase for PduC, PduD and PduE; propionaldehyde dehydrogenase for PduP; propanol dehydrogenase for PduQ; adenosyl transferase for PduO; cobalamin reductase for PduS; diol dehydratase reactivase for PduG and PduH; phosphotransacylase for PduL; threonine kinase for PduX; and an unknown function for PduM and PduV[53,55,66]. Established or putative functions for the *eut*-specific gene products are: an unknown function for EutP and EutQ; corrinoid adenosyltransferase and cobalamin recycling for EutT; phosphotransacetylase for EutD; aldehyde dehydrogenase for EutE; a possible chaperone function for EutJ; alcohol dehydrogenase for EutG; permease for EutH; ethanolamine ammonia-lyase reactivase for EutA; ethanolamine ammonia lyase for EutB and EutC; and as a transcription activator for EutR[52,67]. In addition to several shell proteins and certain *pdu* and *eut* homologues, based on sequence homology, the *V. furnissii* bacterial microcompartment cluster also encodes a pyruvate formate lyase (*pfl2*) and a pyruvate formate lyase activator (*pflC*)[55].

(reviewed in REF. 38), whereas the second part involves the carboxysome, in which bicarbonate is converted to CO_2 in the viccinity of RuBisCO. Models for how the carboxysome contributes to the CCM predict that a carbonic anhydrase needs to be sequestered with RuBisCO in the carboxysome to convert bicarbonate, presumably as it enters the microcompartment from the cytosol, to the RuBisCO substrate CO_2 (BOX 1). Such a mechanism would supply the kinetically inefficient RuBisCO enzyme with a high effective concentration of CO_2 (reviewed in REF. 6). The crucial location of both carbonic anhydrase and RuBisCO in carboxysomes (rather than in the cytosol) was emphasized by the detrimental effects on CO_2 fixation after the ectopic expression of carbonic anhydrase in the cytosol[25]; in this situation, CO_2 is lost by diffusion out of the cell. In α-carboxysomes, CsoS3 represents only a few percent of the total carboxysome protein content and is associated with the shell[33] (FIG. 2a). Recombinant CsoS3 proteins from *H. neapolitanus*, and those from the marine cyanobacteria *Prochlorococcus marinus* subsp. *pastoris* str. CCMP1986 (also known as *Prochlorococcus marinus* MED4), *Prochlorococcus marinus* str. MIT 9313 and *Synechococcus* sp. WH 8102 that were produced in *Escherichia coli*, were recently shown to have carbonic anhydrase activity[39]. Based on the lack of detectable

Box 1 | **Carbon concentration and the carboxysome**

In the first part of the carbon-concentrating mechanism (CCM), bicarbonate is concentrated inside the cell by transporters in the cell membrane (reviewed in REF. 38). The carboxysome is involved in the second part of the CCM (see the figure for a schematic that shows the function of the carboxysome and its role in the CCM), and enhances CO_2 fixation by co-localizing the two enzymes ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) and carbonic anhydrase (CA) inside a thin shell that is assembled from thousands of protein subunits. Bicarbonate is thought to enter the carboxysome through pores in the proteinaceous shell, where it is converted to CO_2 for use by RuBisCO. The five-carbon compound ribulose-1,5-bisphosphate (RuBP) must also enter the carboxysome, and the three-carbon compound 3-phosphoglycerate (3-PGA; two molecules of which are formed by the addition of CO_2 to RuBP) must exit the carboxysome. Glycolate metabolism has been shown to be important in *Synechocystis* sp. PCC 6803 (REFS 64,65). If the phosphoglycolate arises from carboxysomal RuBisCO, it would imply that some O_2 enters the carboxysome and that phosphoglycolate exits into the cytosol. The two enzymes in the carboxysome and their reactions are shown in reactions 1 and 2.



sequence similarity to known carbonic anhydrases, CsoS3 was initially classified as a novel type of carbonic anhydrase, but was later assigned to the β-class of carbonic anhydrases based on structural similarities[40]; the CsoS3 protein was consequently renamed CsoSCA (carboxysome shell carbonic anhydrase). Organisms with β-type carboxysomes do not encode a homologue of CsoSCA. Instead, in most cyanobacteria that contain β-carboxysomes, the carbonic anhydrase CcaA is present. CcaA is found in partially purified preparations of β-carboxysomes[69], and has been shown to associate with other components of the shell[42,43].

Colocalization of RuBisCO and carbonic anhydrase is a crucial feature that occurs within the carboxysome, but a key question has been whether colocalization of the two enzymes in the proper configuration is sufficient to produce the observed CO_2-fixing efficiencies or whether diffusive loss of CO_2 from the carboxysome would also need to be mitigated[44]. The outer shell could provide this role, as recent experiments have shown that the outer shell provides a barrier to the diffusion of CO_2 and bicarbonate substrates[45,46]. However, how this diffusive barrier impedes CO_2 passage without preventing the entry and exit of larger substrates and products remains a puzzle that has focused attention on the composition and structure of the shell.

The carboxysome shell

The enzymatic components of the carboxysome — RuBisCO and carbonic anhydrase — are surrounded by a thin protein shell. In α-carboxysomes, the main shell components are the CsoS1 proteins. These small proteins (typically 9–11 kDa) are encoded by at least two paralogous genes in all carboxysome-forming bacteria (FIG. 2a, TABLE 1). The sequence similarity between different paralogues of the main shell proteins is generally high within a given organism. In *H. neapolitanus*, the main shell proteins CsoS1A and CsoS1C differ by only two amino acids, whereas CsoS1B is 90% identical in its amino acid sequence and has 12 more amino acids at its carboxyl terminus. A recent quantification of carboxysome gene transcripts in *H. neapolitanus* revealed transcripts for all three *csoS1* paralogues (*csoS1A–C*) at concentrations that correlated with the numbers of the respective protein products in the α-carboxysome[47]. The main shell proteins (CsoS1A–C) account for approximately 17% of the carboxysome by weight, whereas RuBisCO accounts for approximately 70% of the carboxysome by weight. In *Synechocystis* sp. PCC 6803, an organism that produces β-carboxysomes, the homologues of the CsoS1 proteins are named CcmK1–4 and another homologue, CcmO, is also encoded. Crystal structures have recently been determined for representative shell proteins from both α-carboxysomes and β-carboxysomes, and have provided insight into how the shell functions (discussed below).

In addition to the main shell proteins, a few other proteins do not seem to be enzymatic and probably have structural roles in the shell (TABLE 1). In α-carboxysomes, one of these proteins is CsoS2. The proposal that CsoS2 is located in the shell has been supported by immuno-electron microscopy, which indicated that the protein is located mainly at the periphery of carboxysomes[33].

Table 1 | **Carboxysome proteins***

Protein	Function
α-type carboxysome	
CsoSCA (also known as CsoS3)	Carboxysomal carbonic anhydrase; dehydration of bicarbonate to form CO_2 (REFS 34,39,40)
CsoS1A–C‡¶	Main shell proteins[26,30]; members of the BMC protein family; form hexamers that further assemble into a tightly packed layer that contains narrow pores[61]; conserved across other non-carboxysome microcompartments
CsoS2§	Unknown
CsoS4A and CsoS4B (previously known as OrfA and OrfB, respectively)#	Crystal structures have revealed that CsoS4A is a pentamer that probably forms the vertices of the shell[35]; conserved across other non-carboxysome microcompartments
Form 1A RuBisCO large and small subunits‖ (REF. 12)	CO_2 fixation; formation of two molecules of the three-carbon compound 3-PGA from the five-carbon compound RuBP and CO_2
β-type carboxysome	
CcaA	Carboxysomal carbonic anhydrase; dehydration of bicarbonate to form CO_2 (REFS 41,42,69)
CcmK1–4‡¶ and CcmO¶	Main shell proteins[26,30]; members of the BMC protein family; form hexamers that further assemble into a tightly packed layer that contains narrow pores[60]; conserved across other non-carboxysome microcompartments
CcmL#	Crystal structures have revealed that this protein is a pentamer that probably forms the vertices of the shell[35]; conserved across other non-carboxysome microcompartments
CcmN	Unknown
CcmM	Probably a scaffold or organizing protein[42,43]; contains separate regions that are homologous to γ-carbonic anhydrases and the RuBisCO small subunit[26]
Form 1B RuBisCO large and small subunits‖	CO_2 fixation; formation of two molecules of 3-PGA from RuBP and CO_2

*Proteins were chosen to be representative of the two types of carboxysome, although some species variations exist. ‡Proteins are present in multiple paralogous copies in a single genome (for example, CcmK1–4 or CsoS1A–C). CcmO contains two tandem BMC domains in one polypeptide chain. §Migrates as two distinct forms by SDS–PAGE, which has been attributed to differential glycosylation[33]. ‖Proteins are homologous. ¶Proteins are homologous. #Proteins are homologous. 3-PGA, 3-phosphoglycerate; BMC, bacterial microcompartment; RuBisCO, ribulose-1,5-bisphosphate carboxylase/oxygenase; RuBP, ribulose-1,5-bisphosphate

A molecular mass of 92 kDa was calculated from the genomic sequence, whereas in SDS gels the protein appeared as two bands that corresponded to two different molecular weights (85 kDa and 130 kDa), which has been attributed to differential glycosylation[33]. The function of CsoS2 is unknown, but a high isoelectric point value has been noted (>9)[16]; the positive charge on CsoS2 could facilitate diffusion of the negatively charged substrates and products of carbon fixation (bicarbonate, 3-phosphoglycerate (3PGA) and ribulose-1,5-bisphosphate (RuBP)) across the microcompartment shell. Homologues of CsoS2 are not present in β-carboxysomes, and the functions of other β-carboxysome proteins, such as CcmM and CcmN, are still unclear. However, recent biochemical and genetic experiments have established protein–protein interactions between CcmM, other carboxysome proteins, including CcmN and CcaA, and RuBisCO[42,43], which has implicated CcmM as a key organizing protein.

Finally, a set of homologous proteins — named CcmL in β-carboxysomes, and CsoS4A and CsoS4B in α-carboxysomes — is encoded in operons together with other carboxysome proteins, but is not present in detectable numbers in isolated carboxysome preparations (that is, CsoS4A and CsoS4B were not evident in purified *H. neapolitanus* carboxysomes) (TABLE 1). Nonetheless, the importance of these proteins was highlighted by genetic experiments in which the deletion of *ccmL* in *S. elongatus* PCC 7942 yielded mutants that contained elongated carboxysomes[26]. This suggested that the CcmL and CsoS4A or CsoS4B proteins have an important structural role in the carboxysome shell. Indeed, special structural roles for CcmL and CsoS4A and an explanation for their low abundance have been suggested by recent crystallographic studies[35] (discussed below). In addition, mRNA transcripts of *csoS4B* (previously *orfB*) and the homologous *csoS4A* (previously *orfA*) were detected in a recent study by Cai and colleagues[47].

The diversity of bacterial microcompartments

For many years, carboxysomes were the only known polyhedral microcompartments in bacteria. However, in 1994, homologues of carboxysome shell proteins were reported in *Salmonella enterica* serovar Typhimurium[48], in which they constitute part of a cluster of genes that are involved in coenzyme B_{12}-dependent metabolism of 1,2-propanediol (FIG. 2b). By encapsulating the enzymes that are necessary for this process within a protein shell, the propanediol utilization microcompartment (encoded by the *pdu* operon) presumably protects the cell from propionaldehyde, a toxic intermediate[49]. Similar microcompartments are also formed in some enteric bacteria

(including *E. coli* and serovars of *S. enterica*) when they are grown in the presence of ethanolamine. The ethanolamine utilization microcompartment (encoded by the *eut* operon) is thought to sequester acetaldehyde, an intermediate in the degradation of ethanolamine, and might serve to either protect cells from the toxic effects of acetaldehyde or help retain this volatile intermediate, thereby preventing the loss of fixed carbon[50–52]. The microcompartments that are formed during growth on 1,2-propanediol or ethanolamine seem to be less uniform in size and more irregular geometrically than carboxysome microcompartments, but it seems likely, based on homology between the components of their shells, that they are constructed through similar architectural principles. The *pdu* microcompartment and its numerous proteins and enzymes have been functionally characterized (reviewed in REFS 8,53) (FIG. 1d; FIG. 2b). Interestingly, the models for operation of the *pdu* microcompartment require the movement of bulky molecules, such as ATP and B_{12} cofactors, across the shell, which raises other questions about molecular transport.

The observation that diverse microcompartments are composed of proteins that are homologous to the main shell proteins of carboxysomes led to the identification of a protein domain in the shells of all polyhedral bacterial microcompartments. This conserved protein domain, Pfam 00936 (REF. 54), is also known as the bacterial microcompartment (BMC) domain. The conserved BMC domain is approximately 84 amino acids long, although it is sometimes found as part of a larger protein or in tandem copies within a single protein. The BMC domain is found in CsoS1A–C and CcmK1–4 (and CcmO) shell proteins from α-carboxysomes and β-carboxysomes and is also present in multiple proteins that are encoded by the *pdu* and *eut* operons (FIG. 2b). A search of the available protein-sequence databases revealed that proteins with the BMC domain are widely distributed, and 189 bacterial species and strains with BMC gene clusters have been identified to date. Representatives are found in the Actinobacteria, Acidobacteria, Firmicutes, Planctomycetes and Fusobacteria, and all groups of the Proteobacteria, except for the epsilon class. The wide distribution of the BMC domain suggests the existence of related, but diverse, bacterial microcompartments, some of which could have novel metabolic functions[8].

Interestingly, sequences that are typical of bacterial microcompartments do not strictly correlate with species. For example, among the five *Rhodopseudomonas palustris* genomes sequenced, only *R. palustris* BisB18 contains a cluster of recognizable microcompartment genes. This cluster encodes a few BMC-domain proteins, together with several enzymes that are involved in carbon metabolism and other proteins of unknown function. Among the *R. palustris* strains for which genome sequences are available, *R. palustris* BisB18 seems to be especially metabolically versatile, and is further distinguished by its ability to grow anaerobically in the dark (C. Harwood, personal communication). Similarly, among the 16 *Vibrio* genomes sequenced, only *Vibrio furnissii* M1 contains a BMC gene cluster[55]; this cluster is composed of a set of genes that encodes BMC-domain proteins that are similar to those in the *pdu* and *eut* operons, together with several other genes, including those that encode pyruvate formate lyase and its activator (FIG. 2b). Based on sequence homology and experimental evidence, *Vibrio furnissii* M1 is postulated to contain a bacterial microcompartment for metabolizing pyruvate to ethanol[55]. The *R. palustris* BisB18 and *V. furnissii* M1 BMC gene clusters are typical of those that have been identified bioinformatically. Proteins that contain a BMC domain seem to be encoded in operons, together with genes that, based on their annotations, code for enzymes that are involved in the metabolism of small organic compounds. All of the gene clusters that have been shown to contain BMC-domain proteins also encode one or more homologues of CcmL, or CsoS4A and CsoS4B (previously known as OrfA and OrfB), which, together with the BMC-domain proteins, are thought to be important structural components of the microcompartment shell (discussed below).

Collectively, these observations suggest that there are many variations on the bacterial microcompartment theme and that different microcompartments are specialized for different metabolic processes. To date, only a few different microcompartments have been directly observed by electron microscopy (EM). Confirming the existence of novel microcompartments experimentally is challenging, owing to the need to culture potentially novel organisms and determine the conditions under which microcompartments are produced. As a result, comparative genomics approaches will continue to be important in obtaining clues about novel microcompartments and their functions. Parallel structural studies could be crucial in verifying such functions and elucidating the mechanisms of molecular transport.

Structural studies on carboxysomes

Structural studies could provide a better understanding of the biochemical mechanisms and evolutionary origins of the carboxysome and related bacterial microcompartments. Two approaches to three-dimensional structural characterization of carboxysomes have been initiated: EM studies have provided a top-down examination of intact carboxysomes that have been isolated from bacterial cells and, in a complementary strategy, X-ray crystallographic studies on individual protein components have provided a bottom-up understanding of the carboxysome.

EM studies of the carboxysome have faced two main challenges. First, isolating intact carboxysomes has generally been difficult. Although cellular fractions that are enriched in carboxysomes have been obtained from several organisms, including a cyanobacterium (*Synechococcus* sp. WH 8102) that produces α-carboxysomes[56], stable, highly purified carboxysomes have only been obtained from *H. neapolitanus*[20] and two *Nitrobacter* strains[57,58] (all of the α-carboxysome type) and high yields have only been obtained from *H. neapolitanus*. Second, heterogeneity (in terms of size and geometric irregularity) makes it difficult to reconstruct high-resolution models from EM images; cryo-EM is

most straightforward when used to image particles that are homogeneous or symmetrical, or both. Two recent EM studies have partially surmounted the need for homogeneity by using electron-cryotomography methods to produce the first three-dimensional reconstructions of carboxysomes from *H. neapolitanus*[56] and *Synechococcus* sp. WH 8102 (REF. 59). Both of these studies revealed that the carboxysome is icosahedral with nearly flat, triangular faces; earlier studies had suggested that carboxysomes were either icosahedral or dodecahedral, and the tendency of carboxysomes to be visualized as hexagons both in thin sections and in two-dimensional projections had added to this uncertainty. How RuBisCO is organized inside the carboxysome is another question that is fundamental to our understanding of the function of carboxysomes. Both of these EM studies showed that RuBisCO molecules are arranged inside the shell in a few concentric layers that begin at the inner surface of the shell and are spaced approximately 110 Å apart[56,59]. Conclusions from the two studies diverged, however, on a few more detailed points, which remain to be clarified. For example, only one study found that the shell is slightly thicker at the vertices of the icosahedron[56]. EM studies have so far been unable to yield information about the structure or organization of the thousands of individual proteins that constitute the outer shell. Progress towards that goal using EM techniques might require carboxysome preparations that are more homogeneous and have greater geometric regularity, which would enable higher-resolution reconstructions. So far, high-resolution details have required an alternative strategy that is based on X-ray crystallography.

In 2005, X-ray crystallographic studies were initiated on individual carboxysome proteins that are present in the shell in the hope of understanding carboxysome assembly. Although these reductionist studies are still ongoing, they have already been fruitful in clarifying the roles of particular proteins and in understanding a number of mechanistic and architectural features of the carboxysome. As discussed above, the CsoS1 and CcmK proteins (also known as BMC-domain proteins) are the main constituents of the carboxysome shell in the α-type and β-type carboxysomes, respectively. The first crystal structures of carboxysome shell proteins were reported by Kerfeld *et al.*[60], who described the structures of CcmK2 and CcmK4 from *Synechocystis* sp. PCC 6803. The crystal structures of the shell proteins CsoS1A from *H. neapolitanus*[61] and CcmK1 from *Synechocystis* sp. PCC 6803 (REF. 35) were subsequently elucidated. All of these shell proteins were shown to form hexamers that are composed of six identical subunits in a cyclic arrangement around a central axis of symmetry (FIG. 3). This established that the shell of the carboxysome is primarily built from hexameric building blocks. However, whether mixed hexamers (for example, of CsoS1A and CsoS1C) occur is not known. Several large, icosahedral viral capsids have also been shown to be assemblies of mainly hexameric building blocks (reviewed in REF. 62), which made it possible to draw architectural parallels between viruses and carboxysomes.

Three of the CcmK and CsoS1 hexamers visualized so far (CcmK1, CcmK2 and CsoS1A) have been observed to associate in crystals, through tight packing of the hexagonal units, to form two-dimensional layers of molecules (FIG. 3). The conserved nature of this association led to the hypothesis that these layers of molecules form the flat facets of the icosahedral carboxysome shell[60,61]. The features of these thin (~20 Å thick), but tightly packed, molecular layers provide insights into the function of the carboxysome. How small molecules, such as bicarbonate, RuBP and 3-PGA, diffuse in and out of the carboxysome is an important question. Structures of the hexagonal layers of the carboxysome revealed small pores that range from approximately 4 to nearly 7 Å in diameter depending on the protein[60,61]. It has been proposed that these pores serve as the channels for substrate and product diffusion across the shell. Among the hexameric structures solved so far, some variation has been observed in the sizes and properties of the pores, but the tendency of the pore to carry a positive electrostatic potential (owing to lysine or arginine residues) has been consistent. This could provide a transport advantage for the negatively charged bicarbonate compared with uncharged CO_2 and O_2; rapid diffusive loss of CO_2 would compromise efficiency, whereas O_2 would compete with CO_2 in its reaction with RuBisCO (BOX 1; FIG. 3b).

An additional study of the hexameric layers (in which the hexamers were spaced 66–70 Å apart from centre to centre) made it possible to establish the number of shell proteins that are present in a carboxysome and estimate triangulation numbers for the carboxysome shell[61], assuming that parallels can be drawn with viral capsids, which are typically constructed according to the principles of quasi-equivalence[63]. A comparatively small carboxysome shell with a diameter of ~850 Å would have ~3,500 shell subunits (with a triangulation number of ~60), whereas a larger carboxysome with a diameter of 1,100 Å would have ~4,500 shell subunits (with a triangulation number of ~75).

The hexagonal layers that are formed by the main shell proteins provide a good model for the flat parts of the shell, but do not explain how a closed shell can be formed. According to the principles of solid geometry, as well as numerous examples that have been provided by viral capsids, pentameric proteins might be present at the vertices of the icosahedral shell. Insertion of pentameric proteins into flat, hexagonal layers that are formed by the hexameric shell proteins would generate the curvature that would be required to close the shell. Recent crystal structures provided by Tanaka and colleagues[35] have shown that the homologous carboxysomal proteins CcmL from *Synechocystis* sp. PCC 6803 and CsoS4A (also known as OrfA) from *H. neapolitanus*, the functions of which were previously unknown, are both pentamers (FIG. 3). The sizes and shapes of these proteins indicate that they constitute the 12 pentameric vertices of the carboxysome shell[35]. The conclusion that these proteins are present only at the vertices, and therefore only 60 protein molecules are present in each shell, explains why they were not detected earlier in carboxysome preparations.

Icosahedron
A regular geometric solid that has 20 triangular faces and 12 vertices; a large icosahedron can be constructed by assembling hexagons together on the triangular faces with pentagons at the 12 vertices.

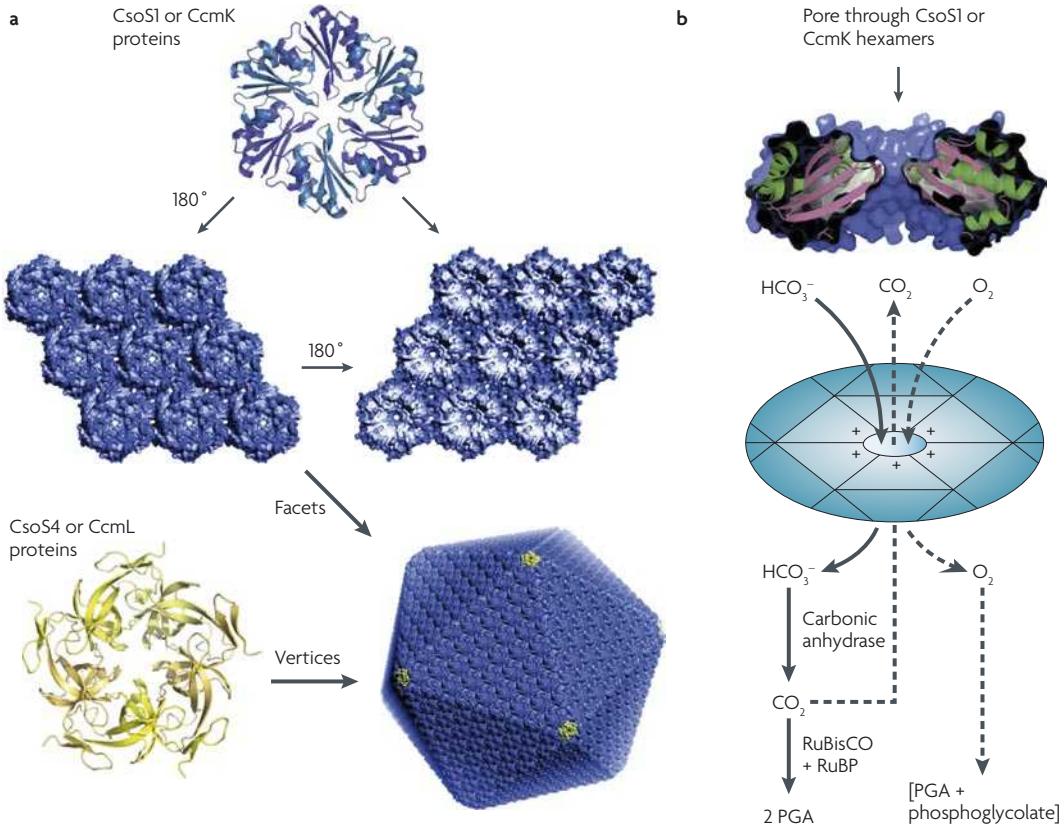


Figure 3 | **High-resolution structures of carboxysome shell proteins. a** | The main constituents of the shell, the hexameric CsoS1 and CcmK proteins, fit together to form a tightly packed molecular layer (top structure; CcmK1 is shown). The two sides of the layer have different geometric and chemical properties, but which side faces outward has not yet been determined. CsoS4A (previously known as OrfA) and CcmL are pentameric carboxysome shell proteins (lower left structure; CcmL is shown) that are proposed to form the vertices of the shell. The lower right structure shows one possible model for the arrangement of hexameric and pentameric proteins in the carboxysome shell. How particular paralogues of these shell proteins and other proteins (such as CsoS2 and CsoSCA in α-carboxysomes and CcmM, CcmN and CcaA in β-carboxysomes) are arranged in the shell is not yet well understood. **b** | Narrow pores through the hexagonal layer are proposed to allow diffusion of substrates and products into and out of the carboxysome. A shell hexamer is shown rotated 90° from that in panel **a** and sectioned through the middle, with the pore (arrow) running vertically along the axis of symmetry. The diagram below the shell hexamer indicates how a positive electrostatic potential is created in the pore by charged side chains and how this might provide an advantage for the passage of negatively charged molecules, such as bicarbonate. Negatively charged ribulose-1,5-bisphosphate (RuBP) and phosphoglycerate (PGA) (not shown) must also cross the shell. Dashed arrows indicate the passage of neutral molecules (CO_2 and O_2), which would be impeded by the tight packing of the shell and would not be promoted by the charged pores. Molecular images were generated using PyMOL[68]. RuBisCO, ribulose-1,5-bisphosphate carboxylase/oxygenase. The lower right structure in panel **a** was adapted, with permission, from REF. 35 © (2008) American Association for the Advancement of Science.

These early structural studies have led to key insights about the architecture and function of carboxysomes, and the detection of hexameric and pentameric building blocks has provided a basic understanding of their underlying self-assembly principles. The tight packing of these units suggests that the shell provides a crucial function in the control of molecular diffusion, whereas detailed features, such as charged pores, have provided clues about selectivity in the diffusion of negatively charged molecules. This selectivity could contribute to the ability of the carboxysome to maintain a high effective CO_2 concentration. However, despite these advances, we are only beginning to understand bacterial microcompartments in atomic level detail.

Future directions

Knowledge of the roles and structures of many of the individual components of the carboxysome has led to a preliminary understanding of how the microcompartment functions, but many important questions remain to be clarified. The distribution of various homologous proteins in the shell is not understood, nor is it certain whether this multiplicity has evolved owing to structural or biochemical factors. Furthermore, the organization of the shell proteins compared with the enzymes that are encapsulated in microcompartments is just beginning to be elucidated[42,43,56,59]. Transport models for the diffusion of molecules across the shell have been surmised from crystal structures of isolated

shell proteins. However, further studies are required to reinforce these proposed models, particularly for more complex microcompartments, which involve larger and more varied molecules. An important piece of information that is missing from our picture of molecular transport is the 'sidedness' of the shell: which side of the molecular layer faces out towards the cytosol and which side faces inward? Some of these questions are ripe for investigation by structure-guided mutagenesis. Finally, it should be noted that even in the well-studied carboxysome, there are some proteins for which their functions and structures have not been elucidated. We therefore acknowledge that the views presented here may need to be modified when our understanding is more complete.

The test of our understanding of microcompartments will be our ability to successfully model their complex behaviour within the cell, either mathematically or computationally, and then validate these models experimentally. This challenge will require integrated knowledge of the self-assembly and spatial organization of carboxysome proteins and their molecular transport and enzymatic activities. Eventually, it might be possible to exploit an understanding of bacterial microcompartments to design new microcompartments with modified properties or novel enzymatic activities, which could result in potentially useful applications in bioengineering and nanotechnology. From such investigations, we would probably learn more about the origin of these exquisite structures and how their varied forms have evolved.

1. Gitai, Z. The new bacterial cell biology: moving parts and subcellular architecture. *Cell* **120**, 577–586 (2005).
2. Shively, J. M. (ed.) *Microbiology Monographs: Complex Intracellular Structures in Prokaryotes* (Springer, Berlin/Heidelberg, 2006).
3. Lowe, J. & Amos, L. A. Crystal structure of the bacterial cell-division protein FtsZ. *Nature* **391**, 203–206 (1998).
4. van den Ent, F., Amos, L. & Lowe, J. Bacterial ancestry of actin and tubulin. *Curr. Opin. Microbiol.* **4**, 634–638 (2001).
5. van den Ent, F., Amos, L. A. & Lowe, J. Prokaryotic origin of the actin cytoskeleton. *Nature* **413**, 39–44 (2001).
6. Cannon, G. C. *et al.* Microcompartments in prokaryotes: carboxysomes and related polyhedra. *Appl. Environ. Microbiol.* **67**, 5351–5361 (2001).
7. Price, G. D., Sültemeyer, D., Klughammer, B., Ludwig, M. & Badger, M. R. The functioning of the CO_2 concentrating mechanism in several cyanobacterial strains: a review of general physiological characteristics, genes, proteins, and recent advances. *Can. J. Bot.* **76**, 973–1002 (1998).
8. Bobik, T. A. Polyhedral organelles compartmenting bacterial metabolic processes. *Appl. Microbiol. Biotechnol.* **70**, 517–525 (2006).
9. Huang, X., Holden, H. M. & Raushel, F. M. Channeling of substrates and intermediates in enzyme-catalyzed reactions. *Annu. Rev. Biochem.* **70**, 149–180 (2001).
10. Drews, G. & Niklowitz, W. Beiträge zur Cytologie der Blaualgen. II. Zentroplasma und granulare Einschlüsse von *Phormidium uncinatum*. *Arch. Mikrobiol.* **24**, 147–162 (1956).
11. Shively, J. M. Inclusion bodies of prokaryotes. *Annu. Rev. Microbiol.* **28**, 167–187 (1974).
12. Shively, J. M., Ball, F., Brown, D. H. & Saunders, R. E. Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of *Thiobacillus neapolitanus*. *Science* **182**, 584–586 (1973).
Reported the first isolation and description of polyhedral bodies and proposed the name carboxysome.
13. Codd, G. A. Carboxysomes and ribulose bisphosphate carboxylase/oxygenase. *Adv. Microbial. Physiol.* **29**, 115–164 (1988).
14. Codd, G. A. & Marsden, W. J. N. The carboxysomes (polyhedral bodies) of autotrophic prokaryotes. *Biol. Rev.* **59**, 389–422 (1984).
15. Shively, J. M. *et al.* Functional inclusions in prokaryotic cells. *Int. Rev. Cytol.* **113**, 35–100 (1988).
16. Heinhorst, S., Cannon, G. C. & Shively, J. M. in *Complex Intracellular Structures in Prokaryotes* (ed. Shively, J. M.) 141–165 (Springer-Verlag, Berlin, 2006).
17. Shively, J. M., Ball, F. L. & Kline, B. W. Electron microscopy of the carboxysomes (polyhedral bodies) of *Thiobacillus neapolitanus*. *J. Bacteriol.* **116**, 1405–1411 (1973).
18. Shively, J. M. & English, R. S. The carboxysome, a prokaryotic organelle: a mini-review. *Can. J. Bot.* **69**, 957–962 (1991).
19. Cannon, G. C., English, R. S. & Shively, J. M. *In situ* assay of ribulose-1,5-bisphosphate carboxylase/oxygenase in *Thiobacillus neapolitanus*. *J. Bacteriol.* **173**, 1565–1568 (1991).
The first direct evidence that carboxysomal RuBisCO exists in a fully activated form *in vitro*, which demonstrated that carboxysomes are metabolically active and are not just storage bodies for the enzyme.
20. Cannon, G. C. & Shively, J. M. Characterization of a homogeneous preparation of carboxysomes from *Thiobacillus neapolitanus*. *Arch. Microbiol.* **134**, 52–59 (1983).
Reported the detailed biochemical characterization of carboxysomes that were purified almost to homogeneity.
21. Buedeker, R. F., Cannon, G. C., Kuenen, J. G. & Shively, J. M. Relations between D-ribulose-1,5 bisphosphate carboxylase, carboxysomes, and CO_2 fixing capacity in the obligate chemolithotroph *Thiobacillus neapolitanus* grown under different limitations in the chemostat. *Arch. Microbiol.* **124**, 185–189 (1980).
22. Turpin, D. H., Miller, A. G. & Canvin, D. T. Carboxysome content of *Synechococcus leopoliensis* (cyanophyta) in response to inorganic carbon. *J. Phycol.* **20**, 249–253 (1984).
23. Price, G. D. & Badger, M. R. Isolation and characterization of high CO_2-requiring-mutants of the cyanobacterium *Synechococcus* PCC7942: two phenotypes that accumulate inorganic carbon but are apparently unable to generate CO_2 within the carboxysome. *Plant Physiol.* **91**, 514–525 (1989).
24. Reinhold, L., Zviman, M. & A, Kaplan, A. quantitative model for inorganic carbon fluxes and photosynthesis in cyanobacteria. *Plant Physiol. Biochem.* **27**, 945–954 (1989).
25. Price, G. D. & Badger, M. R. 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**, 505–513 (1989).
Experimentally established that carbonic anhydrase activity must be localized to the carboxysome rather than to the cytosol for CO_2 fixation.
26. Price, G. D., Howitt, S. M., Harrison, K. & Badger, M. R. Analysis of a genomic DNA region from the cyanobacterium *Synechococcus* sp. strain PCC7942 involved in carboxysome assembly and function. *J. Bacteriol.* **175**, 2871–2879 (1993).
Identified several genes that are required for carboxysome formation in cyanobacteria.
27. Suzuki, E., Fukuzawa, H. & Miyachi, S. Identification of a genomic region that complements a temperature-sensitive, high CO_2-requiring mutant of the cyanobacterium, *Synechococcus* sp. PCC7942. *Mol. Gen. Genet.* **226**, 401–408 (1991).
28. English, R. S., Jin, S. & Shively, J. M. Use of electroporation to generate a *Thiobacillus neapolitanus* carboxysome mutant. *Appl. Environ. Microbiol.* **61**, 3256–3260 (1995).
29. Holthuizen, Y. A., Van Breemen, J. F. L., Kuenen, J. G. & Konigs, W. N. Protein composition of the carboxysomes of *Thiobacillus neapolitanus*. *Arch. Microbiol.* **144**, 398–404 (1986).
30. English, R. S., Lorbach, S. C., Qin, X. & Shively, J. M. Isolation and characterization of a carboxysome shell gene from *Thiobacillus neapolitanus*. *Mol. Microbiol.* **12**, 647–654 (1994).
The first study to correlate a shell protein with its corresponding gene.
31. Cannon, G. C., Heinhorst, S., Bradburne, C. E. & Shively, J. M. Carboxysome genomics: a status report. *Funct. Plant. Biol.* **29**, 175–182 (2002).
Pointed out that *Prochlorococcus* marine cyanobacteria contain carboxysome operons that are related to those of chemolithotrophic bacteria.
32. Shively, J. M., Lorbach, S. C., Jin, S. & Baker, S. H. in *Microbial Growth on C1 Compounds* (eds Lidstrom, M. & Tabita, F. R.) 56–63 (Kluwer, Dordrecht, 1996).
33. Baker, S. H. *et al.* The correlation of the gene *cso*S2 of the carboxysome operon with two polypeptides of the carboxysome in *Thiobacillus neapolitanus*. *Arch. Microbiol.* **172**, 233–239 (1999).
34. Baker, S. H., Williams, D. S., Aldrich, H. C., Gambrell, A. C. & Shively, J. M. Identification and localization of the carboxysome peptide Csos3 and its corresponding gene in *Thiobacillus neapolitanus*. *Arch. Microbiol.* **173**, 278–283 (2000).
35. Tanaka, S. *et al.* Atomic-level models of the bacterial carboxysome shell. *Science* **319**, 1083–1086 (2008).
Showed that certain previously uncharacterized carboxysome proteins assemble as pentamers, postulated that these proteins form the vertices of the carboxysome shell and provided the first atomic level models for the assembled shell.
36. Field, C. B., Behrenfeld, M. J., Randerson, J. T. & Falkowski, P. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**, 237–240 (1998).
37. Badger, M. R., Hanson, D. T. & Price, G. D. Evolution and diversity of CO_2 concentrating mechanisms in cyanobacteria. *Funct. Plant Biol.* **29**, 407–416 (2002).
38. Badger, M. R. & Price, G. D. CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *J. Exp. Bot.* **54**, 609–622 (2003).
39. So, A. K. *et al.* A novel evolutionary lineage of carbonic anhydrase (ε class) is a component of the carboxysome shell. *J. Bacteriol.* **186**, 623–630 (2004).
The first study to show the presence of carbonic anhydrase in carboxysomes as a protein that is tightly associated with, or is part of, the carboxysome shell. Owing to sequence divergence, CsoS3 could not be shown to be homologous to known families of carbonic anhydrases.
40. Sawaya, M. R. *et al.* The structure of β-carbonic anhydrase from the carboxysomal shell reveals a distinct subclass with one active site for the price of two. *J. Biol. Chem.* **281**, 7546–7555 (2006).
The structure reported in this paper revealed that the α-carboxysome carbonic anhydrase (CscoS3 or CsoSCA) is a member of the β-family of carbonic anhydrases, but with the loss of an active site and addition of a domain of unknown function.
41. Price, G. D., Coleman, J. R. & Badger, M. R. Association of carbonic anhydrase activity with carboxysomes isolated from the cyanobacterium *Synechococcus* PCC7942. *Plant Physiol.* **100**, 784–793 (1992).
42. Cot, S. S., So, A. K. & Espie, G. S. A multiprotein bicarbonate dehydration complex essential to carboxysome function in cyanobacteria. *J. Bacteriol.* **190**, 936–945 (2008).
Used various approaches to identify several protein–protein interactions in the β-carboxysome,

identified CcmM as a key organizing component and argued that the CcaA carbonic anhydrase is a component of the *Synechocystis* sp. PCC 6803 carboxysome.
43. Long, B. M., Badger, M. R., Whitney, S. M. & Price, G. D. Analysis of carboxysomes from *Synechococcus* PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. *J. Biol. Chem.* **282**, 29323–29335 (2007).
Identified protein–protein complexes between proteins from the β-carboxysome that involve RuBisCO, CcaA and CcmM.
44. Reinhold, L., Kosloff, R. & Kaplan, A. A model for inorganic carbon fluxes and photosynthesis in cyanobacterial carboxysomes. *Can. J. Bot.* **69**, 984–988 (1991).
45. Heinhorst, S. *et al.* Characterization of the carboxysomal carbonic anhydrase CsoSCA from *Halothiobacillus neapolitanus*. *J. Bacteriol.* **188**, 8087–8094 (2006).
The first study to provide biochemical evidence that the carboxysome shell acts as a barrier for CO_2.
46. Dou, Z. *et al.* 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**, 10377–10384 (2008).
47. Cai, F., Heinhorst, S., Shively, J. M. & Cannon, G. C. Transcript analysis of the *Halothiobacillus neapolitanus cso* operon. *Arch. Microbiol.* **189**, 141–150 (2008).
48. Chen, P., Andersson, D. I. & Roth, J. R. The control region of the *pdu/cob* regulon in *Salmonella typhimurium*. *J. Bacteriol.* **176**, 5474–5482 (1994).
First report of the presence and putative function of a bacterial microcompartment shell protein gene in a heterotrophic organism.
49. Havemann, G. D., Sampson, E. M. & Bobik, T. A. PduA is a shell protein of polyhedral organelles involved in coenzyme B_{12}-dependent degradation of 1,2-propanediol in *Salmonella enterica* serovar Typhimurium LT2. *J. Bacteriol.* **184**, 1253–1261 (2002).
Showed that bacterial microcompartments were formed by *S. typhimurium* in the presence of 1,2-propanediol, which confirmed that bacterial microcompartments are involved in metabolic processes other than CO_2 fixation. PduA was also localized to the shell.
50. Penrod, J. T. & Roth, J. R. Conserving a volatile metabolite: a role for carboxysome-like organelles in *Salmonella enterica*. *J. Bacteriol.* **188**, 2865–2874 (2006).
51. Stojiljkovic, I., Baumler, A. J. & Heffron, F. Ethanolamine utilization in *Salmonella typhimurium*: nucleotide sequence, protein expression, and mutational analysis of the *cchA cchB eutE eutJ eutG eutH* gene cluster. *J. Bacteriol.* **177**, 1357–1366 (1995).
52. Brinsmade, S. R., Paldon, T. & Escalante-Semerena, J. C. Minimal functions and physiological conditions required for growth of *Salmonella enterica* on ethanolamine in the absence of the metabolosome. *J. Bacteriol.* **187**, 8039–8046 (2005).
53. Havemann, G. D. & Bobik, T. A. Protein content of polyhedral organelles involved in coenzyme B_{12}-dependent degradation of 1,2-propanediol in *Salmonella enterica* serovar Typhimurium LT2. *J. Bacteriol.* **185**, 5086–5095 (2003).
Purified and characterized the Pdu microcompartment.
54. Bateman, A. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **32**, D138–D141 (2004).
55. Wackett, L. P., Frias, J. A., Seffernick, J. L., Sukovich, D. J. & Cameron, S. M. Genomic and biochemical studies demonstrating the absence of an alkane-producing phenotype in *Vibrio furnissii* M1. *Appl. Environ. Microbiol.* **73**, 7192–7198 (2007).
56. Schmid, M. F. *et al.* Structure of *Halothiobacillus neapolitanus* carboxysomes by cryo-electron tomography. *J. Mol. Biol.* **364**, 526–535 (2006).
Provided the first three-dimensional reconstructions of an α-carboxysome by cryo-electron tomography.
57. Ebert, A. *Ribulose-1,5-bisphosphate carboxylase in* Nitrobacter. Thesis, Univ. Hamburg (1982).
58. Shively, J. M., Bock, E., Westphal, K. & Cannon, G. C. Icosahedral inclusions (carboxysomes) of *Nitrobacter agilis*. *J. Bacteriol.* **132**, 673–675 (1977).
59. Iancu, C. V. *et al.* The structure of isolated *Synechococcus* strain WH8102 carboxysomes as revealed by electron cryotomography. *J. Mol. Biol.* **372**, 764–773 (2007).
Provided three-dimensional reconstructions of α-carboxysomes by cryo-electron tomography, together with models for the packing of RuBisCO enzymes inside the carboxysome.
60. Kerfeld, C. A. *et al.* Protein structures forming the shell of primitive bacterial organelles. *Science* **309**, 936–938 (2005).
Provided the first atomic resolution crystal structures of the proteins that constitute the bulk of the carboxysome shell, which revealed hexameric shell protein assemblies that were arranged in tightly packed two-dimensional layers, with small pores that were suggestive of molecular transport.
61. Tsai, Y. *et al.* Structural analysis of CsoS1A and the protein shell of the *Halothiobacillus neapolitanus* carboxysome. *PLoS Biol.* **5**, 1345–1354 (2007).
Provided the first crystal structure of a shell protein from an α-carboxysome, together with a comparison of the structure and packing of shell proteins from the previously elucidated β-carboxysome.
62. Johnson, J. E. & Speir, J. A. Quasi-equivalent viruses: a paradigm for protein assemblies. *J. Mol. Biol.* **269**, 665–675 (1997).
63. Caspar, D. L. & Klug, A. Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* **27**, 1–24 (1962).
64. Eisenhut, M. *et al.* The plant-like C2 glycolate cycle and the bacterial-like glycerate pathway cooperate in phosphoglycolate metabolism in cyanobacteria. *Plant Physiol.* **142**, 333–342 (2006).
65. Eisenhut, M. *et al.* Long-term response toward inorganic carbon limitation in wild type and glycolate turnover mutants of the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Plant Physiol.* **144**, 1946–1959 (2007).
66. Liu, Y. *et al.* PduL is an evolutionarily distinct phosphotransacylase involved in B_{12}-dependent 1,2-propanediol degradation by *Salmonella enterica* serovar Typhimurium LT2. *J. Bacteriol.* **189**, 1589–1596 (2007).
67. Kofoid, E., Rappleye, C., Stojiljkovic, I. & Roth, J. The 17-gene ethanolamine (*eut*) operon of *Salmonella typhimurium* encodes five homologues of carboxysome shell proteins. *J. Bacteriol.* **181**, 5317–5329 (1999).
68. DeLano, W. L. *The PyMOL User's Manual* (DeLano Scientific, San Carlos, California, 2002).
69. So, A. K. & Espie, G. S. Cloning, characterization and expression of carbonic anhydrase from the cyanobacterium *Synechocystis* PCC6803. *Plant Mol. Biol.* **37**, 202–215 (1998).

Acknowledgements
The authors thank S. Tanaka for figure preparation and T. Bobik for discussions. T.O.Y. and C.A.K. have been supported by grants from the BER programme of the Department of Energy Office of Science and the US Department of Agriculture, G.C.C. and S.H. have been supported by the National Science Foundation (grant numbers MCB-0444568 and DMR-0213883, repectively) and G.C.C. and S.H. have been supported by the T.W. Bennett Foundation at The University of Southern Mississippi.

DATABASES

Entrez Genome Project: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj
Escherichia coli | *Nitrobacter winogradskyi* | *Nitrococcus mobilis* | *Nostoc punctiforme* | *Prochlorococcus marinus* str. MIT 9313 | *Prochlorococcus marinus* subsp. *pastoris* str. CCMP1986 | *Rhodopseudomonas palustris* BisB18 | *Salmonella enterica* serovar Typhimurium | *Synechococcus elongatus* PCC 7942 | *Synechocystis* sp. PCC 6803 | *Synechococcus* sp. WH 8102

FURTHER INFORMATION

Todd O. Yeates's homepage: http://www.doe-mbi.ucla.edu/~yeates/

ALL LINKS ARE ACTIVE IN THE ONLINE PDF