Elucidating Essential Role of Conserved Carboxysomal Protein CcmN Reveals Common Feature of Bacterial Microcompartment Assembly

Received for publication, February 22, 2012, and in revised form, March 20, 2012. Published, JBC Papers in Press, March 29, 2012, DOI 10.1074/jbc.M111.355305

James N. Kinney†§, Annette Salmeen†§, Fei Cai†§, and Cheryl A. Kerfeld†§¶

From the †Department of Energy Joint Genome Institute, Walnut Creek, California 94598 and the ‡Physical Biosciences Division, Lawrence Berkeley National Laboratory and the §Department of Plant and Microbial Biology, University of California, Berkeley, California 94720

Background: CcmN is a conserved carboxysomal protein of unknown function in β-cyanobacteria. Results: The N-terminal domain of CcmN binds the encapsulated protein CcmM; the C-terminal peptide of CcmN interacts with the major carboxysomal shell protein. Conclusion: CcmN is an essential carboxysomal component; its C-terminal peptide is essential for carboxysome biogenesis. Significance: Diverse bacterial microcompartment-encapsulated proteins contain a similar peptide, suggesting that it is a common feature of bacterial microcompartment biogenesis.

Bacterial microcompartments are organelles composed of a protein shell that surrounds functionally related proteins. Bioinformatic analysis of sequenced genomes indicates that homologs to shell protein genes are widespread among bacteria and suggests that the shell proteins are capable of encapsulating diverse enzymes. The carboxysome is a bacterial microcompartment that enhances CO2 fixation in cyanobacteria and some chemoautotrophs by sequestering ribulose-1,5-bisphosphate carboxylase/oxygenase and carbonic anhydrase in the microcompartment shell. Here, we report the in vitro and in vivo characterization of CcmN, a protein of previously unknown function that is absolutely conserved in some shell. Deletion of the peptide abolishes carboxysome formation, indicating that its interaction with the shell is an essential step in microcompartment formation. Peptides with similar length and sequence properties to those in CcmN can be bioinformatically detected in a large number of diverse proteins proposed to be encapsulated in functionally distinct microcompartments, suggesting that this peptide and its interaction with its cognate shell proteins are common features of microcompartment assembly.

Bacterial microcompartments (BMCs)2 are organelles consisting of functionally related proteins surrounded by a protein shell (1–4). Models from crystallographic studies have indicated that BMC shells are composed of proteins that form (pseudo) hexamers that tile into layers to form the facets and pentameric proteins that form the vertices of an apparently icosahedral shell (5). Genes encoding shell proteins are found in >20% of sequenced bacterial genomes, where they appear to encapsulate a variety of enzymes, many of which are oxygen-sensitive and/or catalyze reactions that produce toxic or volatile intermediates (2, 4, 6). Because of their structural and functional properties, BMCs have considerable potential for applications in bioengineering and synthetic biology.

Carboxysomes were the first BMCs to be discovered (7) and characterized (8) by transmission electron microscopy (TEM) of cyanobacteria (Fig. 1A, inset). They contain ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) (9) and carbonic anhydrase. The semipermeable protein shell of the carboxysome allows the diffusion of bicarbonate into the interior, where it is converted to CO2, a substrate for Rubisco, by carbonic anhydrase. The shell also functions as a selective barrier to retain CO2 in the interior (10). It also presumably excludes O2, an alternative substrate for Rubisco, thereby enhancing CO2 fixation. Given that all cyanobacteria contain carboxysomes and that cyanobacteria are considered the most abundant photosynthetic organisms on Earth (11), a substantial portion of global CO2 fixation takes place in carboxysomes.

There are two types of carboxysomes, distinguished by the form of Rubisco they encapsulate and by their gene organization. α-Carboxysomes contain Form IA Rubisco and are found in picocyanobacteria (also known as α-cyanobacteria) and some chemoautotrophs; β-carboxysomes contain Form IB Rubisco and are found in all other cyanobacteria (β-cyanobac-

---

1 To whom correspondence should be addressed: Dept. of Energy Joint Genome Inst., 2800 Mitchell Dr., Walnut Creek, CA 94598. Tel.: 925-296-5691; E-mail: ckerfeld@lbl.gov.

2 The abbreviations used are: BMC, bacterial microcompartment; TEM, transmission electron microscopy; Rubisco, ribulose-1,5-bisphosphate carboxylase/oxygenase; Syn7942, S. elongatus PCC 7942; HCR, high CO2-requiring; CFP, cyan fluorescent protein; HMM, hidden Markov model.
teria). α- and β-carboxysomes also differ in the type of carbonic anhydrase (CsoSCA or CcmM [UniProt Q03513]/CcaA, respectively) that is encapsulated. Both types of carboxysomal gene clusters contain an additional gene encoding a protein of unknown function: CsoS2 in α and CcmN (UniProt P46204) in β. CsoS2 and CcmN do not share any sequence homology.

Among cyanobacterial carboxysomes, the β-carboxysome is the most extensively characterized. β-Carboxysomes are coded for by one of the few gene clusters conserved in β-cyanobacteria (12), with additional carboxysomal genes dispersed throughout the genome (Fig. 2A). Crystal structures of the shell proteins have led to a proposed model for the carboxysome shell (5, 13). In addition, the crystal structures of the β-carboxysomal Rubisco (14) and the γ-carbonic anhydrase domain of CcmM (15) are known, but structures of complexes of the encapsulated proteins have not been determined. Some of the protein-protein interactions within the β-carboxysome have been deduced: CcmM has been shown to interact with the carboxysomal carbonic anhydrase (CcaA), Rubisco, and CcmN via yeast two-hybrid analysis (16); and CcmN was also identified as an interaction partner with CcmM through in vitro pulldown assays using recombinant protein (17).

Despite being conserved in genetic location and sequence, little is known about the function of CcmN. Here, we report the in vivo and in vitro characterization of CcmN, demonstrating that it localizes to the carboxysome and is essential for carboxysome biogenesis and function. We show that CcmN consists of two functional regions. The N terminus forms a stable complex with CcmM that interacts with Rubisco. A short C-terminal peptide interacts with the essential carboxysomal shell protein CcmK2 (UniProt Q03511) and is essential for carboxysome biogenesis. Shell proteins are the defining feature of all BMCs; we show that peptides similar to the one found in CcmN are present in a large number of proteins known or presumed to be encapsulated in diverse BMCs. Among these is a peptide found in the N terminus of an enzyme of the pdu (1,2-propanediol utilization) microcompartment of Salmonella typhimurium LT2; it was sufficient to target a fluorescent protein to the pdu microcompartment, but the nature of the protein-protein interactions underlying this targeting was not established (18). The peptides we identified are found at the N or C terminus or between domains of many BMC-encapsulated proteins, suggesting that they are a widespread structural feature of protein interaction with BMC shells.

**EXPERIMENTAL PROCEDURES**

**Bacterial Strain Construction, Media, and Growth Conditions—**Synechococcus elongatus PCC 7942 (Syn7942) was purchased from the Pasteur Culture Collection and cultured in BG-11 liquid or BG-11 agar (19) at 30 °C under a constant illumination of ~100 micro-einsteins/m²/s in ambient or 3% CO₂ for high CO₂-requiring (HCR) mutants. Syn7942 mutants were generated via homologous recombination as described previously (20, 21). All fluorescent constructs were under the control of the native allophycocyanin promoter (papcA). Constructs and primers are listed in supplemental Tables S1 and S2, respectively. Fully segregated transformants were verified by PCR.

**TEM of Syn7942—**Cyanobacterial cells were fixed using 2% EM-grade glutaraldehyde and 0.2 M sodium cacodylate (pH 7.2). Fixed cells were embedded in Spurr’s resin and sectioned according to standard methods (22). Thin sections were stained
with 2% methanolic uranyl acetate and Reynolds lead citrate and imaging on a JEOL 1200-EX microscope at 80 kV.

Fluorescence Microscopy—Log-phase cells with YFP or cyan fluorescent protein (CFP) fluorophore were spotted on a thin BG-11 agar pad for imaging with a Zeiss Axioplan 2 microscope using a 100× oil immersion objective. Images were visualized and analyzed with NIH ImageJ; the JACoP plug-in was used for co-localization analyses (24).

In Vivo Immunoprecipitation of YFP-tagged Proteins—7.5 μg of goat anti-GFP polyclonal antibody (Abcam) in 200 μl of PBS and 0.02% Tween 20 (pH 7.4) was added to 50 μl of protein G-conjugated Dynabeads (Invitrogen) and incubated at 4 °C overnight. Washing and elution were carried out following the immunoprecipitation protocol using Bio-Rad Micro Bio-Spin chromatography columns to effectively remove the lysate and washes from the resin.

Western Blotting—Gels were transferred to a PVDF membrane at 70 V for ~1–2 h at 4 °C. Primary rabbit anti-CcmK2 antiserum was diluted 1:2500, and secondary HRP-conjugated goat anti-rabbit IgG (Santa Cruz Biotechnology) was diluted 1:20,000 in TBS with 0.01% Triton X-100. The blot was developed using SuperSignal West Pico substrate (Pierce) and imaged using a Bio-Rad ChemiDoc XRS+ system.

Sequence Analysis—CcmN amino acid sequences from cyanobacteria were obtained from public and private databases in Integrated Microbial Genomes. Alignments were made using MUSCLE (25) and viewed with Jalview (26). Jpred (27) was used for secondary structure predictions. A hidden Markov model (HMM) and HMM logos were created from the alignment of all CcmN orthologs using HMMER Version 3.0 and LogoMatP (28), respectively. A visual search for BMC gene clusters was done by searching Integrated Microbial Genomes for Pfam00936 domains in bacterial genomes and viewing each hit and its gene neighborhood. Putative encapsulated metabolic pathways were inferred based on the annotation of the enzymes in the BMC gene cluster and pathways checked using the KEGG database. Helical wheel projections were created using a web-based server from the Zidovetski lab at University of California, Riverside.

Construction of CcmN Homology Model—SWISS-MODEL (29) was used in automated mode using the CcmN protein sequence from Syn7942 as the input and the CcmM structure (Protein Data Bank code 3KWD, Chain A) from Thermosynechococcus elongatus as the template.
Role of CcmN in the Carboxysome

RESULTS

CcmN Localizes to the Carboxysome—Although the gene for CcmN is conserved in all β-cyanobacterial carboxysomal gene clusters, its localization to the carboxysome had not been demonstrated. To confirm CcmN localization to the carboxysome, we cotransformed wild-type Syn7942 with YFP fused to CcmN and CFP fused to the large subunit of Rubisco (RbcL) (UniProt Q31NB3) (30). When each fluorophore was imaged individually, fluorescent puncta could be seen in the cells. When these images were overlaid, the fluorescence intensity of the spots aligned with thresholded Mander’s coefficients of 0.422 and 0.407 for the CFP and YFP channels, respectively (Fig. 3, A–C). These values are similar to co-localization statistics of the shell protein CcmK4 fused to YFP with an RbcL-CFP fusion (supplemental Fig. S1).

CcmN Is Essential for Carboxysome Formation—To determine whether CcmN is an essential carboxysomal protein, a knock-out mutant (7942ΔN) was created in Syn7942. The mutant exhibited an HCR phenotype (supplemental Fig. S2), unable to grow at ambient CO₂ levels but with WT-like growth rates at 3% CO₂, indicative of impaired carboxysome function. Reintroduction of CcmN restored the ability of the mutant to grow at ambient CO₂ (supplemental Fig. S2).

TEM revealed that the 7942ΔN strain did not form carboxysomes but instead contained large, mostly polar aggregates of protein (Fig. 1A). Introduction of YFP-CcmN into the 7942ΔN mutant resulted in fully formed carboxysomes (Fig. 1B). We confirmed the results of the TEM studies with fluorescence microscopy. Expression of YFP-CcmN in the 7942ΔN mutant led to an even distribution of fluorescent puncta (Fig. 3D), indicating restoration of proper carboxysome formation and localization. In contrast, expression of RbcL-CFP in the 7942ΔN mutant did not restore proper carboxysome formation. RbcL-CFP resulted in single fluorescent spots localized at the poles (Fig. 3E), which are consistent with the polar masses observed by TEM. RbcL-CFP expressed in WT cells resulted in evenly distributed fluorescent puncta in each cell (Fig. 3B). Expression of RbcL-CFP in 7942ΔN and WT cells did not alter the background strains’ HCR and WT growth phenotypes, respectively (supplemental Fig. S2).

CcmN Interacts with CcmM through Its N-terminal Region—An HMM logo of all CcmN orthologs (Fig. 2B) revealed that CcmN can be subdivided into two conserved regions: the first, a conserved N-terminal region of ~120 amino acids that contains six bacterial hexapeptide repeat domains (Pfam00132). Hexapeptide repeat domains are also found in the N-terminal domain of CcmM, facilitating construction of a model for the N-terminal region of CcmN (Fig. 2C). The N-terminal domain of CcmN is followed by a variable length linker region enriched in Pro and Ser residues and a C-terminal peptide of 18 conserved amino acids (Fig. 2B). In some CcmN orthologs, the peptide is followed by another poorly conserved region of up to 24 residues.

To correlate function with the different regions of the CcmN primary structure, we coexpressed cleavable GST-CcmN, GST-CcmNΔ18 (lacking the conserved 18 amino acids on the C terminus), and GST-CcmNΔ39 (lacking both the variable region and conserved peptide) fusions (Fig. 2B) with CcmM in E. coli and were able to co-purify stable CcmM-CcmN complexes after on-column cleavage of the GST tag (Fig. 4 and supplemental Fig. S3). These data indicate that CcmN interacts with CcmM through the N-terminal region of CcmN.

Conserved C-terminal Peptide of CcmN Is Crucial for Proper Carboxysome Assembly—To determine the function of the C-terminal peptide of CcmN, we constructed a mutant lacking 18 amino acids at the C terminus of CcmN in Syn7942 (7942ΔN18). The 7942ΔN18 mutant exhibited an HCR phenotype (supplemental Fig. S2). TEM of the 7942ΔN18 mutant revealed that the cells lacked carboxysomes; instead, they contained polar masses of protein larger than and without the distinct shape of carboxysomes (Fig. 5A). When we transformed this strain with YFP-CcmNΔ18, fluorescent puncta were present at the poles, consistent with the TEM images (Fig. 5B).
Wild-type growth at ambient CO₂ levels and proper carboxysome formation were restored by reintroduction of the full-length \textit{ccmN} gene in the 7942N/H9004 strain (Fig. 5, C and D). Collectively, these data indicate that the C-terminal peptide of CcmN is essential for carboxysome formation.

**Conserved C-terminal Region of CcmN Binds Major Shell Component CcmK2 in Vivo and in Vitro**—To identify protein interaction partners specific to the N-terminal region and C-terminal peptide of CcmN, the WT and 7942N/H9004 mutant strains were transformed with YFP-CcmN, YFP-CcmN/H900418, or YFP-CcmN/H900439, and their lysates were screened for interacting proteins using anti-GFP antibody to precipitate the YFP-tagged proteins. CcmM, RbcL, and RbcS (UniProt Q31NB2) were precipitated by both full length and truncated CcmN fusion proteins (Fig. 6A), as identified by Western blotting, molecular weight on SDS-PAGE, and/or mass spectrometry. Subsequent Western blotting with anti-CcmK2 antibody showed that a shell protein was present only in lanes that contained full-length CcmN (Fig. 6A), indicating that the C-terminal peptide interacted with shell proteins. The amount of CcmK2 pulled down by full-length CcmN appeared to be greater in the WT experiment than in the mutant background.

To verify that the peptide interacts with the essential carboxysomal shell protein CcmK2, GST-CcmN and GST-CcmNΔ39, and their lysates were screened for interacting proteins using anti-GFP antibody to precipitate the YFP-tagged proteins. CcmM, RbcL, and RbcS (UniProt Q31NB2) were precipitated by both full length and truncated CcmN fusion proteins (Fig. 6A), as identified by Western blotting, molecular weight on SDS-PAGE, and/or mass spectrometry. Subsequent Western blotting with anti-CcmK2 antibody showed that a shell protein was present only in lanes that contained full-length CcmN (Fig. 6A), indicating that the C-terminal peptide interacted with shell proteins. The amount of CcmK2 pulled down by full-length CcmN appeared to be greater in the WT experiment than in the mutant background.

To verify that the peptide interacts with the essential carboxysomal shell protein CcmK2, GST-CcmN and GST-CcmNΔ39 were each coexpressed with CcmK2 in \textit{E. coli}. Only full-length CcmN was able to pull down CcmK2 in vitro (Fig. 6B).

**Bioinformatic Identification of Conserved Peptides in Encapsulated Proteins of Diverse BMCs**—CcmK2 contains the BMC domain (Pfam00936), the signature domain found in the most abundant protein of BMC shells. Gene clusters that have the potential to encode proteins that form BMCs can be identified bioinformatically by searching for conserved groups of genes.
Role of CcmN in the Carboxysome

![Diagram A](Image)

![Diagram B](Image)

**FIGURE 7.** Conserved properties of peptide in various encapsulated proteins of known and bioinformatically identified BMCs. **A,** an overview sequence alignment of 303 BMC-associated peptides (supplemental Table S4) shows the conservation by property of the peptides. Blue, hydrophobic; green, polar uncharged; red, charged; orange, Pro and Gly. The consensus helical prediction is underscored with a yellow bar. **B,** helical wheel representation of the consensus predicted helix. The corresponding amino acid residues from the peptides of CcmN (inner), the N terminus of PduP of *S. typhimurium* LT2 (*, center), and the interdomain of the B₁₂-independent diol dehydratase of *R. palustris* Bis818 (†, outer) are mapped on the helix.

that contain orthologs to Pfam00936 and the putative vertex proteins (Pfam03319) (2). The function of a BMC that is identified bioinformatically may be inferred from the annotations of other genes in the cluster. We classified BMC gene clusters from currently available genome sequence data into 10 different functional categories (supplemental Table S3). Given the homology among shell proteins and the demonstration that the CcmN peptide interacts with the CcmK2 shell protein, we searched for homologs to the peptide in other gene products in BMC gene clusters. In all 10 types, one or more genes encoded proteins that contained a region similar to the CcmN peptide; these were found at either the N or C terminus or between domains of a presumably encapsulated protein. All were short (13–22 amino acids), and a portion of the peptide was predicted to be α-helical with high confidence in most (Fig. 7A and supplemental Table S3). Notably, the α-helix had a conserved face formed by four hydrophobic amino acids flanked by two polar/charged amino acids (Fig. 7B). As in CcmN, a poorly conserved region rich in Pro and Ser separated the peptides from the functional domains of the proteins.

**DISCUSSION**

We have shown that a previously uncharacterized protein, CcmN, is indispensable for β-carboxysome biogenesis. Our data show that CcmN interacts directly with the carboxysome-encapsulated protein CcmM and the carboxysomal shell protein CcmK2 through two distinct regions. The N-terminal domain of CcmN interacts with CcmM (Fig. 4 and supplemental Fig. S3), which binds to other encapsulated carboxysomal proteins, Rubisco and the carboxysomal carbonic anhydrase CcaA (16, 17). In the pulldown experiments (Fig. 6A), bands attributed to Rbcl, RbcS, and CcmM were evident, whereas an expected band for CcaA at 30 kDa was not observed.

A yeast two-hybrid study reported an interaction between the N-terminal domain of CcmM and CcmN (17). A comparison of their (proposed) structures (Fig. 2C) readily suggests hypotheses for how CcmN and the N-terminal domain of CcmM interact within the carboxysome. Both proteins contain hexapeptide repeats in their N-terminal domains, which are known to adopt a left-handed β-helical fold (31). The N-terminal domain of CcmM, a member of the γ-carbonic anhydrase family, crystallized as a trimer (15), similar to other hexapeptide repeat-containing proteins (32). A structural homology model of CcmN can be constructed using the structure of the N-terminal domain of CcmM (Fig. 2C) (15). Given their structural similarity, CcmN can be envisioned as interacting with the CcmM trimer or CcmM replacing CcmM monomers to form a heterotrimer. Also of note is the typically observed lack of carboxy

anhydride activity in purified CcmM (15); this could be due to the necessity of the interaction with the CcmN N-terminal domain to achieve activity.

The conserved C-terminal peptide of CcmN is essential for carboxysome formation; deletion of the peptide results in an HCR phenotype (supplemental Fig. S2), absence of carboxysomes, and, in most cells, a single, shapeless, polar protein aggregate (Fig. 5, A and B). The importance of the CcmN peptide for carboxysome assembly is also supported by an early study reporting the characterization of a spontaneous mutant of Syn7942; a point mutation of the conserved glycine (G146D) (Fig. 2B) found in the peptide region caused an HCR phenotype and resulted in carboxysomes with poorly defined edges when visualized by TEM (33). We attempted to show that the CcmN C-terminal peptide (both with and without the CcmN linker region) could target YFP to the carboxysome, but invariably, the constructs did not express well, resulting in formation of fluorescent puncta in only a small percentage of cells.

Based on our interaction data (Fig. 6) from both Syn7942 and *E. coli*, the C-terminal peptide region in CcmN interacts with the essential carboxysomal shell protein CcmK2. The decrease in detectable CcmK2 in the mutant background pulldown assay compared with the WT (Fig. 6A) could be due to lower expression of the carboxysomal gene cluster caused by the mutagenesis or an abundance of the truncated CcmN complex interfering with the proper assembly of carboxysomal components.

All BMC shell proteins have recognizable sequence homology, and many proteins identified experimentally or bioinformatically as encapsulated in BMCs contain peptides with sequence properties similar to those of the peptide in CcmN.
(Fig. 7A). Among 10 types of BMCs (supplemental Table S3), one or more of the putatively encapsulated proteins in each contains a peptide similar to that in CcmN. The peptides range from 13 to 22 residues in length with a predicted propensity for α-helix formation, most with high confidence. The predicted helix has a conserved hydrophobic face formed by four residues, whereas the rest of the helix is mainly polar or charged (Fig. 7B). The peptide is always separated from the rest of the protein by a poorly conserved region of variable length that is predicted to be unstructured. The presence of this structurally disordered region could be necessary to reduce any steric hindrance between the region of the protein that interacts with shell and the functional domain of the protein, which presumably interacts with other encapsulated proteins.

Notably, one of the peptides containing these features is the N terminus of PduP, a propionaldehyde dehydrogenase. The PduP peptide was shown to target GFP to the pdu BMC (18) after it was noted to be present in sequences of BMC-encapsulated propionaldehyde dehydrogenases and absent in sequences of cytosolic forms of the enzyme. Deletion of PduP still resulted in BMCs that lacked only the dehydrogenase function, suggesting that, in contrast to the CcmN peptide, the PduP peptide is not essential for pdu BMC formation. Unlike the carboxysome, in which only one component protein contains an identifiable peptide, the pdu BMC gene cluster encodes two additional proteins containing similar peptides that may compensate for the loss of the PduP peptide.

In addition, peptides identified by sequence features similar to those of the CcmN peptide are also found at the C terminus (in an aldolase of Planctomyces limnophilus) or between domains (in a B12-independent diol dehydratase of Rhodopseudomonas palustris BisB18) of BMC-encapsulated proteins. Interestingly, the proteins containing the peptide domain tend to be associated with enzymes for the first or last step of the hypothesized encapsulated series of reactions, consistent with positioning these enzymes adjacent to the BMC shell, which is assumed to play a key role in substrate and product flux.

The presence of this peptide in many (putatively) encapsulated proteins suggests that it interacts with conserved sequence features in all shell proteins. The most conserved amino acids among BMC shell proteins are found at the edges of hexamers (23), where they are presumably important for interaction with adjacent hexamers to form the facets of the shell. Although speculative, this leads to a model for BMC assembly in which proteins that contain the peptide interact with the shell proteins as they assemble. In the carboxysome, for example, as CcmK2 hexamers form facets, binding sites for the CcmN peptide may be created at the 2- or 3-fold axes between adjacent hexamers. Other proteins to be encapsulated, such as CcmM, can then be recruited to the assembling shell via interaction with the N-terminal domain of CcmN. Alternatively, shell assembly may commence after recruitment of shell proteins to peptide motifs displayed by assemblies of interior components.

CcmN plays an essential role in the formation and organization of a bacterial organelle that plays a key role in global CO2 fixation. The occurrence of gene products containing similar peptides in diverse BMC gene clusters suggests that such shell-protein-peptide interactions are a common feature of BMC organization. This feature potentially offers a versatile approach to adhering molecules to BMC shell proteins for applications such as assembling enzyme-scaffold complexes or designed enzymatic nanoreactors.

Acknowledgments—We thank Dr. Susan Golden for providing constructs used in Syn7942 and imaging guidance, Dr. David Savage for fluorescently tagged CcmK4 and RbcL, Dr. Gustaf Sandh for the creation of the pAcmCcm and pCcmNA1.8 constructs, the UC Berkeley Electron Microscopy Lab for assistance with TEM, and Seth Axen for helpful discussions and figure preparation.

REFERENCES
1. Yeates, T. O., Kerfeld, C. A., Heinhorst, S., Cannon, G. C., and Shively, J. M. (2008) Protein-based organelles in bacteria: carboxysomes and related microcompartments. Nat. Rev. Microbiol. 6, 681–691.
2. Kerfeld, C. A., Heinhorst, S., and Cannon, G. C. (2010) Bacterial microcompartments. Ann. Rev. Microbiol. 64, 391–408.
3. Cannon, G. C., Bradburne, C. E., Aldrich, H. C., Baker, S. H., Heinhorst, S., and Shively, I. M. (2001) Microcompartments in prokaryotes: carboxysomes and related polyhedra. Appl. Environ. Microbiol. 67, 5351–5361.
4. Bobik, T. A. (2006) Polyhedral organelles compartmenting bacterial metabolic processes. Appl. Microbiol. Biotechnol. 70, 517–525.
5. Tanaka, S., Kerfeld, C. A., Sawaya, M. R., Cai, F., Heinhorst, S., Cannon, G. C., and Yeates, T. O. (2008) Atomic level models of the bacterial carboxysome shell. Science 319, 1083–1086.
6. Penrod, J. T., and Roth, J. R. (2006) Conserving a volatile metabolite: a role for carboxysome-like organelles in Salmonella enterica. J. Bacteriol. 188, 2865–2874.
7. Drews, G., and Niklowitz, W. (1956) Beiträge zur cytologie der blaualgen. Arch. Microbiol. 24, 147–162.
8. Gantt, E., and Conti, S. F. (1969) Ultrastructure of blue-green algae. J. Bacteriol. 97, 1486–1493.
9. Shively, J. M., Ball, F., Brown, D. H., and Saunders, R. E. (1973) Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of Thiothrix近距离. Acta Crystallogr. D Biol. Crystallogr. 30, 936–938.
10. Cai, F., Menon, B. B., Cannon, G. C., Curry, K. J., Shively, J. M., and Heinhorst, S. (2009) The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO2 leakage barrier. PLoS ONE 4, e7521.
11. Partensky, F., Hess, W. R., and Vaulot, D. (1999) Prochlorococcus, a marine photosynthetic prokaryote of global significance. Microbiol. Mol. Biol. Rev. 63, 106–127.
12. Frangeul, L., Quillardet, P., Castets, A. M., Humbert, J. F., Matthijis, H. C., Cortez, D., Tolonen, A., Zhang, C. C., Gribaldo, S., Kehr, J. C., Zillig, E., Ziemert, N., Becker, S., Talla, E., Latifi, A., Billaut, A., Lepelletier, A., Dittmann, E., Bouchier, C., and de Marsac, N. T. (2008) Highly plastic genome of Microcystis aeruginosa PCC 7806, a ubiquitous toxic freshwater cyanobacterium. BMC Genomics 9, 274.
13. Kerfeld, C. A., Sawaya, M. R., Tanaka, S., Nguyen, C. V., Phillips, M., Beeby, M., and Yeates, T. O. (2005) Protein structures forming the shell of primitive bacterial organelles. Science 309, 936–938.
14. Newman, J., Branden, C. I., and Jones, T. A. (1993) Structure determination and refinement of ribulose-1,5-bisphosphate carboxylase/oxygenase from Synechococcus PCC 6301. Acta Crystallogr. D Biol. Crystallogr. 49, 548–560.
15. Peña, K. L., Castel, S. E., de Araujo, C., Espie, G. S., and Kimber, M. S. (2010) Structural basis of the oxidative activation of the carboxysomal γ-carbonic anhydrase, CcmM. Proc. Natl. Acad. Sci. U.S.A. 107, 2455–2460.
16. Long, B. M., Badger, M. R., Whitney, S. M., and Price, G. D. (2007) Analysis of carboxysomes from Synechococcus PCC 7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. J. Biol. Chem. 282, 29323–29335.
17. Cot, S. S., So, A. K., and Espie, G. S. (2008) A multiprotein bicarbonate dehydration complex essential to carboxysome function in cyanobacteria. *J. Bacteriol.* **190**, 936–945
18. Fan, C., Cheng, S., Liu, Y., Escobar, C. M., Crowley, C. S., Jefferison, R. E., Yeates, T. O., and Bobik, T. A. (2010) Short N-terminal sequences package proteins into bacterial microcompartments. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 7509–7514
19. Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M. and Stanier, R. Y. (1979) Generic assignments, strain histories, and properties of pure cultures of cyanobacteria. *Microbiology* **111**, 1–61
20. Andersson, C. R., Tsinoremas, N. F., Shleton, J., Lebedeva, N. V., Yarrow, J., Min, H., and Golden, S. S. (2000) Application of bioluminescence to the study of circadian rhythms in cyanobacteria. *Methods Enzymol.* **305**, 527–542
21. Clerico, E. M., Ditty, J. L., and Golden, S. S. (2007) Specialized techniques for site-directed mutagenesis in cyanobacteria. *Methods Mol. Biol.* **362**, 155–171
22. Byers, B., and Goetsch, L. (1991) Preparation of yeast cells for thin-section electron microscopy. *Methods Enzymol.* **194**, 602–608
23. Yeates, T. O., Crowley, C. S., and Tanaka, S. (2010) Bacterial microcompartment organelles: protein shell structure and evolution. *Annu. Rev. Biophys.* **39**, 185–205
24. Bolte, S., and Cordelières, F. P. (2006) A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224**, 213–232
25. Edgar, R. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797
26. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191
27. Cole, C., Barber, J. D., and Barton, G. J. (2008) The Jpred 3 secondary structure prediction server. *Nucleic Acids Res.* **36**, W197–W201
28. Schuster-Böckler, B., and Bateman, A. (2005) Visualizing profile-profile alignment: pairwise HMM logos. *Bioinformatics* **21**, 2912–2913
29. Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381–3385
30. Savage, D. F., Afonso, B., Chen, A. H., and Silver, P. A. (2010) Spatially ordered dynamics of the bacterial carbon fixation machinery. *Science* **327**, 1258–1261
31. Raetz, C. R., and Roderick, S. L. (1995) A left-handed parallel β-helix in the structure of UDP-N-acetylglucosamine acyltransferase. *Science* **270**, 997–1000
32. Schulz, E. C., Bergfeld, A. K., Ficner, R., and Mühlenhoff M. (2011) Crystal structure analysis of the polysialic acid specific O-acetyltransferase NeuO. *PLoS ONE* **6**, e17403
33. Friedberg, D., Kaplan, A., Ariel, R., Kessel, M., and Seijffers, J. (1989) The 5′-flanking region of the gene encoding the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase is crucial for growth of the cyanobacterium *Synechococcus* sp. strain PCC 7942 at the level of CO2 in air. *J. Bacteriol.* **171**, 6069–6076