THE PROTEIN SOCIETY

Bacterial microcompartment shells of diverse functional types possess pentameric vertex proteins

Nicole M. Wheatley,[1] Soheil D. Gidaniyan,[2] Yuxi Liu,[3] Duilio Cascio,[2] and Todd O. Yeates[1,2,3]*

[1]Molecular Biology Institute, University of California, Los Angeles, California
[2]UCLA-DOE Institute for Genomics and Proteomics, University of California, Los Angeles, California
[3]Department of Chemistry and Biochemistry, University of California, Los Angeles, California

Received 11 January 2013; Revised 18 February 2013; Accepted 22 February 2013
DOI: 10.1002/pro.2246
Published online 4 March 2013 proteinscience.org

Abstract: Bacterial microcompartments (MCPs) are large proteinaceous structures comprised of a roughly icosahedral shell and a series of encapsulated enzymes. MCPs carrying out three different metabolic functions have been characterized in some detail, while gene expression and bioinformatics studies have implicated other types, including one believed to perform glycyl radical-based metabolism of 1,2-propanediol (Grp). Here we report the crystal structure of a protein (GrpN), which is presumed to be part of the shell of a Grp-type MCP in *Rhodospirillum rubrum* F11. GrpN is homologous to a family of proteins (EutN/PduN/CcmL/CsoS4) whose members have been implicated in forming the vertices of MCP shells. Consistent with that notion, the crystal structure of GrpN revealed a pentameric assembly. That observation revived an outstanding question about the oligomeric state of this protein family: pentameric forms (for CcmL and CsoS4A) and a hexameric form (for EutN) had both been observed in previous crystal structures. To clarify these confounding observations, we revisited the case of EutN. We developed a molecular biology-based method for accurately determining the number of subunits in homo-oligomeric proteins, and found unequivocally that EutN is a pentamer in solution. Based on these convergent findings, we propose the name bacterial microcompartment vertex for this special family of MCP shell proteins.

Keywords: OCAC; oligomeric state determination; EutN; bacterial microcompartments; protein assembly; capsid; glycyl radical; BMV; pentameric vertex protein; GrpN

Introduction

Microcompartments (MCPs) are polyhedrally shaped supramolecular protein assemblies that physically encapsulate select metabolic pathways in prokaryotes (reviewed in Refs. 1–5). This encapsulation serves to concentrate and separate specific enzymes and their metabolic intermediates from the cytoplasm, thereby increasing reaction efficiency, retaining volatile intermediates, and/or protecting the cellular milieu from toxic intermediates.[5] The carboxysome MCP increases carbon-fixation efficiency by encapsulating carbonic anhydrase and RuBisCO together, so the CO_2 produced by the first enzyme can be delivered at high concentration to the second enzyme.[4,6] Cofactor B_{12}-dependent MCPs for

Abbreviations: BMC, bacterial microcompartment (hexameric) shell protein; BMV, bacterial microcompartment vertex protein (pentameric); Grp, glycyl radical-based propanediol utilizing MCP; MCP, bacterial microcompartment; OCAC, oligomeric characterization by addition of charge; OCAM, oligomeric characterization by addition of mass.

Soheil D. Gidaniyan and Yuxi Liu contributed equally to this work.

Grant sponsor: NIH grant; Grant numbers: R01AI08114, P41RR015301, P41GM103403; Grant sponsor: Ruth L. Kirschstein National Research Service Award; Grant number: GM007185; Grant sponsor: DOE Grant; Grant numbers: DE-FC02-02ER63421, DE-AC02-06CH11357.

**Correspondence to:* Todd O. Yeates, UCLA Department of Chemistry and Biochemistry, 611 Charles Young, Dr. East, Los Angeles, CA 90095-1569. E-mail: yeates@mbi.ucla.edu

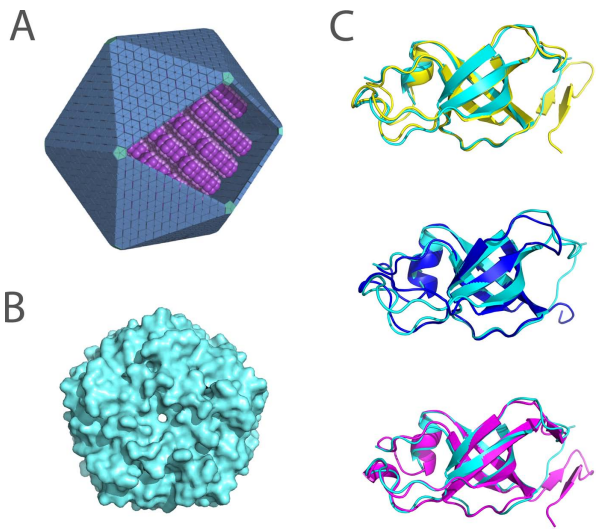

Figure 1. Comparison of the structure of GrpN with other bacterial microcompartment vertex (BMV) proteins. (A) An idealized model of an MCP showing pentameric units at the vertices of the polyhedral shell. (B) Space-filling model of the pentameric structure of GrpN. (C) Superposition of a GrpN monomer with CcmL (top, yellow), CsoS4A (middle, blue), and EutN (bottom, magenta). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

ethanolamine utilization (Eut) and 1,2-propandiol utilization (Pdu) function by retaining volatile or toxic aldehyde compounds—propionaldehyde and acetaldehyde—that occur as intermediates in those pathways.[5,7–11] Bioinformatic studies have suggested the existence of an MCP for glycyl radical-based 1,2-propanediol metabolism, referred to hereafter as Grp.[12] Sequence analysis and gene expression data suggest that the Grp MCP, like the Pdu MCP, houses enzymes for metabolizing 1,2-propanediol, but that it uses a glycyl radical enzyme for the key dehydration reaction rather than B_{12}-dependent enzymes.[12,13] The Grp MCP has not yet been studied in detail.

Structural analyses have answered a number of questions regarding the geometry and mechanism of MCP assembly and function (reviewed in Ref. 2, 3). Current models describe MCPs as being comprised of sheets of hexameric shell proteins (belonging to the BMC family of proteins) forming polyhedral facets,[14–16] along with a much smaller number of special pentameric proteins (belonging to the EutN/PduN/CcmL/CsoS4 protein family) placed at the vertices of the polyhedral shell.[17] This family of presumptive vertex proteins appears distinct in sequence and structure from the BMC family of proteins. Although crystallographic studies of (hexameric) BMC shell proteins have consistently supported this model of MCP organization, experiments on the vertex proteins have provided a less coherent picture. Two crystal structures of carboxysome vertex proteins, CcmL and CsoS4A, are indeed pentameric[17] (PDB accession code 2QW7 and 2RCF). However, a third structure of a presumptive vertex protein, EutN from *E. coli* (PDB code 2Z9H), revealed a hexameric quaternary assembly with nearly hexagonal shape.[18] Likewise, multiple genetic studies have shown that deleting the vertex proteins compromises the formation of closed polyhedral shells,[19–21] yet normal looking shells could be formed with lower efficiency in one study.[22] These equivocal results have complicated the interpretation of the architectural role of this protein family.

Electron microscopy studies indicate that carboxysomes tend to be more geometrically regular and more nearly icosahedral in shape compared to other types of MCPs that have been visualized[23,24]; this has raised the possibility that pentameric units required for symmetric icosahedral architecture might be present only in carboxysome MCPs. Here we provide evidence that this family of proteins serves as the pentameric vertex element across multiple divergent types of MCPs. Specifically, we report the pentameric crystal structure of GrpN from the Grp MCP, and show also that EutN is a pentamer in solution, contrary to previous crystallographic findings.

Results and Discussion

An operon encoding enzymes and shell proteins for a presumptive Grp MCP has been described in *R. rubrum*.[12] One of the encoded proteins, referred to hereafter as GrpN (Gene ID: 12642037), was identified as a representative of the EutN/PduN/CcmL/CsoS4 family of shell proteins (Pfam domain: PF03319), hereafter described as bacterial microcompartment vertex (or BMV) proteins. GrpN was overexpressed in *E. coli* and purified by metal affinity chromatography on the basis of a polyhistidine tail added to the protein. Single crystals were obtained in space group I23, and an X-ray structure was determined by molecular replacement and refined at a resolution of 3.2 Å. The final atomic model is 96% complete; residues 65–70 are in a disordered loop and could not be modeled accurately.

As expected, the tertiary fold follows closely that reported earlier for other members of this protein family (Fig. 1). The core structure of GrpN is comprised of five antiparallel beta strands that curve to create a small beta barrel. The C-termini and loop regions between beta strands extend away from the beta-barrels, making contacts with adjoining subunits. The atomic coordinates align well to CcmL, CsoS4A, and EutN, with RMSD values of 0.73 Å, 0.81 Å, and 0.80 Å, respectively, over C-alpha positions. Stabilization of the quaternary structure is largely mediated by contacts between loop regions from adjacent beta-barrels; minor contributions are made by barrel-to-barrel contacts. Structural comparisons between BMV structures show that the N-termini and the small beta barrels are highly conserved, while the C-termini and regions adjoining

Table 1. *X-ray Data Collection and Model Refinement Statistics*

Statistics	Value
Wavelength (Å)	0.9791
Resolution range (Å)	19.8–3.2 (3.314–3.2)
Space group	I23
Unit cell (Å)	$a=b=c=150.7$
Total reflections recorded	49,986 (3856)
Unique reflections	9,248 (700)
Multiplicity	5.4 (5.5)
Completeness (%)	96.6 (99.6)
Mean I/sigma(I)	21.0 (3.0)
Wilson B-factor (Å^2)	84.4
R-meas[25]	7.3% (62%)
Model R-work	0.267 (0.340)
Model R-free	0.288 (0.381)
Number of non-hydrogen atoms	2,930
Macromolecules	2,926
Chloride	3
Water	1
Protein residues	423
Geometric deviations (rms)	
Bonds (Å)	0.004
Angles (°)	0.69
Ramachandran favored (%)	95
Ramachandran outliers (%)	0
Clashscore[26]	3813.35
Average B-factor (Å^2)	87.6
Protein atoms	87.6
Solvent (water)	82.2

Statistics for the highest-resolution shell are shown in parentheses.

the core beta sheets differ slightly in length and secondary structure content. For example, GrpN and CsoS4A lack C-terminal beta hairpins present in both CcmL and EutN.

GrpN crystallized as a pentamer in the asymmetric unit of the crystal (Table I). The pentameric unit of GrpN matches those reported earlier for CcmL from the beta-type carboxysome and CsoS4A from the alpha-type carboxysome. One surprising aspect of the GrpN crystal structure is the arrangement of pentamers in the unit cell. Twelve crystallographically-related pentamers are arranged in a nearly icosahedral fashion. This 60-subunit icosahedral arrangement is unlikely to reflect the natural assembly of the much larger MCP; in native MCPs, pentameric units are presumably surrounded by BMC-type hexamers. In addition, the interactions between pentamers in the crystal involve relatively small sites of contact. Nonetheless, this spontaneous assembly could be an interesting starting point for engineering a novel icosahedral cage. The design of self-assembling protein cages by various methods has been discussed in a series of recent studies.[27–31] The observation that GrpN is a pentamer establishes that other MCPs besides carboxysomes possess special pentameric shell proteins to explain the closure of an otherwise flat layer of hexameric units provided by the BMC-type proteins.

Our finding that GrpN forms a pentameric assembly left the earlier reported hexameric structure for *E. coli* EutN as an outlying observation. The hexameric state reported for EutN could either indicate an unusual structural role for this homolog, or it could represents a spurious or minor oligomeric form selected during crystallization. To distinguish between these two possibilities, we devised a novel method to determine the oligomeric state of EutN in solution (Fig. 2).

The method we developed for determining oligomeric states is based on the change in charge-to-mass ratio that occurs after proteolytic cleavage of a small, charged N- or C-terminal sequence that is genetically appended at the end of the native protein subunit [Fig 2(A)]. Following different degrees of exposure to protease, a homo-oligomer composed of n subunits can be partially proteolyzed to generate a total of $n+1$ distinct charge forms. The distinct charge forms can be enumerated by native gel electrophoresis. To be consistent with the nomenclature of established methods, we refer to our approach as OCAC, an acronym for *oligomeric characterization by addition of charge* [Fig. 2(B)]. A related method described earlier, OCAM (*oligomeric characterization*

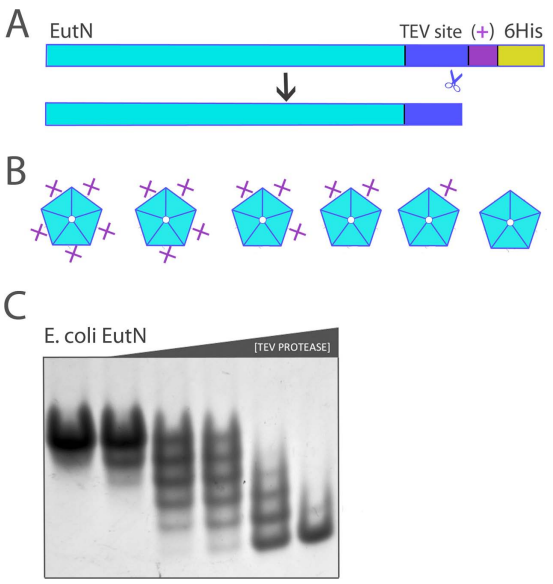

Figure 2. Application of the OCAC (oligomeric characterization by the addition of charge) method to the EutN shell protein. (A) Diagram of primary structure of EutN before and after cleavage with TEV protease. The plus symbol "+" denotes one additional positive charge. (B) For an n-oligomer, $n+1$ possible charge states exist. Shown are the six possible charge-states of a pentamer. (C) EutN OCAC native gel. The engineered EutN(+) oligomer was incubated with TEV protease for 15 min, quenched with iodoacetamide, and then run on a native gel. From left to right, TEV protease concentrations are 0.0, 0.002, 0.01, 0.02, 0.1, and 0.5 mg/mL. The presence of six distinct bands shows that EutN is a pentamer. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

by addition of mass), was developed to determine the oligomeric state of membrane proteins.[32] OCAM determines oligomeric state based on changes in the mobility of protein complexes on blue native gels, which arise from differential removal of relatively massive domains fused to the native protein subunits. In contrast, OCAC requires addition of relatively small terminal extensions to the native protein: essentially a protease cleavage site and as few as one or two directly adjacent charged amino acids. As with OCAM, the success of OCAC depends on a slow rate of exchange between subunits in different oligomeric complexes. Under conditions where we applied the method, native gels were capable of distinguishing between oligomers with just one unit charge difference. Another method based on counting distinct oligomeric forms was described several years ago, wherein the chemical modification of a genetically engineered cysteine residue by a charged reagent resulted in mobility shifts on SDS-PAGE.[33] Unlike this earlier method, OCAC uses limited proteolysis to create alternative charge forms and does not require the complex to be stable in the presence of sodium dodecyl sulfate. The OCAC method also avoids the potential challenges of using chemical approaches when multiple cysteine residues are present in a protein.

We applied the OCAC method to examine the oligomeric state of EutN, which we had earlier reported to be a hexamer based on its crystal structure. EutN from *E. coli* was cloned to include a C-terminal TEV site followed by LEKK-6His. The resulting net charge difference between uncleaved and cleaved EutN is one charge unit (at pH 8.6 of the native gel running buffer). After recombinant expression and purification, this construct was subjected to a series of increasing concentrations of TEV protease, and then run on a native gel.

Surprisingly, the results clearly indicated that EutN is a stable pentamer in solution [Fig. 2(C)]. Six sharp bands are obtained by separation on a native gel. These arise from oligomers having 0, 1, 2, 3, 4, or 5 tails cleaved. This pentameric behavior contradicts two independent crystal structures in the PDB showing EutN in a hexameric arrangement[18](PDB ID 2Z9H and 2HD3). A comparison of conditions used to obtain crystals of EutN did not suggest a clear explanation for the observed difference between crystalline and solution oligomer states. Although we see no evidence in the OCAC method for hexameric assembly, it is possible that a minor hexameric species of EutN exists in solution, which happens to crystallize preferentially. If the protein could equilibrate between pentameric and (minor) hexameric forms in solution, a low energy crystal form could drive the protein to a hexameric configuration by mass action. Such a scenario, however, would be difficult to reconcile with the conclusion that the pentameric form is stable at least over the time course of the OCAC experiment. Under one interpretation of the data, the hexameric form of EutN that is apparently populated under crystallization conditions represents a spurious form. Alternatively, the heterogeneous oligomerization behavior revealed by studies in solution and in crystals could reflect genuine complexity in the way these proteins are used to build microcompartment shells. A finer dissection of oligomeric states in intact MCPs must await higher resolution electron microscopy studies.

Notwithstanding the possibility of complex oligomerization phenomena, our combined results demonstrating the pentameric states of both GrpN and EutN help resolve an outstanding question about microcompartment architecture. Although the role of BMC type hexamers in forming the flat facets of MCPs has been apparent since the first structural studies,[14,16] the universal requirement for pentameric bacterial microcompartment vertex (BMV) proteins in MCPs has remained uncertain.[18,22] Here we confirm the first pentameric assemblies of BMV proteins from MCPs besides the carboxysome: GrpN from *Rhodospirillum rubrum* by X-ray crystallography, and EutN from *E. coli* in solution by the OCAC method. These results help reunify ideas for how different types of MCPs are constructed.

Materials and Methods

Cloning

A gene sequence was designed, using codons optimized for expression in *E. coli*, to encode GrpN with a C-terminal hexahistidine tag, using the online program DNAWorks.[34] The resulting nucleotide sequence was synthesized by Biomatik, and then transferred to pET22b(+) vector using the NdeI restriction site via Isothermal Assembly (a.k.a. Gibson Assembly).[35] We followed the isothermal protocol utilizing 20 nucleotide base pair complementary overhangs.

Full-length EutN was amplified from *E. coli* genomic DNA. Synthetic DNA oligomers, purchased from IDT, were used to add a TEV cleavage site and the positively charged tail onto the C-terminus of EutN with PCR. We refer to the resulting protein as EutN(+). Using Isothermal Assembly, EutN(+) was transferred into a pET22b(+) vector between NdeI and XhoI restriction sites to append a C-terminal 6xHis tag. The sequences of both GrpN and EutN(+) were verified by Laragen, Inc.

Expression and purification

The expression of GrpN-6His and EutN(+) were induced with 1mM IPTG in *E. coli* BL21 cells, shaking at 250 rpm, for 3 to 5 h at 37°C. Cells were spun

down for 5 min at 6000 rpm and stored at –20°C. Cells were suspended in 50 m*M* Tris-HCl pH 7.6, 300 m*M* NaCl, 20 m*M* imidazole, with Protease Inhibitor Cocktail (Sigma, Cat # P8849) and lysed by sonication. Cells were spun down in rotor SS-34 at 16,500 rpm for 30 min, filtered through a 0.2 μm filter, and applied to a Hi-trap Nickel column by syringe at room temperature. Protein was eluted in one step with of 50 m*M* Tris–HCl pH 7.6, 300 m*M* NaCl, 400 m*M* Imidazole. GrpN was dialyzed into 2L of 10 m*M* Tris-HCl pH 7.6, 20 m*M* NaCl for 1 h at 4°C, and then again in 2L of fresh buffer overnight. GrpN-6His was highly soluble, but did show a tendency to precipitate over time. The protein was therefore centrifuged once or twice daily (20,800 rcf for 1–2 min at 4°C) and the pelleted precipitate was discarded.

Crystal structure determination

Initial crystallization screens were performed in 96-well, hanging drop trays, set up with the nanoliter liquid handling Mosquito from TTP LabTech. Upon optimization of condition G8 from Hampton Research screen HR2-110, cube-shaped crystals were obtained within the following condition ranges: 1.4–1.6 *M* ammonium sulfate, 0.1 *M* NaCl, 0.1 *M* HEPES pH 7.0–7.6, at protein concentrations between 20 and 40 mg/mL. Hanging drops were 1:1 well: protein with a total drop size between 2 and 4 μL. Crystals generally took between 1 and 10 days to grow.

Diffraction data extending to 3.2 Å resolution were collected at the Argonne National Laboratory, Advanced Photon Source (APS), beamline 24-ID-C. The structure of GrpN was phased by molecular replacement using the program PHASER.[36] Coordinates for the CcmL pentamer (PDB accession code 2QW7) were used as the search model. The structure was built using the program COOT[37] and refined using PHENIX[38] and BUSTER[39] with a final R_{work} and R_{free} of 0.2671 and 0.2885, respectively. 95% of the backbone dihedral angles are within the favored regions of a Ramachandran diagram. Coordinates and structure factors have been deposited with the PDB with ID code 4I7A.

OCAC assay

Aliquots of purified EutN(+) (in 50 m*M* Tris-HCl pH 7.6, 300 m*M* NaCl and 300 m*M* imidazole) were incubated on ice with fresh 1 m*M* DTT for 30 min. Because of weak UV absorbance at 280 nm, concentrations were adjusted based on band intensity as visualized by SDS-PAGE and Coomassie staining. TEV protease was added to EutN(+) aliquots in a series of dilutions, resulting in the following final concentrations: 0.0, 0.002, 0.01, 0.02, 0.1, and 0.5 mg/mL TEV protease. These TEV: EutN(+) samples were incubated at room temperature for 15 min. Protease reactions were stopped with addition of 5× native loading dye containing 10 m*M* iodoacetamide. Reactions were then run on a native gel (BioRad CAT# 456-1096) at 100 V for 2 h at room temperature.

Acknowledgments

We thank Michael Thompson, Julien Jorda, Thomas Bobik and members of the Yeates lab for helpful discussions. We thank Michael Sawaya and the staff of the UCLA-DOE X-ray Crystallography Core Facility. We thank M. Capel, K. Rajashankar, N. Sukumar, J. Schuermann, I. Kourinov, and F. Murphy at NECAT beamlines 24-ID at APS.

In a new study published after this manuscript was submitted, Petit et al. demonstrate the production of an MCP that appears to metabolize 1,2-PD using a glycyl radical based reaction of the type discussed here and in refs (12, 13). See Petit E, Latouf WG, Coppi MV, Warnick TA, Currie D, Romashko I, Deshpande S, Haas K, Alvelo-Maurosa JG, Wardman C, Schnell DJ, Leschine SB, Blanchard JL. (2013). Involvement of a bacterial microcompartment in the metabolism of fucose and rhamnose by Clostridium phytofermentans. PLoS One 8(1):e54337.

References

1. Kerfeld CA, Heinhorst S, Cannon GC (2010) Bacterial microcompartments. Annu Rev Microbiol 64:391–408.
2. Yeates TO, Thompson MC, Bobik TA (2011) The protein shells of bacterial microcompartment organelles. Curr Opin Struct Biol 21:223–231.
3. Yeates TO, Crowley CS, Tanaka S (2010) Bacterial microcompartment organelles: protein shell structure and evolution. Annu Rev Biophys 39:185–205.
4. Cannon GC, Bradburne CE, Aldrich HC, Baker SH, Heinhorst S, Shively JM (2001) Microcompartments in prokaryotes: carboxysomes and related polyhedra. Appl Environ Microbiol 67:5351–5361.
5. Cheng S, Liu Y, Crowley CS, Yeates TO, Bobik TA (2008) Bacterial microcompartments: their properties and paradoxes. Bioessays 30:1084–1095.
6. Price GD, Coleman JR, Badger MR (1992) Association of carbonic anhydrase activity with carboxysomes isolated from the cyanobacterium *Synechococcus* PCC7942. Plant Physiol 100:784–793.
7. Sampson EM, Bobik TA (2008) Microcompartments for B_{12}-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate. J Bacteriol 190:2966–2971.
8. Penrod JT, Roth JR (2006) Conserving a volatile metabolite: a role for carboxysome-like organelles in *Salmonella enterica*. J Bacteriol 188:2865–2874.
9. Brinsmade SR, Paldon T, Escalante-Semerena JC (2005) Minimal functions and physiological conditions required for growth of *Salmonella enterica* on ethanolamine in the absence of the metabolosome. J Bacteriol 187:8039–8046.
10. Stojiljkovic I, Bäumler A, Heffron F (1995) 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.

11. Bobik T, Xu Y, Jeter R, Otto K, Roth J (1997) Propanediol utilization genes (pdu) of *Salmonella typhimurium*: three genes for the propanediol dehydratase. J Bacteriol 21:6633–6639.
12. Jorda J, Lopez D, Wheatley NM, Yeates TO (2012) Using comparative genomics to uncover new kinds of protein-based metabolic organelles in bacteria. Protein Sci 22:179–195.
13. Scott KP, Martin JC, Campbell G, Mayer C, Flint HJ (2006) Whole-genome transcription profiling reveals genes up-regulated by growth on fucose in the human gut bacterium Roseburia inulinivorans. J Bacteriol 188: 4340–4349.
14. Kerfeld CA, Sawaya MR, Tanaka S, Nguyen CV, Phillips M, Beeby M, Yeates TO (2005) Protein structures forming the shell of primitive bacterial organelles. Science 309:936–938.
15. Samborska B, Kimber MS (2012) A dodecameric CcmK2 structure suggests β-carboxysomal shell facets have a double-layered organization. Structure 20: 1353–1362.
16. Tsai Y, Sawaya MR, Cannon GC, Cai F, Williams EB, Heinhorst S, Kerfeld CA, Yeates TO (2007) Structural analysis of CsoS1A and the protein shell of the *Halothiobacillus neapolitanus* carboxysome. PLoS Biol 5: e144.
17. Tanaka S, Kerfeld CA, Sawaya MR, Cai F, Heinhorst S, Cannon GC, Yeates TO (2008) Atomic-level models of the bacterial carboxysome shell. Science 319: 1083–1086.
18. Tanaka S, Sawaya MR, Yeates TO (2010) Structure and mechanisms of a protein-based organelle in *Escherichia coli*. Science 327:81–84.
19. Price GD, Howitt SM, Harrison K, Badger MR (1993) Analysis of a genomic DNA region from the cyanobacterium *Synechococcus sp. strain* PCC7942 involved in carboxysome assembly and function. J Bacteriol 175: 2871–2879.
20. Cheng S, Sinha S, Fan C, Liu Y, Bobik TA (2011) Genetic analysis of the protein shell of the microcompartments involved in coenzyme B_{12}-dependent 1,2-propanediol degradation by *Salmonella*. J Bacteriol 193: 1385–1392.
21. Parsons JB, Dinesh SD, Deery E, Leech HK, Brindley AA, Heldt D, Frank S, Smales CM, Lünsdorf H, Rambach A, Gass MH, Bleloch A, McClean KJ, Munro AW, Rigby SE, Warren MJ, Prentice MB. (2008) Biochemical and structural insights into bacterial organelle form and biogenesis. J Bacteriol 283:14366–14375.
22. Cai F, Menon BB, Cannon GC, Curry KJ, Shively JM, Heinhorst S (2009) The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO_2 leakage barrier. PLoS ONE 4:e7521.
23. Schmid MF, Paredes AM, Khant HA, Soyer F, Aldrich HC, Chiu W, Shively JM (2006) Structure of *Halothiobacillus neapolitanus* carboxysomes by cryo-electron tomography. J Mol Biol 364:526–535.
24. Iancu CV, Ding HJ, Morris DM, Dias DP, Gonzales AD, Martino A, Jensen GJ (2007) The structure of isolated *Synechococcus strain* WH8102 carboxysomes as revealed by electron cryotomography. J Mol Biol 372: 764–773.
25. Diederichs K, Karplus PA (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. Nat Struct Biol 4:269–275.
26. Word JM, Lovell SC, LaBean TH, Taylor HC, Zalis ME, Presley BK, Richardson JS, Richardson DC (1999) Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. J Mol Biol 285:1711–1733.
27. Lai Y-T, King NP, Yeates TO (2012) Principles for designing ordered protein assemblies. Trends Cell Biol 22:653–661.
28. Lai Y-T, Cascio D, Yeates TO (2012) Structure of a 16-nm cage designed by using protein oligomers. Science 336:1129.
29. Yang Y, Burkhard P (2012) Encapsulation of gold nanoparticles into self-assembling protein nanoparticles. J Nanobiotech 10:42.
30. King NP, Sheffler W, Sawaya MR, Vollmar BS, Sumida JP, André I, Gonen T, Yeates TO, Baker D (2012) Computational design of self-assembling protein nanomaterials with atomic level accuracy. Science 336: 1171–1174.
31. Wörsdörfer B, Woycechowsky KJ, Hilvert D (2011) Directed evolution of a protein container. Science 331: 589–592.
32. Gandhi CS, Walton TA, Rees DC (2011) OCAM: a new tool for studying the oligomeric diversity of MscL channels. Protein Sci 20:313–326.
33. Gouaux J, Braha O, Hobaugh M, Song L, Cheley S, Shustak C, Bayley H (1994) Subunit stoichiometry of staphylococcal alpha-hemolysin in crystals and on membranes: a heptameric transmembrane pore. Proc Natl Acad Sci USA 91:12828–12831.
34. Hoover DM (2002) DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. Nucleic Acids Res 30:43e–43.
35. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA III, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6:343–345.
36. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. J Appl Cryst 40:658–674.
37. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Cryst D60:2126–2132.
38. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Cryst D66:213–221.
39. Headd JJ, Echols N, Afonine PV, Grosse-Kunstleve RW, Chen VB, Moriarty NW, Richardson DC, Richardson JS, Adams PD (2012) Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. Acta Cryst D68:381–390.