Title

Bacterial microcompartments

Journal

Nature Reviews Microbiology, 16(5)

ISSN

1740-1526

Authors

Kerfeld, Cheryl A

Aussignargues, Clement

Zarzycki, Jan

et al.

Publication Date

2018-05-01

DOI

10.1038/nrmicro.2018.10

Peer reviewed

HHS Public Access

Author manuscript

Nat Rev Microbiol. Author manuscript; available in PMC 2018 November 01.

Published in final edited form as:

Nat Rev Microbiol. 2018 May ; 16(5): 277–290. doi:10.1038/nrmicro.2018.10.

Bacterial microcompartments

Cheryl A. Kerfeld[1,2,3,*], **Clement Aussignargues**[1,3], **Jan Zarzycki**[4], **Fei Cai**[2], and **Markus Sutter**[1,2,3]

[1]MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI 48824, USA

[2]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[3]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA

[4]Max-Planck-Institute for Terrestrial Microbiology, D-35043, Marburg, Germany

Abstract

Bacterial microcompartments (BMCs) are self-assembling organelles that consist of an enzymatic core that is encapsulated by a selectively permeable protein shell. The potential to form BMCs is widespread, found across the Kingdom Bacteria. BMCs have crucial roles in carbon dioxide fixation in autotrophs and the catabolism of organic substrates in heterotrophs. They contribute to the metabolic versatility of bacteria, providing a competitive advantage in specific environmental niches. Although BMCs were first visualized more than sixty years ago, it is mainly in the last decade that progress has been made in understanding their metabolic diversity and the structural basis of their assembly and function. This progress has not only heightened our understanding of their role in microbial metabolism but it is also beginning to enable their use in a variety of applications in synthetic biology. In this Review, we focus on recent insights into the structure, assembly, diversity and function of BMCs.

Table of contents blurb

Bacterial microcompartments are self-assembling organelles that consist of an enzymatic core that is encapsulated by a selectively permeable protein shell. In this Review, Kerfeld and colleagues discuss recent insights into the structure, assembly, diversity and function of bacterial microcompartments.

Correspondence to: C.A.K. ckerfeld@lbl.gov.

Competing interests statement

The authors declare no competing interests.

Author contributions

C.A.K., C.A., J.Z., F.C. and M.S. contributed to researching data for the article and its writing. C.A.K. and M.S. wrote the article.

Introduction

Bacterial microcompartments (BMCs) are functional analogs of eukaryotic organelles; a semi-permeable protein shell sequesters an enzymatic core that constitutes a segment of a metabolic pathway[1–8]. Instead of a lipid membrane, the restricting boundary of a BMC is a selectively permeable protein shell. The separation of the encapsulated enzymatic core from the cytosol is thought to protect the cell from toxic metabolic intermediates and prevent unwanted side reactions. Moreover, the co-localization of the encapsulated enzymes enhances flux through multi-step pathways[9] and may increase enzyme stability.

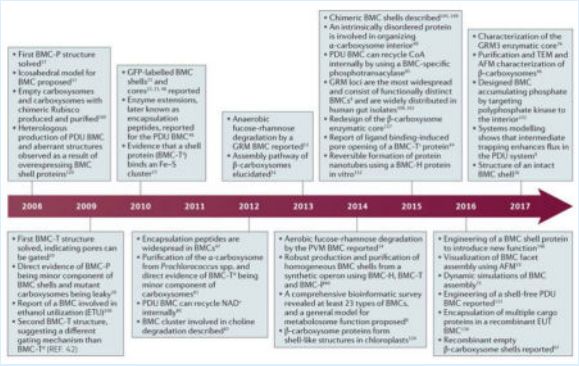
Although the basic architecture of the protein shell is conserved, BMCs are functionally diverse and are involved in both anabolic and catabolic processes. There is currently only one known example of an anabolic BMC — the carboxysome[2–4] — which is found in all cyanobacteria and in some chemotrophic bacteria. Carboxysomes contain the enzymes carbonic anhydrase[10] **[G]** and ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) **[G]** and function to fix carbon dioxide as part of the Calvin-Benson–Bassham cycle **[G]** (Figure 1A). Catabolic BMCs— also known as metabolosomes — are found across Bacterial phyla. They are involved in the metabolism of various organic compounds such as propanediol (PDU; propanediol utilization BMC), ethanolamine (EUT; ethanolamine utilization BMC), fucose and rhamnose[11–14] (Figure 1B). They are typically only expressed when their substrate is present[15, 16]. Metabolosomes frequently enable growth in specific niches and there is an increasing realization of their role in providing a metabolic advantage to pathogenic bacteria[17].

Genes encoding proteins that are structural components of BMCs and gene products for ancillary functions (such as transporters for the uptake of the substrate by the cell and cytoskeletal elements that are presumably involved in positioning the ‘organelle’ within the cytosol) are typically organized in superloci **[G]** [6] (Figure 1C). Accordingly, BMCs can be considered metabolic modules that are encoded in genetic modules that include genes for the organelle and the necessary components for integrating it with other cellular metabolic processes.

Owing to increases in the number of sequenced bacterial genomes and metagenomes in the past decade, the discovery of BMCs no longer depends on electron microscopy studies in conjunction with experimental characterization (reviewed in Refs [1,5,18]). Recognition of the homology among BMC shell proteins using bioinformatics has led to the identification of a multitude of functionally diverse BMCs, found across 19 out of the 29 established bacterial phyla and also in several candidate phyla (Figure 2A)[6]. Moreover, new methods of visualization such as atomic force microscopy[19], the crystal structure determination of an intact BMC shell[20] and labeling BMCs with fluorescent proteins[21–23] have provided new insights into BMC structure, assembly and subcellular localization. This knowledge has enabled the engineering of BMCs for new applications in synthetic biology[24–26]. Combined with biochemical and genetic analyses, these studies have established the new field of BMC biology. This review focuses on the advances of the last decade (Timeline) in which the widespread occurrence of BMCs has become clear and during which study of their structure,

function and assembly has opened up a frontier for the application of BMC-based architectures in bioengineering.

BMC timeline



2008

- First BMC-P structure solved[37]
- Icosahedral model for BMC proposed[37]
- Empty carboxysomes and carboxysomes with chimeric RuBisCO produced and purified[160]
- Heterologous production of PDU BMC and aberrant structures observed as a result of overexpressing BMC shell proteins[120]

2009

- First BMC-T structure solved, indicating pores can be gated[29]
- Direct evidence of BMC-P being minor component of BMC shells and mutant carboxysomes are leaky[39]
- Report of a BMC involved in ethanol utilization (ETU)[150]
- Second BMC-T structure, suggesting a different gating mechanism than BMC-T^{d} [42]

2010

- GFP-labeled BMC shells[21] and cores[22, 23, 48] reported
- Enzyme extensions, later known as encapsulation peptides, reported for the PDU BMC[48]
- Evidence that a shell protein (BMC-T^{s}) binds an iron-sulfur cluster[23]

2012

- Encapsulation peptides are widespread in BMCs[47]
- Purification of the α-carboxysome from *Prochlorococcus* and direct evidence of BMC-T^{d} being minor component of carboxysomes[83]

- PDU BMC can recycle NAD+ internally[80]
- BMC cluster involved in choline degradation described[87]

2013

- Anaerobic fucose/rhamnose degradation by a GRM BMC reported[13]
- Assembly pathway of β-carboxysomes elucidated[51]

2014

- Aerobic fucose/rhamnose degradation by the PVM BMC reported[14]
- Robust production and purification of homogeneous BMC shells from a synthetic operon using BMC-H, BMC-T and BMC-P[40]
- A comprehensive bioinformatic survey revealed at least 23 types of BMCs and a general model for metabolosome function proposed[6]
- B-carboxysome proteins form shell-like structures in chloroplasts[124]

2015

- Chimeric BMC shells described[145, 149]
- An intrinsically disordered protein is involved in organizing α-carboxysome interior[69]
- PDU BMC can recycle coenzyme A internally by using a BMC-specific phosphotransacylase[85]
- GRM loci are the most widespread and consist of functionally distinct BMCs [8] and are widely distributed in human gut isolates[108, 161]
- Redesign of the β-carboxysome enzymatic core[127]
- Report of ligand binding-induced pore opening of a BMC-T^s protein[44]
- Reversible formation of protein nanotubes using a BMC-H protein *in vitro*[152]

2016

- Engineering of a BMC shell protein to introduce new function [148]
- Visualization of BMC facet assembly using AFM [19]
- Dynamic simulations of BMC assembly[71]
- Engineering of a shell-free PDU BMC reported [133]
- Encapsulation of multiple cargo proteins in a recombinant EUT BMC[130]
- Recombinant empty β-carboxysome shells reported [67]

2017

- Characterization of the GRM3 enzymatic core[76]
- Purification and TEM and AFM characterization of β-carboxysomes [46]

- Designed BMC accumulating phosphate by targeting polyphosphate kinase to the interior[132]
- Systems modeling shows that intermediate trapping enhances flux in the PDU system[9]
- Structure of an intact BMC shell[20]

The bacterial microcompartment shell

The shell of all bacterial microcompartments is composed of three types of protein building blocks: BMC-H, BMC-T and BMC-P (Figure 3A). The most abundant type of shell protein is BMC-H **[G]**, containing a single Pfam00936 domain **[G]** that assembles into a cyclic homohexamer (Figure 3A) with a convex and a concave side[27]. BMC-T **[G]** proteins consist of a tandem fusion of two Pfam00936 domains; they form trimers that resemble the BMC-H hexamer in size and shape (that is, they form pseudohexamers)[28–32]. There are two subtypes of BMC-T proteins: a single trimer form, (BMC-T^s) and a second type in which two trimers dimerize across their concave faces, referred to as 'double ' BMC-T^d proteins (Figure 3A). A pore, typically formed at the central symmetry axis of hexamers and pseudohexamers, serves as a channel for metabolites to traverse the shell[27, 29, 33–35].

Whereas BMC-H and BMC-T proteins constitute the facets **[G]** of the shells, pentameric proteins are required to cap the vertices of the polyhedral shell[36, 37] (Figure 3B). This role is served by the third conserved building block, the BMC-P **[G]** proteins, which contain a single Pfam03319 domain and assemble into cyclical pentamers[30, 37, 38]. As icosahedra, regardless of their size, have only twelve vertices, the BMC-P proteins are minor components, yet are essential for the diffusive barrier provided by a completely closed shell[39]. Since the first shell protein crystal structure was reported in 2005[27], more than 40 others, in addition to various shell protein mutants or alternative crystal forms have been determined for both carboxysomes and metabolosomes. Despite large phylogenetic distances among the microorganisms in which they are found, the protein structures are highly conserved, indicative of shared principles of shell assembly. The structure of an intact shell composed of five different proteins from *Haliangium ochraceum*[20,40] provided atomic resolution detail of the protein-protein interactions that mediate shell assembly for the first time. In this shell of ~400 Å diameter, the protein membrane is formed by 540 individual protein chains, constituting a mass of about 6.5 MDa. The volume of this shell is estimated to be able to accommodate ~300 30 kDa proteins. One BMC-T^s and two BMC-T^d pseudohexamers occupy the center of the facets **[G]** and are encompassed by six BMC-H hexamers; the vertices are formed by the BMC-P pentamers (Figure 3B). The structure revealed that the concave sides of the hexamers and the bases of the truncated pyramid-shaped pentamers face the cytosol. The shell is composed of a single layer of protein in which stacked dimers of BMC-T^d proteins are incorporated with the second pseudohexamer protruding outwards. The structure also revealed key interactions between the different shell proteins, such as those that are mediated by lysine and arginine residues that are highly conserved among all BMC-H proteins irrespective of functional type. This indicates that the basic building blocks and architecture of the BMC shell are conserved[20].

The intact shell structure also demonstrated a key structural role for the BMC-T proteins. The BMC-T pseudohexamer interacts with two different BMC-H interfaces (Figure 3C), one in which the hexamers are coplanar and a second in which they adjoin at a 30° angle. At this position, BMC-T can bridge the angled vertices and the flat sheets that form the facets of the shell. BMC-T proteins have two copies of the Pfam00936 domain in each protein chain which enables each subunit in the trimer to participate in two distinct modes of interaction. Indeed, a homohexamer with its strict six-fold symmetry could not readily accommodate two different interactions simultaneously. Larger icosahedral shell models can also be deduced using these same principles[20]. This not only provides insight into the structural foundations of the observed size range of BMCs in bacteria (40–600 nm[14,41]), but also provides a strategy for designing the size of and the surface-to-volume ratios of engineered shells.

The close associations between the shell proteins in BMCs (Figure 3B) indicates that the central pores of the hexamers and pseudohexamers are the only channels to the interior. This is consistent with studies that show changes in BMC function when residues surrounding the pore are mutated[32,33]. The double-stacking BMC-T^d proteins contain a relatively large pore (12–14 Å) that can be open or closed depending on the conformation of the surrounding sidechains[28–31].

The dimerization of the BMC-T^d trimers across their concave faces results in an interior cavity containing conserved pockets in which ligands that are thought to be involved in the gating of the pore are bound[28, 31]. Alternative pore configurations have also been observed for the BMC-T^s pseudohexamers; multiple crystal structures of the metabolosome shell proteins EutL and PduB captured two distinct conformations; either a large central pore or three smaller pores situated around a closed center[35, 42–44]. Ligands bound to the smaller pores observed in the structures were suggested to reflect selectivity for specific metabolites[43] or to provide regulatory control over the opening of the central pore[44]; for example to allow the passage of larger molecules when regeneration of encapsulated cofactors is necessary.

A model for an entirely double-layered β-carboxysomal shell **[G]** was proposed based on biochemically identified interactions between carboxysomal BMC-H proteins[45]. Although two layers of hexamers seems unlikely based on BMC-H protein interactions in the intact shell crystal structure[20] and on the observed thickness of the shell layer in recently published transmission electron microscopy images of intact β-carboxysomes[46], an interaction between the concave sides of hexamers embedded in a BMC facet **[G]** and free cytosolic hexamers could presumably occur and represent an additional, perhaps transient, mechanism to modulate metabolite flow across the shell.

Assembly of bacterial microcompartments

While the overall BMC shell architecture is widely conserved, the encapsulated proteins are structurally and functionally diverse. Do all of these functionally distinct organelles assemble the same way? An understanding of the general principles of the assembly of BMCs is just beginning to emerge, led by studies of both alpha and beta carboxysomes.

Both types contain Form 1 RuBisCO which is a hexadecameric assembly of eight large and eight small subunits. For βcarboxysomes, the best studied example thus far, the core assembles first, followed by encapsulation by the shell (Figure 4A). By contrast, in α-carboxysomes **[G]** the shell and core seem to assemble concomitantly (Figure 4B). Given the observations that many core enzymes have a tendency to aggregate (summarized in REF. [7]) and that many metabolosome enzymes contain extensions that are known to interact with the shells[47–50], it is likely that metabolosomes have a similar assembly pathway where the core is assembled first like β-carboxysomes **[G]** [51]. Interestingly, based on phylogenetic analysis of their BMC-P protein, β-carboxysomes are more closely related to metabolosomes than to α-carboxysomes[2].

The β-carboxysome assembly pathway

Initial clues as to how carboxysomes assemble came from studies of specific domains of two highly conserved β-carboxysome proteins, CcmM and CcmN. Both proteins were shown to be essential for carboxysome formation as the deletion of either one results in the formation of large polar protein aggregates visible in cells in electron micrographs and a phenotype in which CO_2 supplementation is required for growth[47, 51–55]. A requirement for a high level of CO_2 is the typical indicator of loss of or defective carboxysomes. CcmM consist of an amino-terminal γcarbonic anhydrase **[G]** domain[56] and three to five domains at the carboxyl-terminus with amino acid sequence homology to RbcS, the small subunit of RuBisCO [57]. A shorter version of CcmM consisting of only these small subunit-like domains (SSLDs) is also expressed from an internal start site. Both full length CcmM and the shorter isoform are necessary for β-carboxysome formation[58, 59]. The γ-carbonic anhydrase domain of CcmM interacts with the N-terminal domain of CcmN[47, 60] (that contains β-helical domains that are also found in γ-carbonic anhydrases). CcmM also interacts with the additional β-carbonic anhydrase CcaA[58,60,61] (a completely different structural fold that only shares functional homology to γ-carbonic anhydrases), which is encoded in a satellite locus in a subset of β-cyanobacterial genomes[62]. The C-terminal region of CcmN contains an unstructured linker followed by a short (~17 amino acid) segment that is predicted to fold into an amphipathic α-helix[47]. This short peptide is referred to as an encapsulation peptide **[G]** based on its property to localize CcmN to the lumen of the carboxysome shell. Based on the sequence properties and predicted alpha helical structure of the CcmN peptide, Kinney et al[47] predicted that similar extensions were also found on diverse metabolosome proteins[47]. The alpha helical structure and the role of many of these encapsulation peptides in BMC assembly has been experimentally verified[48,49,63–65].

In 2010, three studies employing three different types of BMCs showed that it was possible to fluorescently label BMC cargo and shell proteins without abrogating assembly or function[21–23]. The ability to fluorescently label β-carboxysomal RuBisCO[21] and to control both shell and cargo protein expression and thus organelle formation *in vivo* enabled the visualization of the carboxysome assembly pathway[51,66]. The SSLDs of CcmM cause the aggregation of RuBisCO molecules, most likely by displacing some RbcS subunits of the RbcL-RbcS complex. This procarboxysome is encapsulated by shell proteins that interact with the γ-carbonic anhydrase domain of CcmM[67] or the encapsulation peptide of CcmN[47].

The addition of CcmL pentamers (the BMC-P of β-carboxysomes) closes the shell and the newly assembled carboxysome buds off[51], becoming distributed along the long axis of rod-shaped cyanobacteria through interactions with the cytoskeletal protein ParA[21]. An organized spatial distribution of carboxysomes in the cell is presumably advantageous because it reduces the mean distance that dissolved cytosolic bicarbonate has to diffuse to reach an organelle. It may also help to ensure uniform partitioning of carboxysomes to daughter cells during cell division[21].

The α-carboxysome assembly pathway

In contrast to β-carboxysome assembly, electron cryotomography has captured images of partially assembled α-carboxysome shells with attached core proteins suggests that the enzymes and the shell assemble concomitantly[68] (Figure 4B). Notably, with the exception of Form 1 RuBisCO and shell proteins, the components of α- and β-carboxysomes are structurally distinct. For example, the α-carboxysome contains a conserved intrinsically disordered protein, CsoS2, which is essential for its biogenesis[69,70]. CsoS2 has three sequence regions (N-terminal, middle and C-terminal);repetitive motifs are found in the N-terminal and middle regions. Some organisms express an isoform of CsoS2 that lacks the C-terminal region, CsoS2A, which is not capable of producing intact carboxysomes in a recombinant system in the absence of the full-length protein, CsoS2B[70]. Specific roles have been predicted for each of the regions[69,70]: the N-terminal region of CsoS2 recruits the major shell protein CsoS1, increasing its local concentration until it reaches the threshold that is required for self-assembly. Simultaneously, the middle region coalesces with RuBisCO and the C-terminal region anchors the resulting core to the shell[69].

In a recent *in silico* study[71] two types of BMC assembly could be distinguished. The encapsulation of cargo by a shell was simulated while varying the strength of the interactions between all of the components. When cargo-cargo interactions were relatively weak, a onestep assembly process, as predicted for α-carboxysome biogenesis[69], could be observed. When the strength of protein-protein interactions was increased, the formation of a core followed by its encapsulation in a shell and budding of the newly formed compartment was predicted, which is consistent with the experimentally determined β-carboxysome assembly pathway[51, 66].

Assembly of metabolosomes

Phylogenetic analysis of BMC-P proteins[2], differences in their core components and distinctive assembly pathways collectively suggest that β-carboxysomes are more similar to metabolosomes than to α-carboxysomes (Figure 2B, [2]). The assembly pathway of metabolosomes has yet to be experimentally demonstrated, but several lines of evidence suggest the sequence of events is very similar to β-carboxysome biogenesis. For example, in the glycyl radical enzyme containing microcompartments (GRMs), a subset of the enzymes contain domains with homology to their catalytically active, encapsulated counterparts; this is analogous to the role of the SSLDs of β-carboxysomes. Other examples of such ‘domain mimics’ include catalytically inactive aldehyde dehydrogenase domains or the N-terminal extension of some of the GRM-specific glycyl radical enzymes (GREs) **[G]** that resemble half of the catalytic domain of the GRE[6,8]. Likewise, encapsulation peptides are now an

established feature of many different metabolosome enzymes such as the aldehyde dehydrogenase PduP[72] and the phosphotransacylase PduL[50] of the PDU BMC, and the ethanolamine lyase small subunit (EutC) and ethanol dehydrogenase EutG of the EUT BMC[63].

Recent evidence suggests that at least some encapsulation peptides have a role in nucleating metabolosome cores. This is supported by the observation that they can render proteins insoluble[50, 73–76] or modify their oligomeric state[50]. Furthermore, large polar bodies containing cargo proteins, which are likely pro-metabolosomes, are present in PDU BMC-H protein deletion strains, suggesting that aggregation of the core is not dependent on shell formation[77, 78]. Accordingly, encapsulation peptides may have a dual function in metabolosome biogenesis, acting as a nucleation factor by aggregating proteins to form a pro-metabolosome, and recruiting the shell proteins around this core. In the PDU BMC, a short N-terminal α-helix of the major shell protein PduB is believed to be the primary mediator of the interaction between the core and the shell[79]; interaction between the encapsulation peptide of CcmN and the β carboxysome shell has similarly been demonstrated[47].

Proteins lacking encapsulation peptides or domain mimics are likely encapsulated through an interaction with a partner cargo protein that possesses one of these packaging determinants. For instance, the complex formed by the core enzymes alcohol and aldehyde dehydrogenase presumably targets the alcohol dehydrogenase to the pro-metabolosome of the *Salmonella enterica* PDU BMC[80]. The observation that cofactors such as CoA are bound to their cognate enzyme before their encapsulation in the BMC suggests that this enables the incorporation of cofactors during BMC biogenesis[50].

The functional diversity of BMCs

The observation that the genes encoding for the highly conserved shell proteins cluster with those encoding for the enzymatic core in operons, situated within superloci[6] (Figure 1C) has profoundly advanced the discovery of BMCs. Although the carboxysome, PDU[81] and EUT[82] BMCs were already known in the pre-genomic era, a comprehensive bioinformatic study in 2014[6] identified distinct subtypes of the PDU and EUT BMCs, as well as mergers of EUT and PDU loci, and an array of uncharacterized BMCs. The experimental study of these newly discovered BMCs is just beginning. For example, the Planctomycetes and Verrucomicrobia microcompartment (PVM BMC) from *Planctomyces limnophilus* was shown to be involved in the degradation of L-fucose and L-rhamnose, compounds that the organism is likely exposed to in its typical environmental niche[14]. The association of new BMC functions with specific microbial habitats is of particular interest for emerging applications in microbial community engineering.

The superloci of BMCs

BMC superloci **[G]** (Figure 1C) contain the genes for shell proteins (BMC-H, BMC-T and BMC-P) and encode the core proteins that consist of enzymes and putative scaffolding proteins[6]. In addition they frequently contain genes for regulatory proteins as well as factors to integrate the metabolic activity of the BMC with that of the rest of the cell. and to

maintain its function. For example, α-carboxysome superloci typically include genes encoding RuBisCO activase (CbbX), putative RuBisCO chaperones (CbbO-CbbQ complex), the low-CO_2-inducible bicarbonate transporter (SbtA-SbtB complex), and ChpX/NDH-I_4 complex (with low-affinity and constitutive CO_2 uptake activity)[83]. A recent structural analysis of CbbQ revealed the presence of a typical hexameric AAA+ ATPase domain; the protein was shown to have ATPase activity and to co-migrate with the carboxysomal shell fraction[84].

Metabolosomes

Despite different substrates and diverse functions, metabolosomes share a common enzymatic core composed of a signature enzyme, an aldehyde dehydrogenase (AldDH), an alcohol dehydrogenase (AlcDH) and a phosphotransacylase (PTAC) (Figure 1B). An aldehyde-generating enzyme, which differs between BMC types in reaction and substrate specificities, is referred to as the signature enzyme **[G]** because it defines the initial substrate and, accordingly, determines the main function of the metabolosome[6,7]. The aldehyde intermediate that is generated by the signature enzyme is subsequently processed by a CoA-dependent aldehyde dehydrogenase. Notably, the amino acid sequence of the AldDH can be used to independently reconstruct BMC phylogeny[6, 8]. An alcohol dehydrogenase and a phosphotransacylase (either a PduL-like/pfam06130[50,85] PTAC, which is found mostly associated with BMCs or a EutD-like/pfam01515[86] PTAC, a type that is ubiquitous and often referred to as housekeeping) recycle the NAD^+ and the CoA cofactors (Figure 1B).

Glycyl radical enzyme microcompartments

BMC loci encoding a microcompartment with a glycyl radical enzyme (GRE) as the signature enzyme are the most abundant type of BMC in number, but are only beginning to be experimentally characterized. The GRMs can be subdivided into at least five types (GRM1, GRM2, GRM3, GRM4 and GRM5) according to their locus composition[6, 8]. In contrast to other BMCs, the function of GRM organelles cannot be predicted based solely on the identification of the aldehyde-generating signature enzyme (that is, the GRE). In addition, more in-depth bioinformatic analysis of putative active site residues and/or experimental characterization of the GRE is required due to the functional diversity of this ancient family of enzymes. *In silico* analyses suggests that the five types of GRMs have distinct substrates[8]. Most of the data supporting the bioinformatic assignments of GRM function to date have been drawn from the analysis of GREs in isolation — these studies did not consider their activity within the context of a bacterial organelle. For example, the GREs of GRM1 and GRM2 were demonstrated to represent two different types of choline trimethylamine-lyases[87–89], indicative of a role for these BMCs in choline degradation. The first GRM to be experimentally characterized — the GRM5 — was shown to be involved in the anaerobic degradation of L-fucose and L-rhamnose[13]. This organelle is unusual in that it is associated with a second signature enzyme, an aldolase, which cleaves L-fuculose-1-phosphate or L-rhamnulose-1-phosphate into dihydroxyacetone phosphate and (*S*)-lactaldehyde. The latter is used by a lactaldehyde reductase to produce 1,2-propanediol. This led to the assumption that the GRE of GRM5 was a 1,2-propanediol dehydratase, which was later experimentally verified[90]. This activity was then also proposed for the GREs that are associated with the GRM3 and GRM4; phylogenetic analyses showed that the GREs of

GRM3 and GRM4 cluster closely together with the GREs of GRM5 in a maximum-likelihood tree **[G]** [8]. The GRE of the GRM3 of *Rhodopseudomonas palustris* was recently characterized biochemically and confirmed to be a 1,2-propanediol dehydratase[76]. The organelles of GRM3 and GRM4 are not associated with fuculose phosphate aldolases and some of the corresponding organisms do not possess even a distantly homologous enzyme. Therefore, GRM3 and GRM4 were proposed to directly function in the degradation of 1,2-propanediol from environmental sources[76]. Accordingly, the GRM3 and GRM4 represent vitamin B_{12}-independent alternatives to the canonical PDU BMCs[11].

BMCs awaiting characterization

In the most comprehensive bioinformatic survey to identify BMCs in bacterial genomes to date, there were three types for which a function could not be predicted[6]: MUF (metabolosome of unknown function), MIC (metabolosome with an incomplete core) and BUF (BMC of unknown function). The latter lacks all of the typical core enzymes and might harbor a new type of BMC-associated biochemistry[6]. Another BMC, namely RMM (*Rhodococcus* spp. and *Mycobacterium* spp. microcompartment) has been linked to aminoalcohol or aminoketone metabolism[91]. However, only one of the putatively encapsulated enzymes of RMM has been characterized, the L-1-amino-2-propanol dehydrogenase[92]. Very few representative loci are available for these relatively rare BMC types, making it difficult to predict which genes are conserved near those encoding shell proteins, and therefore likely to functionally associated with the organelle These questions will likely be addressed owing to the increasing number of available sequenced bacterial genomes[93].

An example of a MIC is the BMC from *Haliangium ochraceum* where only a putative AldDH gene is present in the most extensive BMC locus (containing BMC-P, BMC-H and BMC-T protein-encoding genes); an aldolase with a detectable encapsulation peptide is encoded in a distal part of the genome, clustered with two BMC-P homolog genes. The shell proteins from this organism have been shown to form homogeneous shells with a diameter of ~40 nm when recombinantly expressed[20,40]. More recently the range of enzymatic activities that are associated with BMCs has expanded. For example, in the candidate phylum Atribacteria, a BMC superlocus was found that contains 5–6 genes encoding shell proteins and genes annotated as NADH dehydrogenase, sugar isomerase and sugar aldolase in addition to the canonical core enzymes AldDH and PTAC[94]. Another unique type of BMC was recently identified in *Clostridium indolis* DSM 755^T; the locus includes a CoA-transferase gene, and was correspondingly named CoAT BMC[95].

The role of BMCs in pathogenesis

It is now well established that a subset of BMCs contributes to the fitness and virulence of some pathogenic bacteria[17]. The human gut[12,96] and processed food[97,98] are enriched in propanediol and ethanolamine that can be catabolized by the EUT, PDU and some GRM BMCs as a source of energy, carbon and nitrogen. For example, the PDU BMC of *Salmonella typhimurium* was found to contribute to colonization and expansion during infection, possibly by making use of nutrients that are unavailable to the normal flora[17,99,100]. Similarly, the EUT BMCs of *Listeria monocytogenes*, enterohemorrhagic

Escherichia coli O157:H7, *Enterococcus faecalis*, and *S. typhimurium* have been shown to be important for virulence in the host intestine in both *in vivo* and *in vitro* infection models[101–106]. EUT BMCs provide an advantage for enterohemorrhagic *E. coli*, especially when entering stationary phase upon nutrient limitation; bacteria able to metabolize ethanolamine are able to survive[103]. EUT BMCs are also present in other pathogens, such as the virulent multi-drug resistant *Clostridium difficile* 630, but a direct connection to pathogenesis has not yet been established[107].

Another mode by which BMCs contribute to pathogenicity results from the production of trimethylamine (TMA) in the human gut, a byproduct of the catabolism of the choline generated by GRM1 and GRM2. Both GRM1 and GRM2 are widespread in members of healthy human gut microbiomes and may be a substantial source of TMA[108]. TMA is accumulated in the gut and, after its resorption, it is rapidly converted to trimethylamine-N-oxide (TMAO) by flavin monooxygenase FMO3 in the liver[109,110]; it has been established by recent studies that the levels of TMAO in serum positively correlate with many human health issues such as cardiovascular disease, reduced renal function, colorectal cancer and impaired glucose tolerance[109,111–114]. Although a direct relationship between TMAO that is produced by GRM1 and GRM2 and disease has not been established, recent studies found that GRM1 or GRM2 BMCs were required and sufficient for metabolizing choline and for the production of TMA under anaerobic conditions in the human gut[108].

Analyses of the genomes of several different *Clostridium botulinum* isolates revealed numerous BMC-associated genes in putative pathogenicity islands **[G]** [115]. How these BMCs affect toxicity, invasiveness and the colonization of the host has not been studied yet. *C. botulinum* strains may possess multiple types of BMC gene clusters, such as GRM1, a metabolosome of unknown function (MUF), and additional satellite loci[6]. Collectively these data make it clear that BMCs as well as enzymes that function upstream or downstream of the BMC-associated metabolism represent attractive targets for the design of drugs to, for example, disrupt the PDU BMC of *Salmonella typhimurium* or the EUT BMC of enterohemorrhagic *E. coli*. Therefore, it is important when studying the role of these BMCs to consider not only their substrates and products, but also how BMC metabolism integrates with other cellular metabolic processes within the pathogen and how the metabolites interact with the host.

Bacterial microcompartments in bioengineering

BMCs as plug-and-play metabolic modules

The presence of similar BMC loci in distantly related bacteria indicates that BMC gene clusters are frequently horizontally transferred[6, 116,117]. This has likely been facilitated by their genetic modularity. This compact organization and the advent of inexpensive, large scale DNA synthesis methods make it feasible to introduce BMCs as metabolic modules across species[118, 119]. In an early proof-of-concept experiment, the *pdu* operon of *Citrobacter freundii* was introduced into *E. coli*, enabling the resulting mutant strain to grow on 1,2-propanediol[120]. Carboxysomes have also been the subject of extensive bioengineering efforts to enable or enhance CO_2 fixation. The *cso* operon encoding the α carboxysome of the chemoautotroph *Halothiobacillus neapolitanus* has been expressed in *E.*

coli[121] and *Corynebacterium glutamicum*[122], and the observed carboxysome-like particles were shown to have CO_2 fixing RuBisCO activity. The genes that encode for β-carboxysomal shell and core proteins have been expressed in the chloroplasts of *Nicotiana benthamiana*, resulting in the generation of carboxysome-like aggregates[123]. Furthermore, a native RuBisCO-deletion tobacco plant line was able to grow when it was engineered so that a cyanobacterial RuBisCO could be produced in its chloroplasts; the addition of the gene that encodes for the SSLDs of CcmM led to the formation of a structure resembling a pro-carboxysome[124]. This first inter-kingdom heterologous expression of BMC components represents a leap towards the objective of utilizing carboxysomes in plants for enhanced photosynthesis in crops[25,62, 119,123–126] and, more generally, for building bacterial organelles within eukaryotic cells.

Designing new BMC cores

Designing and building BMC cores for new functions is also under intense investigation. Understanding the interactions that occur among protein domains during β-carboxysome assembly enabled the design of a chimeric protein, CcmC, which comprises SSLDs, an encapsulation peptide and a β-carbonic anhydrase — the essential elements for pro-carboxysome formation, shell recruitment and concentration of CO_2 around RuBisCO[127]. CcmC was able to promote β-carboxysome assembly in a carboxysome-mutant strain and the resulting carboxysomes were functional, allowing the cyanobacterium (*Synechococcus elongatus* PCC7942) to grow in air at a comparable rateto the wild type strain[127]. CcmC, by replacing four gene products, simplifies carboxysome assembly by reducing the genetic load on a potential host and reduces the regulatory complexity; this proof-of-concept of domain-based core engineering should facilitate new designs and the introduction of BMCs to new species. It is also possible to target enzymes or proteins of interest into BMC shells by fusing them to encapsulation peptides[40,48,63,67,128–131]. Encapsulating new functions using this strategy was demonstrated by the construction of a synthetic pyruvate-to-ethanol BMC[65] and a polyphosphate accumulation BMC[132], both using recombinant PDU shells in *E. coli*. Additionally, the nucleation role of encapsulation peptides can be used to improve the efficiency of enzymatic pathway reactions by co-localizing enzymes of a multi-step reaction, even in the absence of a shell. Recently, a large enzyme bolus occupying half of the bacterial cell was obtained by fusing two different encapsulation peptides to four enzymes (glycerol dehydrogenase, dihydroxyacetone kinase, methylglyoxal synthase and 1,2-propanediol oxidoreductase) to convert glycerol to 1,2-propanediol[133]. This aggregate was shown to be 2.5 times more efficient in 1,2-propanediol production than enzymes that lacked encapsulation peptides, presumably due to a higher local enzyme concentration and the substrate channeling within the protein aggregate. These properties are consistent with the emerging understanding that encapsulation peptides are also involved in core assembly and not only in interactions with the shell[49, 50].

Fusing encapsulation peptides to individual proteins is a useful method for constructing simple organelles that are composed of one or two proteins but might not be suitable for complex pathways that require many enzymes. There is typically only one or two encapsulation peptides associated with each BMC type, and they may compete for the same shell protein binding partner[130]. The low encapsulation efficiency of non-native cargo that

has been observed in several studies[40, 63,67,134,135] may also be due to the absence of additional protein domain interactions that contribute to the robustness of encapsulation. Several strategies have been explored to overcome these limitations. For example, encapsulation efficiency can be improved by delayed expression of the cargo relative to the shell proteins[136] and by altering the primary structure of encapsulation peptides[137]. To identify the primary structure features that are important for encapsulation efficiency, encapsulation peptides were designed *de novo*[134] and reporter proteins that were fused to encapsulation peptides from GRM and EUT BMCs were encapsulated in a PDU BMC[138], demonstrating the potential for using encapsulation peptides to attach proteins to diverse BMC shells.

In order to target several different proteins to the BMC lumen, encapsulation peptides can be fused to proteins on either their N-terminus or the C-terminus, or within loop regions between domains to generate intra-protein encapsulation peptides. The latter type, the internal encapsulation peptide, has been recently identified in native BMC systems[8]; locating an encapsulation peptide internally will likely be useful for when modifying the N-terminus or the C-terminus of a protein to be encapsulated impacts its function. Likewise, the ability to internally introduce encapsulation peptides increases the range of designs that are possible for constructing chimeric proteins. The same strategy that is based on encapsulation peptides could also be used to target synthetic scaffolding proteins[139] to the lumen of BMCs (Figure 5). This scaffold and the enzymes of the synthetic pathway would be engineered to specifically interact with each other, enabling the precise control of protein stoichiometry[25].

By viewing core proteins as a collection of folds and taking advantage of knowledge of the biochemistry of native systems, it may be possible to engineer the active site of the signature enzyme and other core enzymes of a given BMC to act on different substrates. For example, BMC AldDHs are capable of utilizing different short chain aldehydes[76, 140]. Accordingly, a EUT AldDH was able to functionally replace the native enzyme of a PDU BMC[138]. This strategy could be extended to any core enzyme, by replacing one of the native enzymes with a catalytically distinct member with the same structural fold [7, 25, 118]. Surface features (for example, size and charge of specific amino acids) of the homologous enzyme could be altered to resemble those known to facilitate protein-protein interactions in the native core. Likewise, encapsulation peptides could be grafted onto the homolog protein to facilitate encapsulation[141–143] (Figure 5).

Engineering shell proteins

The size and charge of the pores of the different BMC shell proteins define the type of compounds that can traverse the shell[7,8,18,144]. Mutations of the residues that flank the pores of shell proteins from the PDU metabolosome[32, 33] and carboxysomes[145] showed that altering the size and charge of the pore to influence permeability does not impair shell formation. This provides the basis for the further rational design of channels that are specific to desired metabolites that are not native to natural BMCs; this is a key design requirement as the shell serves as the interface between the lumen of the BMC and its cellular environment.

Indications of the presence of Fe-S clusters **[G]** in BMC-H[8, 146] and BMC-T[120, 135, 147] proteins suggest that electron transfer across the shell occurs in some types of BMCs. In developing synthetic BMCs for bioengineering applications, the incorporation of electron relays through shell proteins would greatly expand the possible repertoire of potential core chemistries. For example, this would enable the encapsulation of redox reactions that can be connected to a source of electrons in the cytosol, or conversely, to an electron sink that is located outside of the compartment. Recently, a BMC-T^s shell protein was engineered to bind a redox active 4Fe-4S cluster in its pore, which was stable through redox cycling and exhibited properties similar to the clusters of low-potential bacterial ferredoxins **[G]**[148]. This was also the first example of conferring a completely new function to a shell protein. Engineering different shell proteins with diverse metal-cofactors and fine-tuning their redox potential allows for the conversion of the passive BMC shell barrier into an active interface.

Strong sequence and structural conservation among shell proteins from functionally distinct BMCs enables mixtures of shell proteins from different types of BMCs to assemble together. For example, chimeric shells were obtained by replacing the major shell protein CcmK2 of β-carboxysomes with the CsoS1 protein of α-carboxysomes[145], and EutM was successfully integrated into a PDU BMC[33]. Likewise, chimeric shells were produced by the simultaneous expression of the *pdu* and *eut* operons[149], and a β-carboxysome shell protein (CcmO) has been incorporated in the PDU BMC[149], highlighting the potential use of BMC shell proteins as generic structural building blocks (Figure 5).

Shell proteins for novel bioarchitectures

When shell proteins are expressed outside of their native cellular context they form a diverse array of aberrant structures, including filaments[63,120,150], rosettes[19,107,120] and nanotubes[23,151,152]. These architectures provide an alternative use for shell proteins in the construction of modular scaffolds that could be used to enhance metabolic efficiency *in vivo* and *in vitro*[153].

The re-purposing of shell proteins to form programmable scaffolds is predicated on understanding the specific structural determinants that underlie the interactions among the proteins forming the facets of BMC shells, and how their expression levels influence the resulting architecture. New insights have emerged from the study of two non-natural BMC architectures: nanotubes and two-dimensional layers. Nanotubes can be formed by the expression of a single BMC-H proteins, PduA[151] or RmmH from *Mycobacterium smegmatis*[152]. For the latter, nanotube assembly is reversible and regulated in a concentration-dependent manner[152]. Likewise, after their purification, the rosettes formed *in vivo* by the BMC-H protein of *H. ochraceum* roll-out into flat, single hexamer-thick sheets in which hexamers are uniformly oriented[19, 40]. In addition, computational simulations were used to study the self-assembly of the shell protein CcmK2 *in silico*[154].

To further expand the design space of BMC-based architectures, two-dimensional chimeric scaffolds could be constructed from shell proteins from diverse BMC types, and functionalized with enzymes using strategies developed for three-dimensional compartments (Figure 5). Shell proteins engineered to form empty shells or cages[155,156] could also be

repurposed as cargo carriers for the delivery of molecules, with potential application in human health and phytonanotechnology[157,158].

Conclusions

In the last 10 years, the understanding of the BMCs has greatly increased (TIMELINE). Progress has been made in characterizing their structures, assembly mechanisms and encapsulated reactions. Although the number and types of BMCs encoded in genomes across the Bacterial Kingdom continues to grow, there are only six with experimentally confirmed functions, namely α- and β-carboxysomes and the EUT, PDU, GRM and PVM metabolosomes[11–14, 57, 159]. In addition to experimental characterization of the diversity of their functions, many fundamental questions remain (Box 1). The plasticity of the BMC architecture provides the foundation for diverse applications in synthetic biology, including their use in programming the metabolism of members of microbial consortia (synthetic ecology). It seems likely that the study and engineering of bacterial organelles will become a field comparable in size to those focused on their eukaryotic counterparts.

Box 1

Outstanding questions and major challenges in the field

Questions

Are there shared strategies for organizing core enzymes for substrate and product channeling within bacterial microcompartments (BMCs)s?

How dynamic are BMCs once assembled? Can damaged core proteins be repaired —or replaced? Can the shell composition be altered to tune permeability? Do BMCs change their size in response to environmental conditions?

Are the core enzymes active prior to encapsulation? How is complete encapsulation of enzymes that produce toxic intermediates ensured?

Why do some BMC loci contain multiple genes encoding BMC-P (pentamer forming) proteins, given they are assumed to be only a minor component of the shell?

Do any of the BMC shell proteins assemble into hetero-oligomers?

How is the internal environment (pH, oxygen tension) established?

How do the organelles interact with the cytoskeleton? Is there a specific subcellular site for initiation of BMC formation?

Challenges

A systematic characterization of the permeability of BMC shells

Developing a systems-level understanding of BMC function that includes the role of proteins encoded in the superloci

Development *in vitro* assembly triggers to facilitate characterization of designed BMCs

Discover reagents that precisely control number and location of proteins to be encapsulated in shells

Acknowledgments

This work was supported by the National Institutes of Health, National Institute of Allergy and Infectious Diseases (NIAID) grant 1R01AI114975-01 with infrastructure support from the U.S. Department of Energy, Basic Energy Sciences, Contract DE-FG02-91ER20021.

Glossary

Carbonic anhydrase
Enzyme that catalyzes conversion of bicarbonate to carbon dioxide (and vice versa); the several subclasses of CAs have distinct structural folds

RuBisCO (ribulose-1,5-bisphosphate carboxylase/oxygenase)
An enzyme that fixes carbon dioxide by reacting it with ribulose 1–5-bisphosphate to create two molecules of phosphoglycerate. Form 1 RuBisCO is composed of eight small and eight large subunits

Calvin-Benson-Bassham Cycle
A pathway for producing phosphoglyceraldehyde from CO2

Alpha carboxysomes
One of two types of carboxysome; alpha carboxysomes encapsulate Form1A RuBisCO and are found primarily in marine cyanobacteria and chemoautotrophs

Beta carboxysome
The type of carboxysome found in ecophysiologically diverse cyanobacteria, it encapsulates Form 1B RuBisCO, the Form found in higher plants

Pfam 00936 domain
An ~90 amino acid sequence that folds into an alpha beta structure that oligomerizes into a hexamer

BMC-H
A type of shell protein containing a single Pfam00936 domain that form cyclic homohexamers

BMC-T
A shell protein that contain two Pfam00936 domains that form cyclic homodimers or pseudohexamers

BMC-P
A type of protein containing the Pfam03319 domain that forming the homopentamers and function as the pentagonal vertices of the BMC shell

Encapsulation peptide

One or more short (~17 amino acid) amphipathic helix that targets cargo proteins to the interior of BMCs; they are typically located at the N- or C-terminus of a protein and connected by an unstructured linker

γ-carbonic anhydrase
A subclass of carbonic anhydrases with a characteristic structure of three chains that each form a left handed β helix and a metal ion active site

Glycyl radical enzymes (GREs)
A class of enzymes that uses radicals of glycine and cysteine for catalysis; they are highly oxygen sensitive and require an activating enzyme containing an iron-sulfur cluster to generate the glycyl radical

Superlocus
Region on the chromosome that contains one or more operons encoding genes for a bacterial microcompartment and ancillary proteins that support the function of the organelle

Signature enzyme
An enzyme of a metabolosome that is specific to the initial substrate of the BMC

Maximum-likelihood tree
A phylogenetic tree constructed using a computationally intense method that searches for the tree that has the highest probability of producing the observed data

Pathogenicity island
segments of chromosomes that encode virulence factors and are found in pathogenic microbes but absent in closely related, non-pathogenic strains

Iron sulfur clusters
Metal clusters composed of non-heme iron and sulfur atoms; in proteins they function to transfer electrons over a wide range of potentials

Ferredoxins
A family of proteins that contain iron-sulfur clusters to mediate electron transfer

References

1. Kerfeld CA, Heinhorst S, Cannon GC. Bacterial microcompartments. Annu Rev Microbiol. 2010; 64:391–408. [PubMed: 20825353]
2. Kerfeld CA, Melnicki MR. Assembly, function and evolution of cyanobacterial carboxysomes. Curr Opin Plant Biol. 2016; 31:66–75. [PubMed: 27060669]
3. Rae BD, Long BM, Badger MR, Price GD. Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO_2 fixation in cyanobacteria and some proteobacteria. Microbiol Mol Biol Rev. 2013; 77:357–79. [PubMed: 24006469]
4. Turmo A, Gonzalez-Esquer CR, Kerfeld CA. Carboxysomes: metabolic modules for CO_2 fixation. FEMS Microbiol Lett. 2017; 364
5. Chowdhury C, Sinha S, Chun S, Yeates TO, Bobik TA. Diverse bacterial microcompartment organelles. Microbiol Mol Biol Rev. 2014; 78:438–68. [PubMed: 25184561]

6. Axen SD, Erbilgin O, Kerfeld CA. A taxonomy of bacterial microcompartment loci constructed by a novel scoring method. PLoS Comput Biol. 2014; 10:e1003898. Comprehensive survey of BMC types and subtypes by bioinformatic analysis of superloci. [PubMed: 25340524]
7. Kerfeld CA, Erbilgin O. Bacterial microcompartments and the modular construction of microbial metabolism. Trends Microbiol. 2015; 23:22–34. [PubMed: 25455419]
8. Zarzycki J, Erbilgin O, Kerfeld CA. Bioinformatic characterization of glycyl radical enzyme-associated bacterial microcompartments. Appl Environ Microbiol. 2015; 81:8315–29. [PubMed: 26407889]
9. Jakobson CM, Tullman-Ercek D, Slininger MF, Mangan NM. A systems-level model reveals that 1,2-Propanediol utilization microcompartments enhance pathway flux through intermediate sequestration. Plos Computational Biology. 2017; 13
10. Cannon GC, Heinhorst S, Kerfeld CA. Carboxysomal carbonic anhydrases: Structure and role in microbial CO_2 fixation. Biochim Biophys Acta. 2010; 1804:382–92. [PubMed: 19818881]
11. Bobik TA, Havemann GD, Busch RJ, Williams DS, Aldrich HC. The propanediol utilization (pdu) operon of *Salmonella enterica* serovar Typhimurium LT2 includes genes necessary for formation of polyhedral organelles involved in coenzyme B12-dependent 1,2-propanediol degradation. J Bacteriol. 1999; 181:5967–5975. [PubMed: 10498708]
12. 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. 1999; 181:5317–29. [PubMed: 10464203]
13. Petit E, et al. Involvement of a bacterial microcompartment in the metabolism of fucose and rhamnose by *Clostridium phytofermentans*. PLoS One. 2013; 8:e54337. First experimental characterization of a GRM organelle. [PubMed: 23382892]
14. Erbilgin O, McDonald KL, Kerfeld CA. Characterization of a planctomycetal organelle: a novel bacterial microcompartment for the aerobic degradation of plant saccharides. Appl Environ Microbiol. 2014; 80:2193–205. Functional characterization of a PVM BMC involved in the degradation of fucose and rhamnose. [PubMed: 24487526]
15. Roof DM, Roth JR. Autogenous regulation of ethanolamine utilization by a transcriptional activator of the eut operon in Salmonella typhimurium. J Bacteriol. 1992; 174:6634–43. [PubMed: 1328159]
16. Bobik TA, Ailion M, Roth JR. A single regulatory gene integrates control of vitamin B12 synthesis and propanediol degradation. J Bacteriol. 1992; 174:2253–66. [PubMed: 1312999]
17. Jakobson CM, Tullman-Ercek D. Dumpster Diving in the Gut: Bacterial Microcompartments as Part of a Host-Associated Lifestyle. PLoS Pathog. 2016; 12:e1005558. [PubMed: 27171216]
18. Yeates TO, Kerfeld CA, Heinhorst S, Cannon GC, Shively JM. Protein-based organelles in bacteria: carboxysomes and related microcompartments. Nat Rev Microbiol. 2008; 6:681–91. [PubMed: 18679172]
19. Sutter M, et al. Visualization of Bacterial Microcompartment Facet Assembly Using High-Speed Atomic Force Microscopy. Nano Lett. 2016; 16:1590–5. [PubMed: 26617073]
20. Sutter M, Greber B, Aussignargues C, Kerfeld CA. Assembly principles and structure of a 6.5-MDa bacterial microcompartment shell. Science. 2017; 356:1293–1297. Structure of a complete BMC shell reveals specific interactions among shell proteins. [PubMed: 28642439]
21. Savage DF, Afonso B, Chen AH, Silver PA. Spatially ordered dynamics of the bacterial carbon fixation machinery. Science. 2010; 327:1258–61. Shows alignment of β-carboxysomes along the long axis of *S. elongatus* by fluorescent labeling and correlates this to interaction with the cytoskeleton. [PubMed: 20203050]
22. Menon BB, Heinhorst S, Shively JM, Cannon GC. The carboxysome shell is permeable to protons. J Bacteriol. 2010; 192:5881–6. [PubMed: 20870775]
23. Parsons JB, et al. Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. Mol Cell. 2010; 38:305–15. [PubMed: 20417607]
24. Frank S, Lawrence AD, Prentice MB, Warren MJ. Bacterial microcompartments moving into a synthetic biological world. Journal of Biotechnology. 2013; 163:273–279. [PubMed: 22982517]

25. Gonzalez-Esquer CR, Newnham SE, Kerfeld CA. Bacterial microcompartments as metabolic modules for plant synthetic biology. Plant J. 2016; 87:66–75. [PubMed: 26991644]
26. Plegaria JS, Kerfeld CA. Engineering nanoreactors using bacterial microcompartment architectures. Curr Opin Biotechnol. 2017; 51:1–7. [PubMed: 29035760]
27. Kerfeld CA, et al. Protein structures forming the shell of primitive bacterial organelles. Science. 2005; 309:936–8. [PubMed: 16081736]
28. Cai F, et al. The Structure of CcmP, a Tandem Bacterial Microcompartment Domain Protein from the beta-Carboxysome, Forms a Subcompartment Within a Microcompartment. Journal of Biological Chemistry. 2013; 288:16055–16063. [PubMed: 23572529]
29. Klein MG, et al. Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport. J Mol Biol. 2009; 392:319–33. Reports the first structure of a double stacking BMC-T protein and a possible pore gating mechanism. [PubMed: 19328811]
30. Mallette E, Kimber MS. A Complete Structural Inventory of the Mycobacterial Microcompartment Shell Proteins Constrains Models of Global Architecture and Transport. J Biol Chem. 2017; 292:1197–1210. [PubMed: 27927988]
31. Larsson AM, Hasse D, Valegard K, Andersson I. Crystal structures of beta-carboxysome shell protein CcmP: ligand binding correlates with the closed or open central pore. J Exp Bot. 2017
32. Chowdhury C, et al. Selective molecular transport through the protein shell of a bacterial microcompartment organelle. Proc Natl Acad Sci U S A. 2015; 112:2990–5. [PubMed: 25713376]
33. Slininger Lee MF, Jakobson CM, Tullman-Ercek D. Evidence for Improved Encapsulated Pathway Behavior in a Bacterial Microcompartment through Shell Protein Engineering. ACS Synthetic Biology. 2017
34. Tsai Y, et al. Structural analysis of CsoS1A and the protein shell of the *Halothiobacillus neapolitanus* carboxysome. PLoS Biol. 2007; 5:e144. [PubMed: 17518518]
35. Tanaka S, Sawaya MR, Yeates TO. Structure and mechanisms of a protein-based organelle in *Escherichia coli*. Science. 2010; 327:81–4. [PubMed: 20044574]
36. Iancu CV, et al. The structure of isolated *Synechococcus* strain WH8102 carboxysomes as revealed by electron cryotomography. J Mol Biol. 2007; 372:764–73. [PubMed: 17669419]
37. Tanaka S, et al. Atomic-level models of the bacterial carboxysome shell. Science. 2008; 319:1083–6. [PubMed: 18292340]
38. Sutter M, Wilson SC, Deutsch S, Kerfeld CA. Two new high-resolution crystal structures of carboxysome pentamer proteins reveal high structural conservation of CcmL orthologs among distantly related cyanobacterial species. Photosynth Res. 2013; 118:9–16. [PubMed: 23949415]
39. Cai F, et al. The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO_2 leakage barrier. PLoS One. 2009; 4:e7521. Provides evidence that the pentameric BMC-P proteins are essential for formation of an intact carboxysome shell. [PubMed: 19844578]
40. Lassila JK, Bernstein SL, Kinney JN, Axen SD, Kerfeld CA. Assembly of robust bacterial microcompartment shells using building blocks from an organelle of unknown function. J Mol Biol. 2014; 426:2217–28. [PubMed: 24631000]
41. Liberton M, Austin JR 2nd, Berg RH, Pakrasi HB. Unique thylakoid membrane architecture of a unicellular N2-fixing cyanobacterium revealed by electron tomography. Plant Physiol. 2011; 155:1656–66. [PubMed: 21173021]
42. Sagermann M, Ohtaki A, Nikolakakis K. Crystal structure of the EutL shell protein of the ethanolamine ammonia lyase microcompartment. Proc Natl Acad Sci U S A. 2009; 106:8883–7. Describes the structure of a single layered BMC-T protein with three smaller pores around the central axis of symmetry. [PubMed: 19451619]
43. Pang A, Liang M, Prentice MB, Pickersgill RW. Substrate channels revealed in the trimeric *Lactobacillus reuteri* bacterial microcompartment shell protein PduB. Acta Crystallogr D Biol Crystallogr. 2012; 68:1642–52. [PubMed: 23151629]
44. Thompson MC, Cascio D, Leibly DJ, Yeates TO. An allosteric model for control of pore opening by substrate binding in the EutL microcompartment shell protein. Protein Sci. 2015; 24:956–75. [PubMed: 25752492]

45. Samborska B, Kimber MS. A Dodecameric CcmK2 Structure Suggests beta-Carboxysomal Shell Facets Have a Double-Layered Organization. Structure. 2012; 20:1353–1362. [PubMed: 22748766]
46. Faulkner M, et al. Direct characterization of the native structure and mechanics of cyanobacterial carboxysomes. Nanoscale. 2017; 9:10662–10673. Purification and structural characterization of β-carboxysomes. [PubMed: 28616951]
47. Kinney JN, Salmeen A, Cai F, Kerfeld CA. Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly. J Biol Chem. 2012; 287:17729–36. Identifies a peptide extensions for targeting in β-carboxysomes and predicts counterparts in diverse BMCs. [PubMed: 22461622]
48. Fan C, et al. Short N-terminal sequences package proteins into bacterial microcompartments. Proc Natl Acad Sci U S A. 2010; 107:7509–14. [PubMed: 20308536]
49. Aussignargues C, Paasch BC, Gonzalez-Esquer R, Erbilgin O, Kerfeld CA. Bacterial microcompartment assembly: The key role of encapsulation peptides. Commun Integr Biol. 2015; 8:e1039755. [PubMed: 26478774]
50. Erbilgin O, Sutter M, Kerfeld CA. The Structural Basis of Coenzyme A Recycling in a Bacterial Organelle. PLoS Biol. 2016; 14:e1002399. [PubMed: 26959993]
51. Cameron JC, Wilson SC, Bernstein SL, Kerfeld CA. Biogenesis of a bacterial organelle: the carboxysome assembly pathway. Cell. 2013; 155:1131–40. Describes the assembly pathway of β-carboxysomes using knockout mutants and fluorescently labeled reporters. [PubMed: 24267892]
52. Ludwig M, Sultemeyer D, Price GD. Isolation of ccmKLMN genes from the marine cyanobacterium, *Synechococcus* sp PCC7002 (Cyanobacteria), and evidence that CcmM is essential for carboxysome assembly. Journal of Phycology. 2000; 36:1109–1118.
53. Emlyn-Jones D, Woodger FJ, Andrews TJ, Price GD, Whitney SM. A *Synechococcus* PCC7942 ΔccmM (Cyanophyceae) Mutant Pseudoreverts to Air Growth without Regaining Carboxysomes. J Phycol. 2006; 42:769–777.
54. Woodger FJ, Badger MR, Price GD. Sensing of inorganic carbon limitation in *Synechococcus* PCC7942 is correlated with the size of the internal inorganic carbon pool and involves oxygen. Plant Physiology. 2005; 139:1959–1969. [PubMed: 16306144]
55. Berry S, Fischer JH, Kruip J, Hauser M, Wildner GF. Monitoring cytosolic pH of carboxysome-deficient cells of *Synechocystis* sp. PCC 6803 using fluorescence analysis. Plant Biol (Stuttg). 2005; 7:342–7. [PubMed: 16025406]
56. Pena KL, Castel SE, de Araujo C, Espie GS, Kimber MS. Structural basis of the oxidative activation of the carboxysomal gamma-carbonic anhydrase, CcmM. Proc Natl Acad Sci U S A. 2010; 107:2455–60. [PubMed: 20133749]
57. Price GD, Howitt SM, Harrison K, Badger MR. Analysis of a genomic DNA region from the cyanobacterium *Synechococcus* sp. strain PCC7942 involved in carboxysome assembly and function. J Bacteriol. 1993; 175:2871–9. [PubMed: 8491708]
58. Long BM, Badger MR, Whitney SM, Price GD. Analysis of carboxysomes from *Synechococcus* PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. J Biol Chem. 2007; 282:29323–35. [PubMed: 17675289]
59. Long BM, Tucker L, Badger MR, Price GD. Functional Cyanobacterial β-Carboxysomes Have an Absolute Requirement for Both Long and Short Forms of the CcmM Protein. Plant Physiol. 2010; 153:285–293. [PubMed: 20304968]
60. Cot SS, So AK, Espie GS. A multiprotein bicarbonate dehydration complex essential to carboxysome function in cyanobacteria. J Bacteriol. 2008; 190:936–45. [PubMed: 17993516]
61. McGurn LD, et al. The structure, kinetics and interactions of the beta-carboxysomal beta-carbonic anhydrase, CcaA. Biochem J. 2016; 473:4559–4572. [PubMed: 27729545]
62. Zarzycki J, Axen SD, Kinney JN, Kerfeld CA. Cyanobacterial-based approaches to improving photosynthesis in plants. Journal of Experimental Botany. 2013; 64:787–798. [PubMed: 23095996]
63. Choudhary S, Quin MB, Sanders MA, Johnson ET, Schmidt-Dannert C. Engineered protein nano-compartments for targeted enzyme localization. PLoS One. 2012; 7:e33342. [PubMed: 22428024]

64. Fan C, Bobik TA. The N-terminal region of the medium subunit (PduD) packages adenosylcobalamin-dependent diol dehydratase (PduCDE) into the Pdu microcompartment. J Bacteriol. 2011; 193:5623–8. Describes encapsulation peptides in the Pdu system and shows targeting of foreign proteins. [PubMed: 21821773]
65. Lawrence AD, et al. Solution structure of a bacterial microcompartment targeting peptide and its application in the construction of an ethanol bioreactor. ACS Synth Biol. 2014; 3:454–65. [PubMed: 24933391]
66. Chen AH, Robinson-Mosher A, Savage DF, Silver PA, Polka JK. The Bacterial Carbon-Fixing Organelle Is Formed by Shell Envelopment of Preassembled Cargo. Plos One. 2013; 8
67. Cai F, Bernstein SL, Wilson SC, Kerfeld CA. Production and Characterization of Synthetic Carboxysome Shells with Incorporated Luminal Proteins. Plant Physiol. 2016; 170:1868–77. [PubMed: 26792123]
68. Iancu CV, et al. Organization, structure, and assembly of alpha-carboxysomes determined by electron cryotomography of intact cells. J Mol Biol. 2010; 396:105–17. [PubMed: 19925807]
69. Cai F, et al. Advances in Understanding Carboxysome Assembly in *Prochlorococcus* and *Synechococcus* Implicate CsoS2 as a Critical Component. Life. 2015; 5:1141–1171. Characterization of the protein responsible for organization of the interior of α-carboxysomes. [PubMed: 25826651]
70. Chaijarasphong T, et al. Programmed Ribosomal Frameshifting Mediates Expression of the α-Carboxysome. J Mol Biol. 2016; 428:153–64. [PubMed: 26608811]
71. Perlmutter JD, Mohajerani F, Hagan MF. Many-molecule encapsulation by an icosahedral shell. Elife. 2016; 5
72. Fan C, Cheng S, Sinha S, Bobik TA. Interactions between the termini of lumen enzymes and shell proteins mediate enzyme encapsulation into bacterial microcompartments. Proc Natl Acad Sci U S A. 2012; 109:14995–5000. [PubMed: 22927404]
73. Tobimatsu T, Kawata M, Toraya T. The N-terminal regions of beta and gamma subunits lower the solubility of adenosylcobalamin-dependent diol dehydratase. Biosci Biotechnol Biochem. 2005; 69:455–62. [PubMed: 15784971]
74. Shibata N, et al. Crystal structures of ethanolamine ammonia-lyase complexed with coenzyme B_{12} analogs and substrates. J Biol Chem. 2010; 285:26484–93. [PubMed: 20519496]
75. Akita K, et al. Purification and some properties of wild-type and N-terminal-truncated ethanolamine ammonia-lyase of *Escherichia coli*. J Biochem. 2010; 147:83–93. [PubMed: 19762342]
76. Zarzycki J, Sutter M, Cortina NS, Erb TJ, Kerfeld CA. *In Vitro* Characterization and Concerted Function of Three Core Enzymes of a Glycyl Radical Enzyme - Associated Bacterial Microcompartment. Sci Rep. 2017; 7:42757. [PubMed: 28202954]
77. Havemann GD, Sampson EM, Bobik TA. 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. 2002; 184:1253–1261. [PubMed: 11844753]
78. Cheng S, Sinha S, Fan C, Liu Y, Bobik TA. Genetic analysis of the protein shell of the microcompartments involved in coenzyme B_{12}-dependent 1,2-propanediol degradation by *Salmonella*. J Bacteriol. 2011; 193:1385–92. [PubMed: 21239588]
79. Lehman BP, Chowdhury C, Bobik TA. The N-terminus of the PduB protein binds the protein shell of the Pdu microcompartment to its enzymatic core. Journal of Bacteriology. 2017
80. Cheng S, Fan C, Sinha S, Bobik TA. The PduQ enzyme is an alcohol dehydrogenase used to recycle NAD+ internally within the Pdu microcompartment of *Salmonella enterica*. PLoS One. 2012; 7:e47144. Identifies protein responsible for NADH cofactor recycling in the PDU BMC. [PubMed: 23077559]
81. Chen P, Andersson DI, Roth JR. The control region of the *pdu/cob* regulon in *Salmonella typhimurium*. J Bacteriol. 1994; 176:5474–82. [PubMed: 8071226]
82. Stojiljkovic I, Baumler AJ, 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. 1995; 177:1357–66. [PubMed: 7868611]

83. Roberts EW, Cai F, Kerfeld CA, Cannon GC, Heinhorst S. Isolation and characterization of the *Prochlorococcus* carboxysome reveal the presence of the novel shell protein CsoS1D. J Bacteriol. 2012; 194:787–95. [PubMed: 22155772]

84. Sutter M, et al. Structural Characterization of a Newly Identified Component of alpha-Carboxysomes: The AAA+ Domain Protein CsoCbbQ. Sci Rep. 2015; 5:16243. [PubMed: 26538283]

85. Liu Y, Jorda J, Yeates TO, Bobik TA. The PduL Phosphotransacylase Is Used To Recycle Coenzyme A within the Pdu Microcompartment. J Bacteriol. 2015; 197:2392–9. [PubMed: 25962918]

86. Huseby DL, Roth JR. Evidence that a metabolic microcompartment contains and recycles private cofactor pools. J Bacteriol. 2013; 195:2864–79. [PubMed: 23585538]

87. Craciun S, Balskus EP. Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. Proc Natl Acad Sci U S A. 2012; 109:21307–12. [PubMed: 23151509]

88. Craciun S, Marks JA, Balskus EP. Characterization of choline trimethylamine-lyase expands the chemistry of glycyl radical enzymes. ACS Chem Biol. 2014; 9:1408–13. [PubMed: 24854437]

89. Kalnins G, et al. Structure and Function of CutC Choline Lyase from Human Microbiota Bacterium *Klebsiella pneumoniae*. J Biol Chem. 2015; 290:21732–40. [PubMed: 26187464]

90. LaMattina JW, et al. 1,2-Propanediol Dehydration in *Roseburia inulinivorans*: STRUCTURAL BASIS FOR SUBSTRATE AND ENANTIOMER SELECTIVITY. J Biol Chem. 2016; 291:15515–26. [PubMed: 27252380]

91. Urano N, et al. Genetic analysis around aminoalcohol dehydrogenase gene of *Rhodococcus erythropolis* MAK154: a putative GntR transcription factor in transcriptional regulation. Appl Microbiol Biotechnol. 2011; 89:739–46. [PubMed: 20953603]

92. Kataoka M, et al. A novel NADP+-dependent L-1-amino-2-propanol dehydrogenase from *Rhodococcus erythropolis* MAK154: a promising enzyme for the production of double chiral aminoalcohols. Lett Appl Microbiol. 2006; 43:430–5. [PubMed: 16965375]

93. Land M, et al. Insights from 20 years of bacterial genome sequencing. Funct Integr Genomics. 2015; 15:141–61. [PubMed: 25722247]

94. Nobu MK, et al. Phylogeny and physiology of candidate phylum 'Atribacteria' (OP9/JS1) inferred from cultivation-independent genomics. ISME J. 2016; 10:273–86. [PubMed: 26090992]

95. Biddle AS, et al. The complete genome sequence of *Clostridium indolis* DSM 755(T.). Stand Genomic Sci. 2014; 9:1089–104. [PubMed: 25197485]

96. Lawhon SD, et al. Global regulation by CsrA in *Salmonella typhimurium*. Mol Microbiol. 2003; 48:1633–45. [PubMed: 12791144]

97. Andersson RE. Biogenic-Amines in Lactic Acid-Fermented Vegetables. Lebensmittel-Wissenschaft & Technologie. 1988; 21:68–69.

98. Collier PD, Cromie DDO, Davies AP. Mechanism of Formation of Chloropropanols Present in Protein Hydrolysates. Journal of the American Oil Chemists Society. 1991; 68:785–790.

99. Harvey PC, et al. Salmonella enterica serovar typhimurium colonizing the lumen of the chicken intestine grows slowly and upregulates a unique set of virulence and metabolism genes. Infect Immun. 2011; 79:4105–21. [PubMed: 21768276]

100. Klumpp J, Fuchs TM. Identification of novel genes in genomic islands that contribute to *Salmonella typhimurium* replication in macrophages. Microbiology. 2007; 153:1207–20. [PubMed: 17379730]

101. Joseph B, et al. Identification of *Listeria monocytogenes* genes contributing to intracellular replication by expression profiling and mutant screening. J Bacteriol. 2006; 188:556–68. [PubMed: 16385046]

102. Maadani A, Fox KA, Mylonakis E, Garsin DA. *Enterococcus faecalis* mutations affecting virulence in the *Caenorhabditis elegans* model host. Infect Immun. 2007; 75:2634–7. [PubMed: 17307944]

103. Bertin Y, et al. Enterohaemorrhagic *Escherichia coli* gains a competitive advantage by using ethanolamine as a nitrogen source in the bovine intestinal content. Environ Microbiol. 2011; 13:365–77. [PubMed: 20849446]

104. Srikumar S, Fuchs TM. Ethanolamine utilization contributes to proliferation of *Salmonella enterica* serovar Typhimurium in food and in nematodes. Appl Environ Microbiol. 2011; 77:281–90. [PubMed: 21037291]
105. Thiennimitr P, et al. Intestinal inflammation allows *Salmonella* to use ethanolamine to compete with the microbiota. Proc Natl Acad Sci U S A. 2011; 108:17480–5. [PubMed: 21969563]
106. Kendall MM, Gruber CC, Parker CT, Sperandio V. Ethanolamine controls expression of genes encoding components involved in interkingdom signaling and virulence in enterohemorrhagic *Escherichia coli* O157:H7. MBio. 3:2012.
107. Pitts AC, Tuck LR, Faulds-Pain A, Lewis RJ, Marles-Wright J. Structural Insight into the *Clostridium difficile* Ethanolamine Utilisation Microcompartment. Plos One. 2012; 7
108. Martinez-del Campo A, et al. Characterization and detection of a widely distributed gene cluster that predicts anaerobic choline utilization by human gut bacteria. MBio. 2015; 6
109. Wang Z, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature. 2011; 472:57–63. [PubMed: 21475195]
110. Lang DH, et al. Isoform specificity of trimethylamine N-oxygenation by human flavin-containing monooxygenase (FMO) and P450 enzymes: selective catalysis by FMO3. Biochem Pharmacol. 1998; 56:1005–12. [PubMed: 9776311]
111. Koeth RA, et al. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. Nat Med. 2013; 19:576–85. [PubMed: 23563705]
112. Tang WH, et al. Gut microbiota-dependent trimethylamine N-oxide (TMAO) pathway contributes to both development of renal insufficiency and mortality risk in chronic kidney disease. Circ Res. 2015; 116:448–55. [PubMed: 25599331]
113. Bae S, et al. Plasma choline metabolites and colorectal cancer risk in the Women's Health Initiative Observational Study. Cancer Res. 2014; 74:7442–52. [PubMed: 25336191]
114. Gao X, et al. Dietary trimethylamine N-oxide exacerbates impaired glucose tolerance in mice fed a high fat diet. J Biosci Bioeng. 2014; 118:476–81. [PubMed: 24721123]
115. Bhardwaj T, Somvanshi P. Pan-genome analysis of *Clostridium botulinum* reveals unique targets for drug development. Gene. 2017; 623:48–62. [PubMed: 28450142]
116. Abdul-Rahman F, Petit E, Blanchard JL. The distribution of polyhedral bacterial microcompartments suggests frequent horizontal transfer and operon reassembly. Journal of Phylogenetics & Evolutionary Biology. 2013:1–7.
117. Lawrence JG, Roth JR. Evolution of coenzyme B_{12} synthesis among enteric bacteria: Evidence for loss and reacquisition of a multigene complex. Genetics. 1996; 142:11–24. [PubMed: 8770581]
118. Kerfeld CA. A bioarchitectonic approach to the modular engineering of metabolism. Philos Trans R Soc Lond B Biol Sci. 2017; 372
119. Kerfeld CA. Plug-and-play for improving primary productivity. Am J Bot. 2015; 102:1949–50. [PubMed: 26656128]
120. Parsons JB, et al. Biochemical and structural insights into bacterial organelle form and biogenesis. Journal of Biological Chemistry. 2008; 283:14366–14375. [PubMed: 18332146]
121. Bonacci W, et al. Modularity of a carbon-fixing protein organelle. Proc Natl Acad Sci U S A. 2012; 109:478–83. [PubMed: 22184212]
122. Baumgart M, Huber I, Abdollahzadeh I, Gensch T, Frunzke J. Heterologous expression of the *Halothiobacillus neapolitanus* carboxysomal gene cluster in *Corynebacterium glutamicum*. J Biotechnol. 2017
123. Lin MT, et al. β-Carboxysomal proteins assemble into highly organized structures in *Nicotiana* chloroplasts. Plant J. 2014; 79:1–12. First step of integrating carboxysomes into plants demonstrated by expression of β-carboxysome genes in Nicotiana chloroplasts. [PubMed: 24810513]
124. Lin MT, Occhialini A, Andralojc PJ, Parry MA, Hanson MR. A faster Rubisco with potential to increase photosynthesis in crops. Nature. 2014; 513:547–50. [PubMed: 25231869]
125. Price GD, et al. The cyanobacterial CCM as a source of genes for improving photosynthetic CO_2 fixation in crop species. J Exp Bot. 2013; 64:753–68. [PubMed: 23028015]

126. McGrath JM, Long SP. Can the cyanobacterial carbon-concentrating mechanism increase photosynthesis in crop species? A theoretical analysis. Plant Physiol. 2014; 164:2247–61. [PubMed: 24550242]

127. Gonzalez-Esquer CR, Shubitowski TB, Kerfeld CA. Streamlined Construction of the Cyanobacterial CO_2-Fixing Organelle via Protein Domain Fusions for Use in Plant Synthetic Biology. Plant Cell. 2015; 27:2637–44. Redesign of the enzymatic core of β-carboxysomes. [PubMed: 26320224]

128. Sargent F, et al. A synthetic system for expression of components of a bacterial microcompartment. Microbiology. 2013; 159:2427–36. [PubMed: 24014666]

129. Held M, et al. Engineering formation of multiple recombinant Eut protein nanocompartments in *E. coli*. Sci Rep. 2016; 6:24359. [PubMed: 27063436]

130. Quin MB, Perdue SA, Hsu SY, Schmidt-Dannert C. Encapsulation of multiple cargo proteins within recombinant Eut nanocompartments. Appl Microbiol Biotechnol. 2016; 100:9187–9200. [PubMed: 27450681]

131. Wagner HJ, Capitain CC, Richter K, Nessling M, Mampel J. Engineering bacterial microcompartments with heterologous enzyme cargos. Engineering in Life Sciences. 2017; 17:36–46.

132. Liang M, Frank S, Lunsdorf H, Warren MJ, Prentice MB. Bacterial microcompartment-directed polyphosphate kinase promotes stable polyphosphate accumulation in *E. coli*. Biotechnol J. 2017; 12

133. Lee MJ, Brown IR, Juodeikis R, Frank S, Warren MJ. Employing bacterial microcompartment technology to engineer a shell-free enzyme-aggregate for enhanced 1,2-propanediol production in *Escherichia coli*. Metab Eng. 2016; 36:48–56. [PubMed: 26969252]

134. Jakobson CM, Slininger Lee MF, Tullman-Ercek D. De Novo Design of Signal Sequences to Localize Cargo to the 1,2-Propanediol Utilization Microcompartment. Protein Sci. 2017

135. Parsons JB, et al. Characterisation of PduS, the pdu metabolosome corrin reductase, and evidence of substructural organisation within the bacterial microcompartment. PLoS One. 2010; 5:e14009. [PubMed: 21103360]

136. Jakobson CM, et al. Tuning the Catalytic Activity of Subcellular Nanoreactors. J Mol Biol. 2016; 428:2989–96. [PubMed: 27427532]

137. Kim EY, Tullman-Ercek D. A rapid flow cytometry assay for the relative quantification of protein encapsulation into bacterial microcompartments. Biotechnol J. 2014

138. Jakobson CM, Kim EY, Slininger MF, Chien A, Tullman-Ercek D. Localization of proteins to the 1,2-propanediol utilization microcompartment by non-native signal sequences is mediated by a common hydrophobic motif. J Biol Chem. 2015; 290:24519–33. [PubMed: 26283792]

139. Dueber JE, et al. Synthetic protein scaffolds provide modular control over metabolic flux. Nat Biotechnol. 2009; 27:753–9. [PubMed: 19648908]

140. Luo LH, et al. Identification and characterization of the propanediol utilization protein PduP of *Lactobacillus reuteri* for 3-hydroxypropionic acid production from glycerol. Appl Microbiol Biotechnol. 2011; 89:697–703. [PubMed: 20890600]

141. Toth-Petroczy A, Tawfik DS. The robustness and innovability of protein folds. Curr Opin Struct Biol. 2014; 26:131–8. [PubMed: 25038399]

142. Bloom JD, Arnold FH. In the light of directed evolution: pathways of adaptive protein evolution. Proc Natl Acad Sci U S A. 2009; 106(Suppl 1):9995–10000. [PubMed: 19528653]

143. Pandya C, Farelli JD, Dunaway-Mariano D, Allen KN. Enzyme promiscuity: engine of evolutionary innovation. J Biol Chem. 2014; 289:30229–36. [PubMed: 25210039]

144. Kinney JN, Axen SD, Kerfeld CA. Comparative analysis of carboxysome shell proteins. Photosynth Res. 2011; 109:21–32. [PubMed: 21279737]

145. Cai F, Sutter M, Bernstein SL, Kinney JN, Kerfeld CA. Engineering Bacterial Microcompartment Shells: Chimeric Shell Proteins and Chimeric Carboxysome Shells. ACS Synthetic Biology. 2015; 4:444–453. Demonstrates that shell proteins can be exchanged between different BMC types. [PubMed: 25117559]

146. Thompson MC, et al. Identification of a unique Fe-S cluster binding site in a glycyl-radical type microcompartment shell protein. J Mol Biol. 2014; 426:3287–304. [PubMed: 25102080]

147. Pang A, Warren MJ, Pickersgill RW. Structure of PduT, a trimeric bacterial microcompartment protein with a 4Fe-4S cluster-binding site. Acta Crystallogr D Biol Crystallogr. 2011; 67:91–6. [PubMed: 21245529]

148. Aussignargues C, et al. Structure and Function of a Bacterial Microcompartment Shell Protein Engineered to Bind a [4Fe-4S] Cluster. J Am Chem Soc. 2016; 138:5262–70. First integration of a function, electron transfer, into a BMC shell protein. [PubMed: 26704697]

149. Sturms R, Streauslin NA, Cheng S, Bobik TA. In *Salmonella enterica*, Ethanolamine Utilization Is Repressed by 1,2-Propanediol To Prevent Detrimental Mixing of Components of Two Different Bacterial Microcompartments. J Bacteriol. 2015; 197:2412–21. [PubMed: 25962913]

150. Heldt D, et al. Structure of a trimeric bacterial microcompartment shell protein, EtuB, associated with ethanol utilization in *Clostridium kluyveri*. Biochem J. 2009; 423:199–207. First report of a ethanol utilization microcompartment. [PubMed: 19635047]

151. Pang A, Frank S, Brown I, Warren MJ, Pickersgill RW. Structural insights into higher order assembly and function of the bacterial microcompartment protein PduA. J Biol Chem. 2014; 289:22377–84. [PubMed: 24873823]

152. Noel CR, Cai F, Kerfeld CA. Purification and Characterization of Protein Nanotubes Assembled from a Single Bacterial Microcompartment Shell Subunit. Advanced Materials Interfaces. 2016; 3

153. Young EJ, et al. Engineering the Bacterial Microcompartment Domain for Molecular Scaffolding Applications. Front Microbiol. 2017; 8:1441. [PubMed: 28824573]

154. Mahalik JP, Brown KA, Cheng X, Fuentes-Cabrera M. Theoretical Study of the Initial Stages of Self-Assembly of a Carboxysome's Facet. ACS Nano. 2016; 10:5751–8. [PubMed: 26906087]

155. King NP, et al. Computational design of self-assembling protein nanomaterials with atomic level accuracy. Science. 2012; 336:1171–4. [PubMed: 22654060]

156. Jorda J, Liu Y, Bobik TA, Yeates TO. Exploring bacterial organelle interactomes: a model of the protein-protein interaction network in the Pdu microcompartment. PLoS Comput Biol. 2015; 11:e1004067. [PubMed: 25646976]

157. Wang P, Lombi E, Zhao FJ, Kopittke PM. Nanotechnology: A New Opportunity in Plant Sciences. Trends Plant Sci. 2016; 21:699–712. [PubMed: 27130471]

158. Tsai SJ, Yeates TO. Bacterial Microcompartments: Insights into the Structure, Mechanism, and Engineering Applications. Molecular Assembly in Natural and Engineered Systems. 2011; 103:1–20.

159. Shively JM, Ball F, Brown DH, Saunders RE. Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of *Thiobacillus neapolitanus*. Science. 1973; 182:584–6. [PubMed: 4355679]

160. Menon BB, Dou Z, Heinhorst S, Shively JM, Cannon GC. *Halothiobacillus neapolitanus* carboxysomes sequester heterologous and chimeric RubisCO species. PLoS One. 2008; 3:e3570. [PubMed: 18974784]

161. Romano KA, Vivas EI, Amador-Noguez D, Rey FE. Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-N-oxide. MBio. 2015; 6:e02481. [PubMed: 25784704]

Key points

Bacterial microcompartments are functional analogs of the lipid bound organelles of eukaryotes. They contain reactions that benefit from being separated from the cytosol.

The delimiting membrane of bacterial microcompartments consists entirely of protein and its components are highly conserved in sequence and structure.

Bacterial microcompartments are found in a wide variety of bacterial species (at least 19 established phyla). They are easily identified in genomes by their tendency to co-localize the associated genes in a large gene cluster, called superlocus.

Carboxysomes, CO2 fixing organelles, were the first identified type of bacterial microcompartment but recently many more have been discovered and characterized; they are involved in catabolizing a variety of nutrients and enable cells to grow in otherwise unavailable niches.

Bacterial microcompartments self-assemble the shell and cargo using different pathways, some build the shell around a cargo bolus while others assemble shell and cargo concomitantly. There are proteins that facilitate cargo aggregation and small encapsulation peptides that specifically associate proteins to the lumen of the shell.

Bacterial microcompartments are linked to pathogenesis of certain bacteria because they give them a growth advantage. The human gut, for example, is enriched in propanediol and ethanolamine, initial substrates of the well characterized PDU and EUT microcompartments.

The knowledge gained from understanding the native functions has led to substantial progress in modifying the shell for bioengineering purposes. Bacterial microcompartment shells can be produced recombinantly and shell proteins and cores have been engineered to adopt new functions.

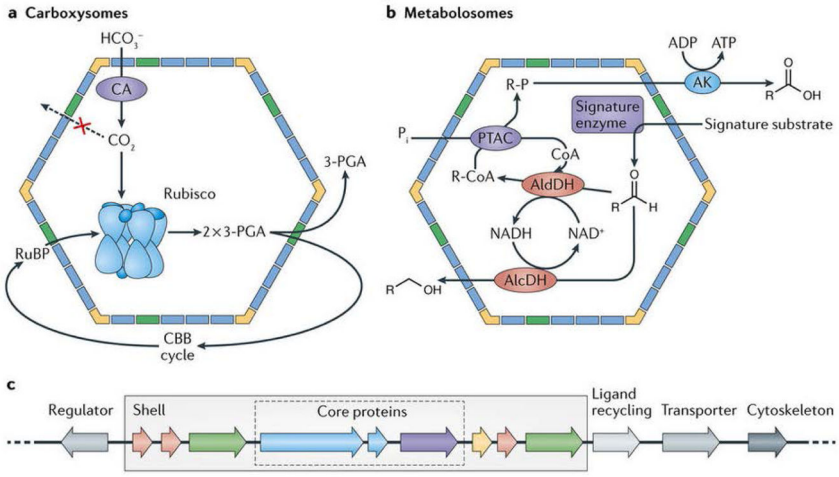


Figure 1. Core biochemistry of carboxysomes and metabolosomes and superloci organization

A) Schematic of the bacterial microcompartment (BMC) shell and encapsulated enzymes and associated flow of substrates/products. α-carboxysomes and β-carboxysomes both encapsulate carbonic anhydrase (CA) and Form 1 RuBisCO to fix CO_2 as part of the Calvin-Benson-Bassham Cycle (CBB). The shell prevents loss of CO2 to the cytoplasm. B) Metabolosomes have more diverse initial substrates compared to carboxysomes, but they typically share a common core biochemistry that is based on a signature enzyme, an aldehyde dehydrogenase (AldDH), an alcohol dehydrogenase (AlcDH) and a phosphotransacylase (PTAC). The signature enzyme generates the aldehyde which is then converted to a product alcohol by the AlcDH. This reaction uses CoA and NAD+ which are recycled in a separate reaction branch that uses AldDH and PTAC to produce a phosphorylated product (R-P). This product is then dephosphorylated by an acetyl kinase (AK) in a reaction that generates ATP. C) Schematic of a typical gene composition in a superlocus encoding a BMC. In addition to genes encoding shell and core proteins, BMC superloci encode proteins for supporting and ancillary functions, like transporters for the signature substrate.

3-PGA, 3-phosphoglycerate; RuBP, ribulose 1,5-bisphosphate;.

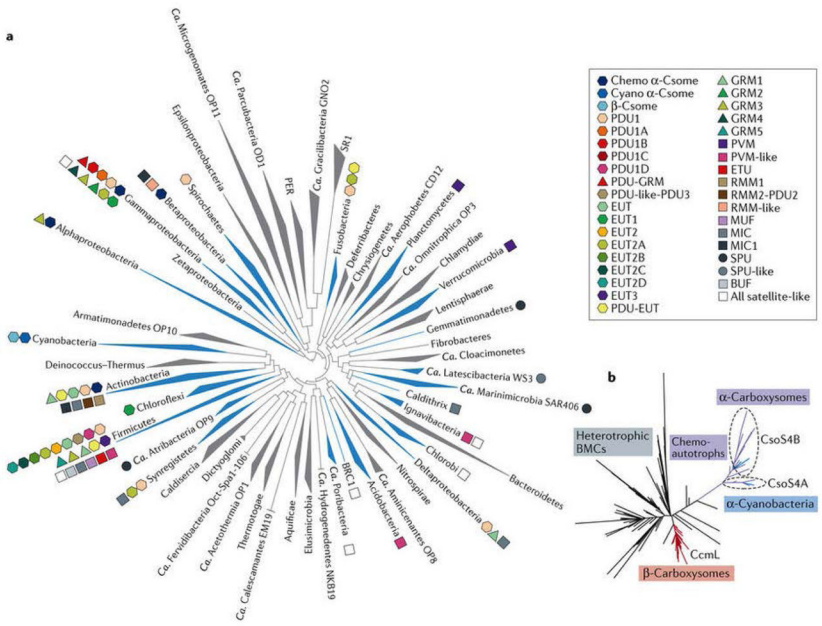


Figure 2. Overview of BMCs in different phyla and tree of shell pentamers

A) Bacterial phyla tree with distribution of bacterial microcompartment (BMC) locus types. Locus types, excluding satellite and satellite-like loci, denoted as colored shapes are adjacent to the phyla in which they appear. For a given phylum, the shape of the triangular wedge represents sequence diversity; the nearest edge represents the shortest branch length from the phylum node to a leaf, while the farthest edge represents the longest branch length from the phylum node to a leaf. **Reproduced with permission from REF.** [6]

B) Phylogenetic tree of BMC-P shell proteins (which occupies the vertex position of BMC shells) shows a large distance between α-carboxysomal and β-carboxysomal homologs. β-carboxysomal BMC-P proteins are more closely related to their counterparts in the metabolosomes of heterotrophic organisms than to the BMC-P proteins of α-carboxysomes. To construct the tree, amino acid sequences were identified in the Uniprot RP75 database with a Hidden Markov Model (HMM) of the BMC-P protein family (Pfam03319). The collected sequences were then made non-redundant with a cutoff of 95% identity, aligned and used to build the phylogenetic tree. **Reproduced with permission from REF.** [2]

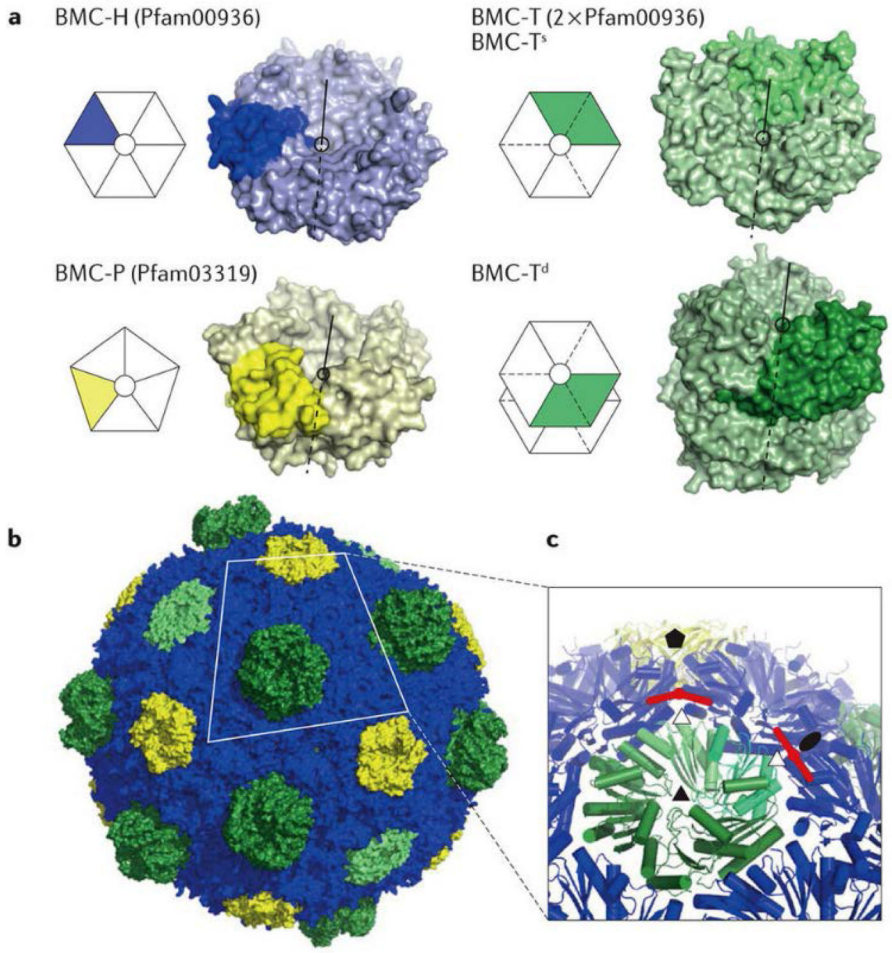


Figure 3. Schematics of BMC structure

A) Bacterial microcompartment (BMC) shells are made of three types of building blocks: BMC-H (blue), BMC-T (green) and BMC-P (yellow). BMC-T has two subtypes based on sequence and oligomer status: single layered BMC-T^s and double layered BMC-T^d. Pores (circles) and central axes of symmetry are shown in the space-filling orientations. B) Three building blocks tile together to form BMC shells. Overview of the *Haliangium ochraceum* (HO) whole shell and the highly conserved planar BMC-domain interactions as observed in the HO BMC shell structure and in the crystal packing of many other BMC-H structures. C) Position of the BMC-T in the shell, interfacing with two different BMC-H edges (planar and tilted, marked with red lines; icosahedral symmetry axis indicated with filled symbols, pseudosymmetry with open symbols, top layer of BMC-T omitted for clarity).

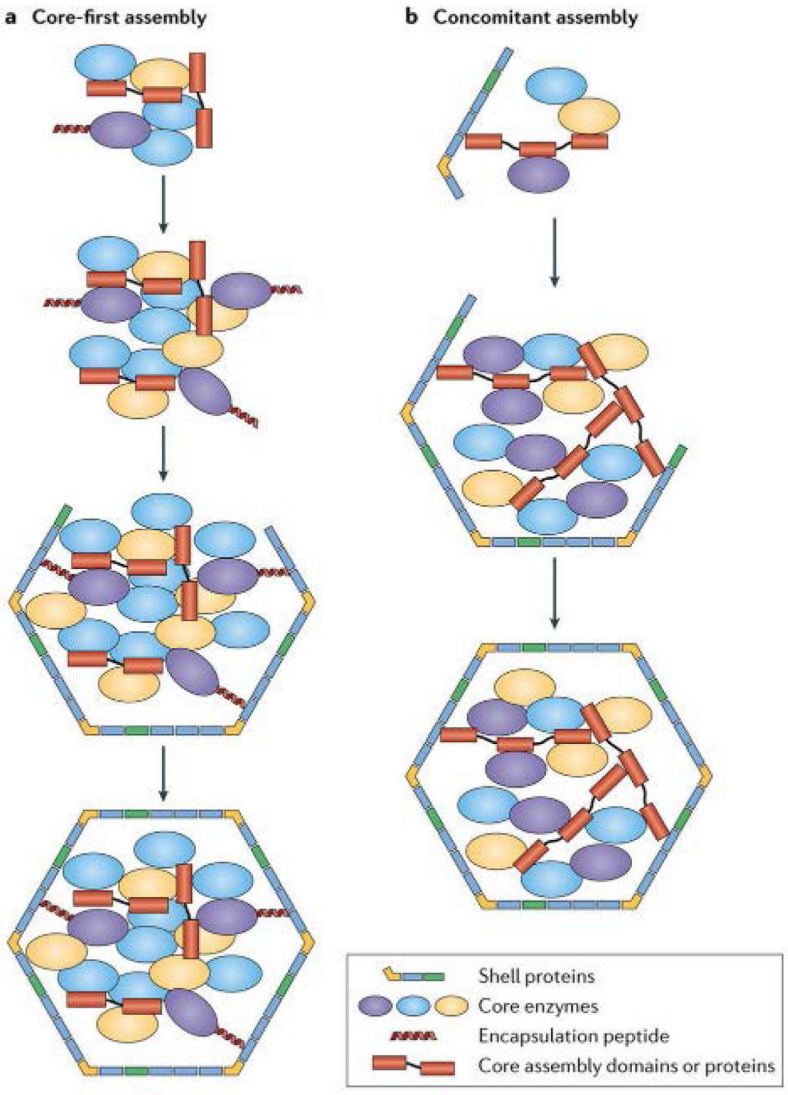


Figure 4. Schematics of bacterial microcompartment (BMC) assembly pathways

A) Core first assembly: core proteins coalesce through protein domain interactions or aggregation of encapsulation peptides. Encapsulation peptides interact with shell proteins which then form the shell around the core. B) Concomitant assembly: core and shell proteins assemble together with the help of core assembly proteins.

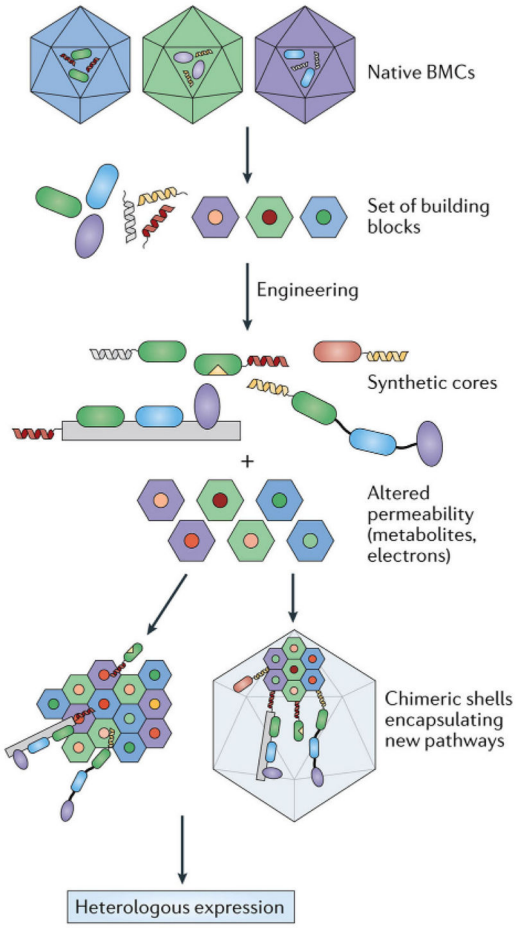


Figure 5. Example of strategies to design synthetic bio-nanoreactors and scaffolding architectures using building blocks from bacterial microcompartments (BMCs)

The natural diversity of BMCs provides a rich collection of building blocks with specific functions, structures, permeability and enzymatic activity. New enzymatic functions can be targeted to the interior of the shell using encapsulation peptides and synthetic BMC cores can be engineered using protein fusions. Modifying the shell proteins enables altered permeability or the introduction of new functions such as electron transfer. Biomaterials, such as scaffolds for metabolic pathways, can also be constructed using the same strategies. Synthetic BMCs and scaffolds can be used *in vivo* and *in vitro*.