Published in final edited form as:
Annu Rev Biophys. 2010 June 9; 39: 185–205. doi:10.1146/annurev.biophys.093008.131418.

Bacterial Microcompartment Organelles: Protein Shell Structure and Evolution

Todd O. Yeates[1,2,3], **Christopher S. Crowley**[2], and **Shiho Tanaka**[3]

Todd O. Yeates: yeates@mbi.ucla.edu

[1]Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095

[2]Molecular Biology Institute, University of California, Los Angeles, California 90095

[3]UCLA-DOE Institute of Genomics and Proteomics, University of California, Los Angeles, California 90095

Abstract

Some bacteria contain organelles or microcompartments consisting of a large virion-like protein shell encapsulating sequentially acting enzymes. These organized microcompartments serve to enhance or protect key metabolic pathways inside the cell. The variety of bacterial microcompartments provide diverse metabolic functions, ranging from CO_2 fixation to the degradation of small organic molecules. Yet they share an evolutionarily related shell, which is defined by a conserved protein domain that is widely distributed across the bacterial kingdom. Structural studies on a number of these bacterial microcompartment shell proteins are illuminating the architecture of the shell and highlighting its critical role in controlling molecular transport into and out of microcompartments. Current structural, evolutionary, and mechanistic ideas are discussed, along with genomic studies for exploring the function and diversity of this family of bacterial organelles.

Keywords

carboxysome; molecular transport; protein assembly; nanocompartments; protein shell

INTRODUCTION

Although they lack the typical membrane-bound organelles of eukaryotic cells, bacterial cells can have highly organized interiors. Electron microscopic investigations of bacteria have revealed a wide variety of subcellular inclusions, many of which remain to be characterized (38, 61). Some serve as storage bodies or granules for special compounds or polymers (4, 65), whereas others perform metabolic activities complex enough to meet the functional definition of an organelle (7, 62, 77). Few of these bacterial organelles have been studied in detail.

Several members of one diverse family of organelles have been either observed directly or inferred to exist from genomic data. Where they have been observed by electron

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

microscopy, they appear as dense, typically polyhedron-shaped bodies ranging in size from 80 to 200 nm (9, 13, 25, 63) (Figure 1). Here we refer to this family of polyhedral organelles as bacterial microcompartments. The current view is that they are proteinaceous assemblies comprising several thousand protein and enzyme subunits, lacking the lipid bilayer of eukaryotic organelles and some viruses, and lacking the nucleic acid component of viruses and some eukaryotic organelles. Each kind of microcompartment consists of two or a few distinct types of enzymes in the interior, acting sequentially in a metabolic pathway, surrounded by a thin protein shell reminiscent of a viral capsid (reviewed in Reference 77). Functionally diverse bacterial microcompartments contain different enzymes, though they have shells assembled from similar proteins bearing a conserved bacterial microcompartment (BMC) protein domain (14, 16, 23, 51). The presence of BMC shell proteins effectively defines the bacterial microcompartment family of organelles.

The various reactions encapsulated by bacterial microcompartments share a common theme. A small metabolite, which is volatile or cytotoxic, is produced by one interior enzyme and consumed by another. Such arrangements are capable of enhancing the flux through desired pathways while limiting alternative, detrimental processes; various simpler protein-based systems for substrate channeling have been elucidated (52, 67). The functions of bacterial microcompartments are reviewed here, with an emphasis on the structure and evolution of their protein shells.

COMPARTMENTALIZED REACTIONS AND PROCESSES

Data have been presented for the existence of bacterial microcompartments carrying out at least six different metabolic pathways. Three of these pathways have been explored experimentally (Figure 2), and genomic searches suggest a wide range of unexplored functions.

CO_2 Fixation in the Carboxysome

Beginning some 50 years ago, dense polyhedral bodies of unknown function were observed in numerous cyanobacteria and chemoautotrophic bacteria (21, 25, 63). In 1973, Shively et al. (62) isolated the inclusions from the chemoautotroph *Halothiobacillus neapolitanus* and showed that they consisted mainly of the CO_2-fixing enzyme ribulose bisphosphate carboxylase/oxygenase (RuBisCO). The inclusions were named carboxysomes because of their inferred function in CO_2 fixation. They became the founding member of the polyhedral bacterial microcompartments. The identification and sequencing of conserved shell proteins led to the possibility of genomic studies (13, 16, 23, 41, 51). Sequence analysis indicated that carboxysomes are present in some chemoautotrophs and all known cyanobacteria. Owing to the abundance of photosynthetic microbes in oceans, carboxysomes contribute substantially to the global CO_2 cycle (2).

The active role of carboxysomes in CO_2 fixation has been established by numerous genetic and biochemical studies (reviewed in References 2 and 28). It is now generally understood that carbonic anhydrase (CA) is colocalized with RuBisCO inside the carboxysome (18, 29, 50) and that this is a key mechanistic feature for enhancing CO_2 fixation when the concentration of inorganic carbon is low, as it is in most natural environments. The carboxysome comprises the second part of a carbon concentrating mechanism (CCM), whose first part is the concentration of bicarbonate inside the cell by active transport across the cytoplasmic membrane and into the cytosol (2). Current models hold that bicarbonate then crosses the protein shell and enters the carboxysome by diffusion. In the lumen of the carboxysome, CA dehydrates bicarbonate to CO_2 in a space where diffusion is restricted and where RuBisCO is highly concentrated, thereby countering the low catalytic efficiency and poor substrate selectivity of RuBisCO for CO_2 (Figure 2). The carboxysome shell is

believed to play a role in allowing or limiting the transport of substrates and products [i.e., bicarbonate, CO_2, 3-phosphoglycerate (3-PG), and ribulose-1,5-bisphosphate] (20, 39, 54, 58) and possibly O_2 (46), which competes with CO_2 in its reaction with RuBisCO.

Two kinds of carboxysomes, α and β, can be delineated on the basis of the identity and genomic organization of their carboxysomal proteins (2) (Figure 2). Form I RuBisCO and shell proteins of the α- and β-carboxysomes are homologous, although their sequences can be grouped by type. On the contrary, the CA enzymes, along with other proteins believed to help organize the carboxysome, differ between α- and β-carboxysomes. Some of the mechanistic features may therefore differ between the two types of carboxysomes.

1,2-Propanediol Utilization in the Pdu Microcompartment

Microcompartments related to but functionally distinct from the carboxysome were first implicated by genomic analysis when a gene encoding a homolog (PduA) of the carboxysome shell proteins was identified in the heterotrophic bacterium *Salmonella enterica* (16). Multiple presumptive shell genes (PduA/B/J/K/T/U) were identified in a large 21-gene operon (Figure 2) for metabolism of 1,2-propanediol (1,2-PD) (8), a compound that arises during the anaerobic breakdown of some sugars. Subsequent studies demonstrated the induction of polyhedral microcompartments by growth on 1,2-PD and confirmed the role of the microcompartment in degrading that compound (8). Isolation of these propanediol utilization (Pdu) microcompartments allowed the identification of enzymes localized to the microcompartment and the development of tentative models for the complex metabolic pathway for compartmentalized 1,2-PD metabolism (26) (Figure 2).

The central feature of the Pdu microcompartment is the breakdown of 1,2-PD to propionaldehyde by an adenosylcobalamin-dependent diol dehydratase, followed by the conversion of propionaldehyde to propionyl-CoA and 1-propanol without escape of the aldehyde into the cytosol. Evidence suggests that propionaldehyde is toxic when it is produced in the cytoplasm in genetic mutants lacking the microcompartment shell. Initial growth on 1,2-PD as a sole carbon source is followed by a period of growth arrest (27), and the buildup of aldehyde correlates with increased DNA damage (57). The findings indicate that the organization of enzymes for forming and consuming propionaldehyde within the protein shell effectively channels the aldehyde through a productive pathway while limiting its cellular toxicity.

Questions remain concerning which molecules must pass through the protein shell, although 1,2-PD presumably passes inward across the shell, and propionaldehyde formed by dehydration of 1,2-PD is presumably oxidized to propionaldehyde-CoA before (or perhaps as) it passes outward through the shell (26) (Figure 2). In addition, it appears likely that cobalamin and possibly nucleotide cofactors may have to cross the shell. This raises an important paradox regarding how large molecules can be transported without allowing the much smaller aldehyde to escape (17). Structural studies on microcompartment shell proteins are providing clues to that puzzle.

Ethanolamine Utilization in the Eut Microcompartment

Several enteric bacteria, including *E. coli* and *Salmonella enterica*, have genes for producing an ethanolamine utilization (Eut) microcompartment. The 17-gene *eut* operon encodes several enzymes and four proteins homologous to microcompartment shell proteins (EutM/K/L/S) (41, 66) (Figure 2). Microcompartments are induced by growth on ethanolamine and have been visualized by electron microscopy (9). The Eut microcompartment appears to be closely related to the Pdu microcompartment in terms of encoded enzymes and chemical reactions. Parallels between the two are therefore useful in developing a model for

metabolism in Eut (48) (Figure 2). Whereas dehydration of the three-carbon 1,2-PD leads to propionaldehyde in the Pdu microcompartment, deamination of the two-carbon ethanolamine leads to acetaldehyde in the Eut microcompartment. Eut microcompartments have not been purified, so we can only surmise which enzymes and reactions occur inside the cell, but it is likely that ethanolamine crosses the protein shell into the microcompartment, where it is converted to acetaldehyde and then to acetyl-CoA before the acetaldehyde can escape into the cytosol. Questions concerning the retention of small molecules and the transport of large cofactors are similar to those noted above for Pdu.

Two different hypotheses, not necessarily mutually exclusive, have been put forward regarding the defect that arises when acetaldehyde is produced in the absence of a microcompartment. Genetic data in *Salmonella* demonstrate that glutathione-based detoxification pathways and DNA polymerase repair functions are essential under conditions in which high concentrations of aldehydes are either added or produced during 1,2-PD or ethanolamine metabolism (55). Alternatively, it has been emphasized that acetaldehyde is sufficiently apolar and volatile (bp = 21°C) to pass readily through the bacterial cell membrane, suggesting that acetaldehyde in the cytosol might lead to detrimental loss of carbon (48). When *Salmonella* was grown at lower concentrations of ethanolamine, acetaldehyde toxicity was not evident. Instead, growth defects associated with microcompartment mutations correlated with evaporative loss of acetaldehyde (48). The potential importance of small-molecule volatility relates closely to the carboxysome. Part of the advantage conferred by the carboxysome is that CO_2 produced inside the carboxysome is more likely to be fixed by RuBisCO than to escape the cell by diffusion across the cell membrane, which it does readily when CO_2 is produced in the cytosol (50). Clarification of the toxicity versus volatility issue for the Eut microcompartment requires further study, but it may be that loss of microcompartment function causes multiple defects that manifest themselves differently under disparate conditions.

Other Presumptive Microcompartments

Few bioinformatics studies on bacterial microcompartments have been reported (see below), but microcompartment shell genes have been identified in some relatively uncharacterized operons coding for enzymes involved in small-molecule metabolism. From this, the likely compartmentalization of certain pathways has been inferred. A microcompartment for metabolizing pyruvate to ethanol has been postulated in *Vibrio furnissii* M1 (74). A microcompartment for fermenting ethanol to acetate by way of an acetaldehyde intermediate in *Clostridium kluyveri* in a pathway for butyrate synthesis, along with a presumably distinct microcompartment for metabolizing glycerol, presumably through a 3-hydroxypropionaldehyde intermediate, has been postulated (60). These putative compartmentalized pathways are generally consistent with the varied themes discussed above: channeling of substrates, mitigation of toxicity, and control of evaporative loss. Bioinformatics approaches can be used to identify a wider range of proteins and enzymes likely to be associated with microcompartments (see below).

PROTEIN SHELL STRUCTURE AND ASSEMBLY

The first electron microscopy studies of isolated carboxysomes (62) indicated that the inclusions were surrounded by a thin membrane of undetermined composition, subsequently established to be a proteinaceous shell (23, 34). Although early impressions left open the possibility that microcompartments could be relatively unorganized packages of enzymes, emerging views of the shell are emphasizing its sophisticated structure and mechanisms.

Gene sequencing and microcompartment purification experiments demonstrated that microcompartments are composed of several protein components in addition to those

presumed to carry out required enzymatic functions (15, 26, 34, 51). The most abundant nonenzymatic components purified from both carboxysomes and Pdu microcompartments are members of a conserved family of small proteins, typically around 100 amino acids in length, that define the BMC protein domain (InterPro domain IPR000249). These BMC proteins are the major constituents of microcompartment shells. Furthermore, in all organisms for which genomic data have revealed a BMC protein for a presumptive microcompartment, multiple paralogous BMC proteins have been found. For example, CcmK1/2/3/4 are present in one type of carboxysome, CsoS1A/B/C in another, PduA/B/J/K/T/U for Pdu, and EutL/M/K/S for Eut (Figure 2). The involvement of multiple distinct BMC proteins in the shell appears to be a strongly conserved feature (5, 39). In addition to the major (BMC-type) shell proteins, other less abundant proteins are likely to be involved in the formation of microcompartment shells. Structural studies on individual shell proteins have begun to illuminate key mechanisms of microcompartment function.

BMC Domains: The Hexameric Building Blocks of the Shell

In a bottom-up approach to studying the shell, individual BMC shell proteins have been purified from heterologous protein expression systems. To date, crystal structures of 12 different BMC proteins have been determined, some in multiple crystal forms and conformations (Supplemental Table 1; follow the Supplemental Material link from the Annual Reviews home page at http://www.annualreviews.org). These include three from the β-type carboxysome (39, 68), three from the α-type (40, 72, 73), one from the Pdu microcompartment (19), four from the Eut microcompartment (56, 70), and one predicted to function in an ethanol-utilizing (Etu) microcompartment (30). From these crystal structures, basic features of assembly and transport have been inferred.

The BMC domain adopts an α/β-fold with a central four-stranded antiparallel sheet flanked by small helices (Figure 3). No similarity is evident between the BMC domain and any viral capsid proteins whose structures are presently known. The BMC fold resembles that seen in various small proteins, including ferredoxin, but whether this reflects a remote evolutionary relationship or a convergence to a relatively simple folding motif is unclear.

The BMC protein structures published to date are homohexamers with cyclic (C_6) symmetry, or some variation on that arrangement (Figure 3). Extensive interactions hold the monomers together within a hexamer. The hexameric arrangements visualized in crystal structures of different BMC domain proteins are essentially superimposable, and biophysical studies confirm that the proteins tend to assemble as hexamers in solution. Numerous lines of evidence therefore show that hexamers comprise the building blocks of bacterial microcompartment shells (39, 72, 78). Most BMC shell proteins adopt structures in which the hexamer has a prominent bowl-shaped depression on one side, specifically the side on which both the N and C termini reside in the typical BMC domain. The C termini are often observed to be flexible or disordered and therefore are difficult to model. The center of the typical BMC hexamer is perforated by a narrow central pore, which has been postulated to allow transport of small metabolites into and out of the microcompartment (39, 72, 78).

Assembly into Molecular Layers

Several different crystal structures of BMC proteins reveal tightly packed two-dimensional layers of molecules (39, 56, 68, 72, 73) (Figure 3) (Supplemental Table 1). The shape of the BMC hexamer is evidently tailored for further side-by-side assembly into a molecular sheet, roughly 20 Å thick on average, with a center-to-center distance between hexamers ranging from 66 to 69 Å. The molecular layer is nearly solid, except for narrow pores through the centers of the hexamers and where hexamers meet at twofold and threefold axes of symmetry.

The conserved sheet structure has been postulated to represent the natural structure of the microcompartment shell. Recent two-dimensional crystallography experiments support this view (22). The three BMC shell proteins from the *Synechocystis* sp. PCC 6803 carboxysome that could be expressed and purified successfully (CcmK1, 2, and 4) were arrayed via their polyhistidine affinity tails on the surface of droplets containing nickelated lipid molecules. Electron microscopy and Fourier transform analysis showed hexagonal molecular packings closely matching those reported in the context of three-dimensional crystals.

Some of the BMC proteins do not appear to form layers by themselves, suggesting alternative behavior or specialized roles. One BMC protein from the *Syn*. 6803 β-type carboxysome, CcmK4, formed linear strips of hexamers in three-dimensional crystals rather than uniformly oriented layers (39). CcmK4 is suspected to be only a minor component of the shell, which appears consistent with its failure to form a layer by itself. Whether the unusual behavior of CcmK4 is biologically relevant—reflecting, for example, a specialized role in forming the edges of a polyhedral shell—remains to be tested.

Overall Shell Architecture

The first crystal structures of BMC proteins explained how facets of a polyhedral shell could be constructed by the packing of hexagonal building blocks, following principles common to various viral capsids, but the overall architecture of the microcompartment shell remained unclear. Electron microscopy studies had led to opposing conclusions that carboxysomes were shaped like an icosahedron (49), a pentagonal dodecahedron (33), or a hexagonal structure with D_6 symmetry (13). Two recent electron cryotomography studies on α-carboxysomes from highly divergent organisms concluded that carboxysomes are at least roughly icosahedral, with nearly flat triangular facets (36, 59). The resolutions of the studies were not high enough to visualize the packing arrangements of individual shell subunits, but the findings were consistent with models proposed on the basis of crystal structures. The shell was confirmed to be a nearly solid layer of molecules. A few layers of enzymes, presumed to be mainly RuBisCO molecules, were visible in the microcompartment interior beginning at the inward facing side of the shell, consistent with earlier studies (33). Less is understood about the shape and organization of other kinds of microcompartments, as these have not yet been studied by similar techniques, but traditional electron microscopy studies suggest that Pdu and Eut microcompartments are not as geometrically regular as carboxysomes.

Minor Shell Components and Vertex Proteins

In different microcompartments, multiple proteins are present that are apparently neither enzymatic nor homologous to the major BMC shell protein family (Figure 2) (Supplemental Table 1). Such proteins are presumed either to help organize the microcompartment enzymes, or to be minor structural components of the shell. There is considerable variation regarding the accessory proteins present in different microcompartments. For example, the proteins CcmM and CcmN are present only in β-carboxysomes, whereas CsoS2 is present only in α-carboxysomes. Even greater variation seems likely in metabolically diverse systems. However, the presence of one minor protein is conserved across microcompartments and is named according to microcompartment type: CcmL (β-carboxysome), CsoS4A/B (α-carboxysome), PduN, or EutN—collectively referred to as the CcmL/EutN protein family.

Proteins from the CcmL/EutN family are found encoded in operons wherever BMC genes are present (Figure 2). Paralogous copies of the gene are often present, paralleling the situation observed for BMC proteins (5). Whether genes from this family encoded bona fide microcompartment genes was initially unclear, because the homologous gene products in the

α-carboxysome operon (CsoS4A/B) were not detected initially in purified *H. neapolitanus* carboxysomes (15). However, deletion of the *ccmL* gene in cyanobacterial β-carboxysomes gave rise to aberrant, elongated tubular carboxysomes (51). Crystal structures of corresponding proteins, CsoS4A and CcmL, from the two types of carboxysomes clarified their structural roles. Both proteins are pentamers whose sizes and shapes are compatible with formation of vertices in an icosahedral shell assembled mainly from BMC hexamers (68). On that basis, rough atomic models of carboxysome shells have been constructed (Figure 3). Models in which the CcmL or CsoS4 proteins occupy icosahedral vertices explain the two key observations noted above. First, vertex proteins would be present in only 60 copies per shell, compared to a few thousand for the BMC proteins and the RuBisCO subunits, explaining difficulties in detection; later studies confirmed the expression of the csoS4A/B genes (11). Second, in the absence of pentagonal vertices, a flat layer of hexamers can roll up into a tube, whereas introducing the (Gaussian) curvature required to close the ends is problematic without pentamers.

Understanding the role of proteins from the CcmL/EutN protein family is complicated, however, by two observations. A recent experiment demonstrates that, in *H. neapolitanus*, deletion mutants lacking the CcmL homologs form only a minority of aberrant, tubular carboxysomes (12). Moreover, the crystal structure of EutN shows a hexameric assembly (24, 68). It is not clear yet whether architectural principles distinct from those in the carboxysome are at work in the Eut microcompartment, but the problem of assembling closed structures from only hexagonal units has been encountered before in other biological contexts (76). Alternatively, despite its behavior in isolation, EutN could be pentameric in the context of the native shell. The reported crystal structure of EutN is more distorted from perfect hexagonal symmetry compared with typical BMC proteins and other homo-oligomeric protein assemblies in general. Similarly, a recent structure of the EutS shell protein revealed a highly distorted hexameric assembly (70). The asymmetry of these Eut shell proteins could relate to the lesser geometric regularity observed in Pdu and Eut compartments compared to the carboxysome.

PROTEIN SHELL EVOLUTION

Bacterial microcompartments have a complex evolutionary history. Operons encoding microcompartment genes are widely but sporadically distributed across the bacterial kingdom (7, 39, 77). Closely related bacterial strains can differ with respect to the presence of microcompartment genes, while close similarities can be seen between genes in more distant lineages. The carboxysome genes from α-type cyanobacteria are more similar to genes from chemoautotrophic bacteria than to those from β-type cyanobacteria. Microcompartment operons have apparently been the subject of horizontal gene transfer events (31). The widespread occurrence of gene transfer events between bacteria, particularly for genes involved in metabolic pathways (37), has been emphasized, with clusters of genes offering a particular advantage (42). Indeed, BMC genes appear to have been transferred often as clusters. In some cases, the underlying genetic events have been clarified by three-dimensional structural elucidation.

Evolution of the BMC Protein Family

A current search of the sequence databases detects some 4300 BMC shell proteins from bacteria belonging to 10 different phyla. BMC shell proteins are particularly common in the Gammaproteobacteria, Firmicutes, and cyanobacteria. A single potentially homologous protein is detected in the Epsilonproteobacteria, in *H. pylori*. Homologues have not been detected in the archaeal or eukaryotic kingdoms, though the possibility of their undetected existence among primitive algae has been raised (10).

BMC gene duplication events are common. In genomes where they appear, the average number of BMC paralogs found is 2.6 (5). About 75% of BMC genes are found immediately adjacent to another BMC gene. Such adjacent BMC genes are often similar in sequence, representing recent gene duplications. In other cases, paralogous BMC proteins within an organism are more divergent than orthologous genes from different organisms, indicating gene duplication prior to speciation or horizontal gene transfer. The evolutionary distances between several different BMC proteins are illustrated in Supplemental Figure 1.

Multiple sequence alignments highlight segments of relatively high conservation. Knowledge of the BMC protein structure makes it possible to map conserved sites to three-dimensional space (Figure 4). The regions of relatively high sequence conservation are around the perimeter of the hexamer; these are the interfaces where multiple hexamers meet in a layer. This is consistent with an increased level of evolutionary constraint in protein-protein interfaces compared to that in solvent-exposed protein surfaces. In contrast, the pore is a region of relatively high divergence with respect to amino acid sequence and loop lengths, suggesting functional specialization for transport. One exception is a widely conserved glycine residue (residue number 38 in CcmK1) that lines the pore.

Protein Fold Variations

The amino acid sequences of BMC proteins have diverged over large evolutionary distances such that recognizing homology is difficult in some cases. Detecting evolutionary relatedness is further complicated by the varied genetic mechanisms that have affected the protein fold. These have been clarified by structural studies (Figure 5). The crystal structure of the PduU shell protein revealed a circularly permuted version of the BMC domain (19). The structure of PduU overlaps with the canonical BMC fold, but its secondary structure elements occur in nonsequential order. If an attempt is made to align the sequence of PduU with a canonical BMC protein, PduU is shorter at its C terminus and longer at the N terminus. This extra N-terminal region of PduU fulfills the structural role of the C-terminal region of the canonical BMC fold. The PduU fold is also distinct because its novel N-terminal segment extends away from the hexagonal disk. Six copies of this segment, one from each monomer, come together to form a tightly wound six-stranded β-barrel that blocks the pore present in the canonical BMC (19).

Genetic events leading to fused BMC domains have also been common, leading to widespread tandem BMC domains. Three tandem BMC domain proteins have been characterized structurally: a widespread carboxysome shell protein called CsoS1D (40), EutL from the Eut microcompartment (56, 70), and EtuB (30). Surprisingly, these tandem BMC domain proteins were revealed to be duplications of a permuted BMC fold closely resembling that of PduU. However, despite the apparently parallel genetic events leading to their similar overall construction, the tandem domains are arranged in a different spatial order in different proteins (Figure 5). Yet another apparently distinct type of tandem BMC domain protein, CcmO, is present in some β-type (cyanobacterial) carboxysomes. The crystal structure of CcmO is not known, but sequence alignments suggest that it might be composed of tandem canonical (unpermuted) BMC domains. In addition to the fusion of multiple BMC domains together, many proteins comprise a BMC domain fused to another distinct domain. Among the numerous domains found fused to BMC domains, only one, EutK, has been characterized structurally. The extra domain present in EutK was revealed to have a helix-turn-helix structure, suggesting a potential nucleic acid binding function (70).

The distinct topological variations observed for BMC proteins are diagrammed in Figure 5. The variation seems remarkable for such a small protein domain. As noted above, multiple BMC domain proteins appear to be present in all microcompartment operons, and varied topologies may be required for distinct structural or functional roles in the shell. The specific

locations and relative arrangements of distinct BMC proteins in native shells have not been deciphered.

MOLECULAR TRANSPORT

Structural studies have focused on the importance of the protein shell in microcompartment function. That the shell proteins have evolved shapes and assembly patterns with such tight packing makes a compelling argument for their roles in controlling molecular transport.

Experimental Studies

The possibility that the protein shells of microcompartments might present important diffusion barriers has long been appreciated (13, 54), but direct experimental tests of the transport properties of the shell have been challenging. Nonetheless, important inferences have been possible. RuBisCO inside the carboxysome shell was less inhibited by O_2 than was RuBisCO expressed in the cytosol (46), suggesting that the carboxysome shell could restrict diffusion of O_2, which is competitive with CO_2. Likewise, the shell provides a barrier to CO_2. A key experiment showed that ectopic expression of CA in the cytosol, which causes dehydration of bicarbonate to CO_2, disrupted CO_2 fixation (50). This demonstrated that localizing CO_2 production to the carboxysome interior is critical, evidently because evaporative loss occurs more rapidly than transport across the shell when CO_2 is produced outside the microcompartment shell. Moreover, experiments on carboxysomes lacking CA showed that cytosolic CO_2 is about three times less kinetically accessible to RuBisCO than is CO_2 produced by CA inside the carboxysome (20). Similarly, the rate of CO_2 hydration by CA (the reverse of the natural reaction) in intact carboxysomes is about three times slower than that in disrupted carboxysomes (29). Experiments on the carboxysome are therefore consistent with a shell that acts as a moderate barrier but does not restrict diffusion severely. More quantitative interpretations are complicated, however, as it is difficult to determine whether the kinetic effects are due entirely to the shell or to a combination of geometric considerations, such as the production of CO_2 in a region of densely packed RuBisCO molecules (53).

Structural Data and Mathematical Models

Detailed kinetic modeling of microcompartments has been limited by an incomplete understanding of the structure of the shell and the organization of the interior enzymes. Although the organizational questions are still murky, crystal structures of the shell components offer a starting point for a quantitative understanding of transport. As noted above, the BMC shell proteins are tightly packed in a molecular layer, leaving but small pores down the center of each hexamer and, in some cases, similarly small pores where multiple hexamers meet. The canonical BMC proteins whose structures have been determined typically have pores with diameters ranging from 4 to 6 Å (39, 68, 69, 72) (Figure 6). These pores are likely large enough to permit the diffusion of small substrates and products of the known microcompartments (e.g., bicarbonate, 3-PG, 1,5-RuBP, ethanolamine, and 1,2-PD). Positive electrostatic potential has been noted in the hexameric pores of the major BMC proteins of the carboxysome (69, 72). It has been suggested that this could confer a kinetic advantage for bicarbonate (HCO_3^-) diffusion into the shell. CO_2 and O_2, whose respective outward and inward passages are undesirable, would not benefit from the same electrostatic advantage as bicarbonate. The pore regions of various BMC hexamers are illustrated in Figure 6.

A definitive demonstration that the pores visualized in BMC protein structures serve as portals for transport has been elusive. Metabolites have not been visualized in the pores by X-ray crystallography, although negatively charged sulfate ions have been visualized in the

hexameric pores in multiple cases (69, 72). Even in the carboxysome, where several shell protein structures are known, there are components potentially residing in the shell whose structures have not been elucidated. Examples include the CsoS2 protein in the α-carboxysome and CcmN and CcmM in the β-carboxysome. The possibility that these proteins provide routes for molecular transport cannot be discounted.

The small size of the pores in BMC shell proteins, and their distant spacing in the shell (nearly 70 Å apart), raises the question of whether such an arrangement would present a diffusive barrier so great as to offset any kinetic advantage gained by sequestering sequential enzymes together. Analytical calculations suggest this is not the case (72). Owing to the complex properties of diffusion, a relatively high porosity can be achieved in a spherical shell (or a polyhedral shell in the present case) when the pores represent only a small fraction of the outer surface. The analysis is parallel to the classic problem of receptors or transporters embedded in a bacterial cell membrane (6). For a microcompartment of typical size, a calculation based on shell proteins with pore diameters in the 4 to 6 Å range shows that such a shell would present only a moderate diffusion barrier (e.g., on the order of 50% porosity). These analytical calculations are consistent with experimental studies of diffusion discussed above. The moderately limiting properties of such a diffusion barrier could be offset by the advantages of enzyme compartmentalization, particularly if the pores provided mechanisms for selective transport of desired substrates and products.

Dynamic Mechanisms

Recent structural data have highlighted potentially dynamic transport processes in microcompartment shells. The structure of a newly identified carboxysome shell protein, CsoS1D, was determined and found to exist in two dramatically different conformations (40) (Figure 6). The two conformations of CsoS1D correspond to an open and closed central pore. The closed form is essentially occluded, whereas the open form has a pore roughly 14 Å in diameter (Figure 6). Likewise, EutL has been visualized in two different conformations, one that is nearly occluded (56, 70) and one that has a roughly 11 Å pore (70). In addition to offering an explanation of how larger molecules might be transported, the two conformations suggest a potential gating mechanism that could control which molecules pass through the shell (40, 70). This would help address the question of how larger molecules might be able to pass through a shell that presents a diffusion barrier to smaller compounds.

A second potentially significant observation is that CsoS1D tends to form a double-disk structure via a face-to-face interaction of two trimers (or pseudohexamers). This creates a large interior cavity. In two different crystal forms, the double-disk is composed of one trimer in the open configuration and a second in the closed configuration, leading the investigators to suggest an alternating access mechanism (40). The discovery of dynamic shell proteins (CsoS1D and EutL) hints that further studies on diverse microcompartment proteins will reveal a multitude of varied pores and transport mechanisms.

MICROCOMPARTMENT BIOINFORMATICS AND GENOMIC CONTEXT

A number of sequence searches for proteins homologous to BMC domain proteins have been conducted. They have led to the conclusion that bacterial microcompartments are widespread among bacteria. The field of bacterial microcompartment research is now ripe for more complex bioinformatics queries based on genomic context and comparative genomics methods.

BMC Proximity and Microcompartment-Related Enzyme Families

The tendency of proteins to be functionally related to each other when they are encoded near each other on a bacterial chromosome—or fused to each other in the case of multidomain proteins—is well established (35, 45). The known microcompartment operons support this finding. Enzyme and protein families showing a tendency to be encoded near BMC shell proteins might therefore by inferred to be involved in compartmentalized functions in certain bacteria. Here we provide a systematic analysis based on that idea. Searching across the known databases of sequenced bacterial genomes, we identified all genes within 10 coding regions (either upstream or downstream) of a BMC domain protein and clustered the resulting set of roughly 8600 distinct proteins (Figure 7). Such an automatic analysis leads to the spurious identification of many individual proteins that might not be truly related to microcompartments. However, the large body of sequence data makes it possible to identify a large of number of protein families that appear near BMC genes in many different genomes. These BMC-proximal protein families can be ascribed microcompartment-related functions with a reasonable statistical likelihood. The list of frequently identified protein families includes several known to be associated with previously characterized microcompartments, along with numerous proteins and enzymes for which roles in microcompartment function might be predicted. They include enzymes involved in butyrate and malate metabolism and proteins involved in iron uptake, nitrate sensing, and drug efflux, among others (Supplemental Tables 2 and 3). These predictions may be useful in guiding future experimental studies.

Genomic Signature for the Discovery of Unrelated Compartments

The varied inclusions visualized in bacteria suggest that other kinds of microcompartments, evolutionarily unrelated to the type discussed in this review (i.e., the BMC type), could exist in nature. A recent bioinformatics study sought to predict novel, uncharacterized microcompartments by analyzing genomic context (5). BMC-type microcompartment operons typically present a simple but rather distinct signature: multiple paralogs of a small protein (in this case the BMC domain) encoded in close proximity to each other (Figure 2). This pattern was used as a query to analyze all known protein families in search of any that tends to follow a similar genomic signature. The analysis retrieved numerous candidates, including a mixture of characterized and uncharacterized protein families (5). Among the identified protein families whose functions had been previously characterized, a majority were confirmed to be involved in forming large protein assemblies of various types (e.g., shells, filaments, and tubular assemblies). A minority of the predictions were of families not known to form large assemblies, and likely represent false positive predictions. The favorable outcome among the characterized families suggests that many of the uncharacterized protein families identified are likewise components of yet unknown bacterial assemblies. Some could represent compartmentalized systems, but experimental studies on the numerous protein families identified are required. Supporting experimental evidence for self-assembly was provided for one family (IPR009482) present in *Archaeoglobus fulgidis* and in other archaea (5).

OPEN QUESTIONS AND FUTURE DIRECTIONS

Research on bacterial microcompartments is accelerating. Open questions abound. Elucidating the higher-level organization of microcompartments is paramount to understanding better their functions and mechanistic features. For the carboxysome, experiments indicate that CA (whose abundance is only about 80 copies per microcompartment) is physically bound to the interior of the shell (3, 64), and RuBisCO appears to be organized in interior layers (33, 36, 59). Further spatial details are murky, but experiments from independent groups have identified key protein-protein interactions in

beta-type carboxysomes. Those studies have focused attention on the protein CcmM as a central player in organizing multiple components (18, 43). CcmM has the remarkable property of showing homology to γ-type CAs in its N-terminal region (1) (despite lacking apparent enzymatic activity) and harboring near its C terminus typically three or four tandem sequence repeats homologous to the small subunit of RuBisCO (44). Both parts of CcmM could interact with other carboxysomal proteins through a kind of mimicry, but structural data are lacking. How other kinds of microcompartments might be organized and how their encapsulated enzymes might be targeted to the interior are not known, but the possibility of special targeting sequences on the enzymes has been put forward (71; T. Bobik, unpublished data). These sequences could interact with the interior surface of the shell. Experiments aimed at determining which side of the molecular layer of BMC proteins faces inward to interact with encapsulated enzymes, and which faces outward to the cytosol, have not been successful. Structural characteristics, combined with the behavior of the flexible C-terminal tails on some BMC domains, suggest that the side of BMC proteins bearing the prominent depression probably faces inward (69), though experimental confirmation is needed. The possibility that formation of the shell might require interactions with interior enzymes has been refuted. Both the carboxysome and the Pdu shell have been assembled following deletion of naturally encapsulated enzymes (7, 47). Because insertion of enzymes after shell assembly is viewed as problematic, it seems likely that interior enzymes assemble simultaneously with the shell when both components are present. Whether microcompartment structures are dynamic—e.g., whether they undergo controlled disassembly, division, or repair—is unknown.

Open questions also surround the evolutionary history of bacterial microcompartments. Architectural similarities to viral capsids are obvious, but no evidence of a common evolutionary origin for the two types of protein shells has been obtained. Likewise, an early claim of nucleic acid in carboxysomes (75), which would provide a further parallel to viruses, was later contradicted (32). Whether nucleic acids might play a role in some microcompartments is an open question; the recent report of a helix-turn-helix motif in the EutK shell protein (70) suggests a potential line of investigation. At the other end of the evolution spectrum, no proteins homologous to BMC proteins have been identified in eukaryotes. If carboxysomes were present in cyanobacteria at the time symbiosis gave rise to algae, then carboxysomes might be expected within algal chloroplasts. RuBisCO is organized in dense pyrenoid or central bodies in some algae, but no biochemical or genomic evidence for a shell has been found (10). Evolutionary connections between microcompartments and similar bodies from viruses and eukaryotes can therefore only be speculated at this time.

In future protein engineering work, structural data on microcompartment shells will guide studies aimed at addressing fundamental mechanistic questions and key experimental obstacles. Carboxysomes with altered pores could address some of the questions articulated above regarding molecular transport across the shell. By engineering stabilized shells, researchers could improve the purification and characterization of microcompartments, enabling more definitive compositional analyses and in vitro activity studies. Finally, systems for targeting enzymes to microcompartment interiors could lead to novel biotechnology applications (47). Diverse bacterial microcompartments, some yet to be discovered, should offer many fruitful lines of investigation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Thomas Bobik and Yingssu Tsai for discussions and comments on the manuscript. We thank Richard Llewellyn, Robert Jefferson, and Tobias Sayre for contributions to the bioinformatics analysis. The work was supported by NIH grant AI081146, NSF grant MCB-0843065, and the BER program of the DOE Office of Science.

Glossary

Bacterial microcompartment	a widely distributed family of typically polyhedron-shaped bacterial organelles that have an outer protein shell and enclose two or more sequentially acting enzymes
Bacterial microcompartment (BMC) protein	the family of conserved shell proteins that defines the bacterial microcompartment family of organelles; they form hexamers that constitute the building blocks of the shell
RuBisCO	ribulose bisphosphate carboxylase/oxygenase
Carboxysome	a particular type of bacterial microcompartment that enhances CO_2 fixation
CA	carbonic anhydrase
3-PG	3-phosphoglycerate
Pdu	propanediol utilization
1,2-PD	1,2-propanediol
Pdu microcompartment	a bacterial microcompartment that metabolizes 1,2-PD without allowing the escape of the propionaldehyde intermediate
Eut	ethanolamine utilization
Eut microcompartment	a bacterial microcompartment that metabolizes ethanolamine without allowing the escape of the acetaldehyde intermediate

LITERATURE CITED

1. Alber BE, Ferry JG. A carbonic anhydrase from the archaeon *Methanosarcina thermophila*. Proc Natl Acad Sci USA. 1994; 91:6909–13. [PubMed: 8041719]
2. Badger MR, Price GD. CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. J Exp Bot. 2003; 54:609–22. [PubMed: 12554704]
3. Baker SH, Williams DS, Aldrich HC, Gambrell AC, Shively JM. Identification and localization of the carboxysome peptide Csos3 and its corresponding gene in *Thiobacillus neapolitanus*. Arch Microbiol. 2000; 173:278–83. [PubMed: 10816046]
4. Bazylinski DA, Frankel RB. Magnetosome formation in prokaryotes. Nat Rev Microbiol. 2004; 2:217–30. [PubMed: 15083157]
5. Beeby M, Bobik TA, Yeates TO. Exploiting genomic patterns to discover new supramolecular protein assemblies. Protein Sci. 2009; 18:69–79. [PubMed: 19177352]
6. Berg HC, Purcell EM. Physics of chemoreception. Biophys J. 1977; 20:193–219. [PubMed: 911982]
7. Bobik TA. Polyhedral organelles compartmenting bacterial metabolic processes. Appl Microbiol Biotechnol. 2006; 70:517–25. [PubMed: 16525780]
8. 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 B_{12}-dependent 1, 2-propanediol degradation. J Bacteriol. 1999; 181:5967–75. [PubMed: 10498708]

9. Brinsmade SR, Paldon T, Escalante-Semerena JC. Minimal functions and physiological conditions required for growth of *Salmonella enterica* on ethanolamine in the absence of the metabolosome. J Bacteriol. 2005; 187:8039–46. [PubMed: 16291677]
10. Burey SC, Poroyko V, Ergen ZN, Fathi-Nejad S, Schuller C, et al. Acclimation to low [CO(2)] by an inorganic carbon-concentrating mechanism in *Cyanophora paradoxa*. Plant Cell Environ. 2007; 30:1422–35. [PubMed: 17897412]
11. Cai F, Heinhorst S, Shively JM, Cannon GC. Transcript analysis of the *Halothiobacillus neapolitanus cso* operon. Arch Microbiol. 2008; 189:141–50. [PubMed: 17899012]
12. Cai F, Menon BB, Cannon GC, Curry KJ, Shively JM, Heinhorst S. The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO_2 leakage barrier. PLoS One. 2009; 4:e7521. [PubMed: 19844578]
13. Cannon GC, Bradburne CE, Aldrich HC, Baker SH, Heinhorst S, Shively JM. Microcompartments in prokaryotes: carboxysomes and related polyhedra. Appl Environ Microbiol. 2001; 67:5351–61. [PubMed: 11722879]
14. Cannon GC, Heinhorst S, Bradburne CE, Shively JM. Carboxysome genomics: a status report. Funct Plant Biol. 2002; 29:175–82.
15. Cannon GC, Shively JM. Characterization of homogenous preparation of carboxysomes from *Thiobacillus neapolitanus*. Arch Microbiol. 1983; 134:52–59.
16. 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]
17. Cheng S, Liu Y, Crowley CS, Yeates TO, Bobik TA. Bacterial microcompartments: their properties and paradoxes. Bioessays. 2008; 30:1084–95. [PubMed: 18937343]
18. 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]
19. Crowley CS, Sawaya MR, Bobik TA, Yeates TO. Structure of the PduU shell protein from the Pdu microcompartment of *Salmonella*. Structure. 2008; 16:1324–32. [PubMed: 18786396]
20. Dou Z, Heinhorst S, Williams EB, Murin CD, Shively JM, Cannon GC. CO_2 fixation kinetics of *Halothiobacillus neapolitanus* mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for CO_2. J Biol Chem. 2008; 283:10377–84. [PubMed: 18258595]
21. Drews G, Niklowitz W. Cytology of cyanophycea. II Centroplasm and granular inclusions of *Phormidium uncinatum*. Arch Mikrobiol. 1956; 24:147–62. [PubMed: 13327992]
22. Dryden KA, Crowley CS, Tanaka S, Yeates TO, Yeager M. Two-dimensional crystals of carboxysome shell proteins recapitulate the hexagonal packing of three-dimensional crystals. Protein Sci. 2009; 18:2629–35. [PubMed: 19844993]
23. English RS, Lorbach SC, Qin X, Shively JM. Isolation and characterization of a carboxysome shell gene from *Thiobacillus neapolitanus*. Mol Microbiol. 1994; 12:647–54. [PubMed: 7934888]
24. Forouhar F, Kuzin A, Seetharaman J, Lee I, Zhou W, et al. Functional insights from structural genomics. J Struct Funct Genomics. 2007; 8:37–44. [PubMed: 17588214]
25. Gantt E, Conti SF. Ultrastructure of blue-green algae. J Bacteriol. 1969; 97:1486–93. [PubMed: 5776533]
26. Havemann GD, Bobik TA. Protein content of polyhedral organelles involved in coenzyme B_{12}-dependent degradation of 1,2-propanediol in *Salmonella enterica* serovar Typhimurium LT2. J Bacteriol. 2003; 185:5086–95. [PubMed: 12923081]
27. 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; 184:1253–61. [PubMed: 11844753]
28. Heinhorst, S.; Cannon, GC.; Shively, JM. Carboxysomes and carboxysome-like inclusions. In: Shively, JM., editor. Complex Intracellular Structures in Prokaryotes. Berlin: Springer-Verlag; 2006. p. 141-65.
29. Heinhorst S, Williams EB, Cai F, Murin CD, Shively JM, Cannon GC. Characterization of the carboxysomal carbonic anhydrase CsoSCA from *Halothiobacillus neapolitanus*. J Bacteriol. 2006; 188:8087–94. [PubMed: 17012396]

30. Heldt D, Frank S, Seyedarabi A, Ladikis D, Parsons JB, et al. Structure of a trimeric bacterial microcompartment shell protein, EtuB, associated with ethanol utilization in *Clostridium kluyveri*. Biochem J. 2009; 423:199–207. [PubMed: 19635047]
31. Hess WR, Rocap G, Ting CS, Larimer F, Stilwagen S, et al. The photosynthetic apparatus of *Prochlorococcus*: insights through comparative genomics. Photosynth Res. 2001; 70:53–71. [PubMed: 16228362]
32. Holthuijzen YA, Maathuis FJM, Kuenen JG, Konings RNH, Konings WN. Carboxysomes of *Thiobacillus neapolitanus* do not contain extrachromosomal DNA. FEMS Microbiol Lett. 1986; 35:193–98.
33. Holthuijzen YA, van Breemen JFL, Konings WN, van Bruggen EFJ. Electron microscopic studies of carboxysomes of *Thiobacillus neapolitanus*. Arch Microbiol. 1986; 144:258–62.
34. Holthuijzen YA, van Breemen JFL, Kuenen JG, Konings WN. Protein composition of the carboxysomes of *Thiobacillus neapolitanus*. Arch Microbiol. 1986; 144:398–404.
35. Huynen M, Snel B, Lathe W, Bork P. Exploitation of gene context. Curr Opin Struct Biol. 2000; 10:366–70. [PubMed: 10851194]
36. Iancu CV, Ding HJ, Morris DM, Dias DP, Gonzales AD, 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. Jain R, Rivera MC, Lake JA. Horizontal gene transfer among genomes: the complexity hypothesis. Proc Natl Acad Sci USA. 1999; 96:3801–6. [PubMed: 10097118]
38. Jensen GJ, Briegel A. How electron cryotomography is opening a new window onto prokaryotic ultrastructure. Curr Opin Struct Biol. 2007; 17:260–67. [PubMed: 17398087]
39. Kerfeld CA, Sawaya MR, Tanaka S, Nguyen CV, Phillips M, et al. Protein structures forming the shell of primitive bacterial organelles. Science. 2005; 309:936–38. [PubMed: 16081736]
40. Klein MG, Zwart P, Bagby SC, Cai F, Chisholm SW, et al. Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport. J Mol Biol. 2009; 392:319–33. [PubMed: 19328811]
41. 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]
42. Lawrence JG, Roth JR. Selfish operons: Horizontal transfer may drive the evolution of gene clusters. Genetics. 1996; 143:1843–60. [PubMed: 8844169]
43. 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]
44. Ludwig M, Sultemeyer D, Price GD. Isolation of ccmKLMN genes from the marine cyanobacterium, *Synechococcus* sp PCC7002 (Cyanophyceae), and evidence that CcmM is essential for carboxysome assembly. J Phycol. 2000; 36:1109–19.
45. Marcotte EM, Pellegrini M, Ng HL, Rice DW, Yeates TO, Eisenberg D. Detecting protein function and protein-protein interactions from genome sequences. Science. 1999; 285:751–53. [PubMed: 10427000]
46. Marcus Y, Berry JA, Pierce J. Photosynthesis and photorespiration in a mutant of the cyanobacterium *Synechocystis* PCC 6803 lacking carboxysomes. Planta. 1992; 187:511–16.
47. 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]
48. Penrod JT, Roth JR. Conserving a volatile metabolite: a role for carboxysome-like organelles in *Salmonella enterica*. J Bacteriol. 2006; 188:2865–74. [PubMed: 16585748]
49. Peters KR. Characterization of a phage-like particle from cells of *Nitrobacter*. II Structure and size. Arch Microbiol. 1974; 97:129–40. [PubMed: 4836296]
50. Price GD, Badger MR. Expression of human carbonic anhydrase in the cyanobacterium *Synechococcus* PCC7942 creates a high CO_2-requiring phenotype: evidence for a central role for carboxysomes in the CO_2 concentrating mechanism. Plant Physiol. 1989; 91:505–13. [PubMed: 16667062]

51. 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; 175:2871–79. [PubMed: 8491708]
52. Raushel FM, Thoden JB, Holden HM. Enzymes with molecular tunnels. Acc Chem Res. 2003; 36:539–48. [PubMed: 12859215]
53. Reinhold L, Kosloff R, Kaplan A. A model for inorganic carbon fluxes and photosynthesis in cyanobacterial carboxysomes. Can J Bot. 1991; 69:984–88.
54. Reinhold L, Zviman X, Kaplan A. A quantitative model for carbon fluxes and photosynthesis in cyanobacteria. Plant Physiol Biochem. 1989; 27:945–54.
55. Rondon MR, Kazmierczak R, Escalante-Semerena JC. Glutathione is required for maximal transcription of the cobalamin biosynthetic and 1,2-propanediol utilization (*cob/pdu*) regulon and for the catabolism of ethanolamine, 1,2-propanediol, and propionate in *Salmonella typhimurium* LT2. J Bacteriol. 1995; 177:5434–39. [PubMed: 7559326]
56. Sagermann M, Ohtaki A, Nikolakakis K. Crystal structure of the EutL shell protein of the ethanolamine ammonia lyase microcompartment. Proc Natl Acad Sci USA. 2009; 106:8883–87. [PubMed: 19451619]
57. Sampson EM, Bobik TA. Microcompartments for B12-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate. J Bacteriol. 2008; 190:2966–71. [PubMed: 18296526]
58. Satoh R, Himeno M, Wadano A. Carboxysomal diffusion resistance to ribulose 1,5-bisphosphate and 3-phosphoglycerate in the cyanobacterium *Synechococcus* PCC7942. Plant Cell Physiol. 1997; 38:769–75.
59. Schmid MF, Paredes AM, Khant HA, Soyer F, Aldrich HC, et al. Structure of *Halothiobacillus neapolitanus* carboxysomes by cryo-electron tomography. J Mol Biol. 2006; 364:526–35. [PubMed: 17028023]
60. Seedorf H, Fricke WF, Veith B, Bruggemann H, Liesegang H, et al. The genome of *Clostridium kluyveri*, a strict anaerobe with unique metabolic features. Proc Natl Acad Sci USA. 2008; 105:2128–33. [PubMed: 18218779]
61. Shively, JM., editor. Microbiology Monographs: Complex Intracellular Structures in Prokaryotes. Berlin/Heidelberg: Springer; 2006.
62. Shively JM, Ball F, Brown DH, Saunders RE. Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of *Thiobacillus neapolitanus*. Science. 1973; 182:584–86. [PubMed: 4355679]
63. Shively JM, Decker GL, Greenawalt JW. Comparative ultrastructure of the thiobacilli. J Bacteriol. 1970; 101:618–27. [PubMed: 5413830]
64. So AK, Espie GS, Williams EB, Shively JM, Heinhorst S, Cannon GC. A novel evolutionary lineage of carbonic anhydrase (epsilon class) is a component of the carboxysome shell. J Bacteriol. 2004; 186:623–30. [PubMed: 14729686]
65. Steinbuchel A, Aerts K, Babel W, Follner C, Liebergesell M, et al. Considerations on the structure and biochemistry of bacterial polyhydroxyalkanoic acid inclusions. Can J Microbiol. 1995; 41(Suppl 1):94–105. [PubMed: 7606669]
66. 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]
67. Sutter M, Boehringer D, Gutmann S, Gunther S, Prangishvili D, et al. Structural basis of enzyme encapsulation into a bacterial nanocompartment. Nat Struct Mol Biol. 2008; 15:939–47. [PubMed: 19172747]
68. Tanaka S, Kerfeld CA, Sawaya MR, Cai F, Heinhorst S, et al. Atomic-level models of the bacterial carboxysome shell. Science. 2008; 319:1083–86. [PubMed: 18292340]
69. Tanaka S, Sawaya MR, Phillips M, Yeates TO. Insights from multiple structures of the shell proteins from the beta-carboxysome. Protein Sci. 2009; 18:108–20. [PubMed: 19177356]
70. Tanaka S, Sawaya MR, Yeates TO. Structure and mechanisms of a protein-based organelle in *Escherichia coli*. Science. 2010; 327:81–84. [PubMed: 20044574]

71. 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]

72. Tsai Y, Sawaya MR, Cannon GC, Cai F, Williams EB, et al. Structural analysis of CsoS1A and the protein shell of the *Halothiobacillus neapolitanus* carboxysome. PLoS Biol. 2007; 5:1345–54.

73. Tsai Y, Sawaya MR, Yeates TO. Analysis of lattice-translocation disorder in the layered hexagonal structure of carboxysome shell protein CsoS1C. Acta Crystallogr D. 2009; 65:980–88. [PubMed: 19690376]

74. Wackett LP, Frias JA, Seffernick JL, Sukovich DJ, Cameron SM. Genomic and biochemical studies demonstrating the absence of an alkane-producing phenotype in *Vibrio furnissii* M1. Appl Environ Microbiol. 2007; 73:7192–98. [PubMed: 17921268]

75. Westphal K, Bock E, Cannon G, Shively JM. Deoxyribonucleic acid in *Nitrobacter* carboxysomes. J Bacteriol. 1979; 140:285–88. [PubMed: 227833]

76. Wright ER, Schooler JB, Ding HJ, Kieffer C, Fillmore C, et al. Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. EMBO J. 2007; 26:2218–26. [PubMed: 17396149]

77. Yeates TO, Kerfeld CA, Cannon GC, Heinhorst S, Shively JM. Protein-based organelles in bacteria: carboxysomes and related microcompartments. Nat Rev Microbiol. 2008; 6:681–91. [PubMed: 18679172]

78. Yeates TO, Tsai Y, Tanaka S, Sawaya MR, Kerfeld CA. Self-assembly in the carboxysome: a viral capsid-like protein shell in bacterial cells. Biochem Soc Trans. 2007; 35:508–11. [PubMed: 17511640]

SUMMARY POINTS

1. Many bacteria have sophisticated proteinaceous inclusions that serve as organelles for specific metabolic pathways.
2. The bacterial microcompartment family of organelles is defined by the presence of the conserved BMC family of shell proteins, a few thousand of which assemble to form the outer shell of the microcompartment.
3. Bacterial microcompartment organelles are widely distributed across the bacterial kingdom, where they have diversified to carry out varied metabolic pathways, typically by encapsulating two or a few sequentially acting enzymes.
4. Three-dimensional structural data for numerous BMC shell proteins indicate a high level of mechanistic sophistication and a complex evolutionary history. The body of structural data illuminates key principles of shell assembly and provides a framework for understanding molecular transport through pores in the protein shell.
5. Bioinformatic analyses, including genomic context methods, suggest that bacterial microcompartments exist with a wide variety of metabolic functions, and that other kinds of microcompartments, evolutionarily unrelated to the BMC family, might also exist.
6. Numerous key questions remain, particularly regarding the higher-level organization of the multiple shell and enzymatic components within a microcompartment.

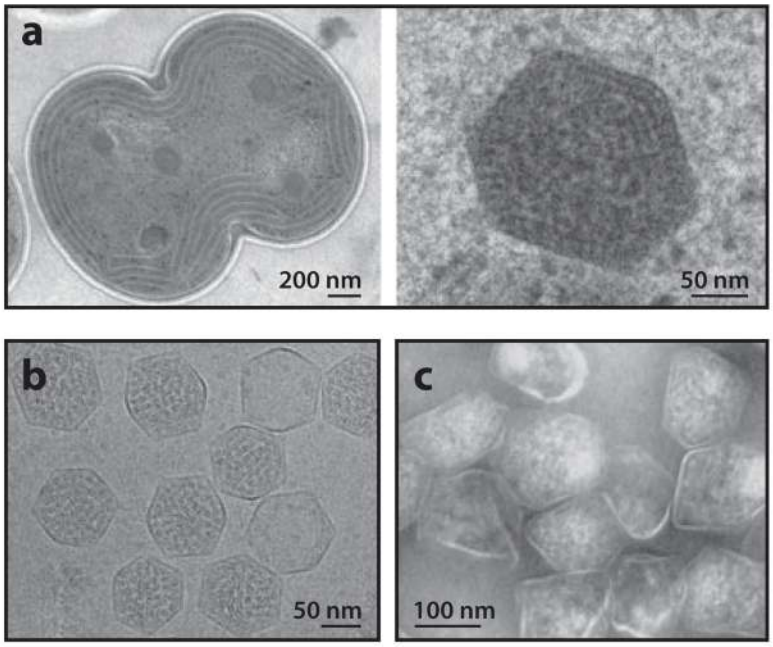

Figure 1. Electron micrographs of various bacterial microcompartments

(*a*) Transmission electron micrographs showing (*left*) a section through a dividing cyanobacterial cell (*Synechocystis* sp. PCC 6803) and (*right*) an enlargement of a single carboxysome on the right (courtesy of Wim Vermaas) (adapted from Reference 39). (*b*) Purified carboxysomes from *Halothiobacillus neapolitanus* (sample courtesy of Sabine Heinhorst and Gordon Cannon, image courtesy of Kelly Dryden and Mark Yeager). (*c*) Isolated Pdu microcompartments from *Salmonella enterica* serovar Typhimurium LT2 (courtesy of Thomas Bobik).

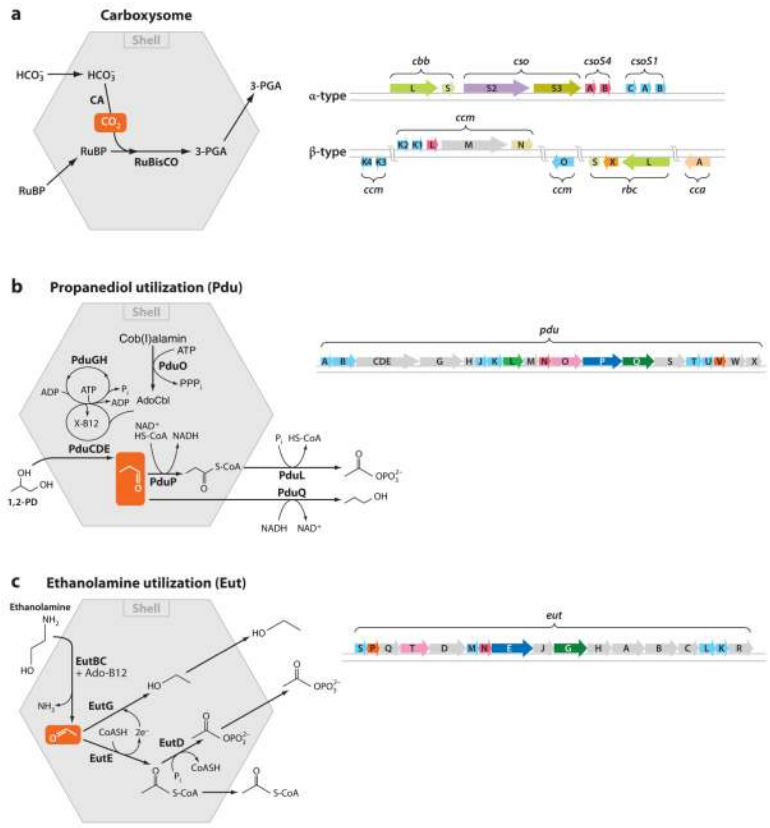

Figure 2.

Gene organization and proposed metabolic pathways for three types of bacterial microcompartments. Genes are colored to indicate their homology. All BMC shell proteins are light blue. For each microcompartment, the key sequestered intermediate is boxed in orange. (*a*) Function of the carboxysome in enhancing CO_2 fixation. See text and Reference 28 for mechanistic details. Gene organizations for α- and β-carboxysomes are on the right. (*b*) A current model for the function of the propanediol utilization (Pdu) microcompartment in metabolizing 1,2-propanediol. See text and Reference 17 for mechanistic details. The gene organization for the *pdu* operon is shown on the right. (*c*) A hypothetical model for the metabolism of ethanolamine in the Eut microcompartment. See text and Reference 48 for mechanistic details. The gene organization of the *eut* operon is shown on the right.

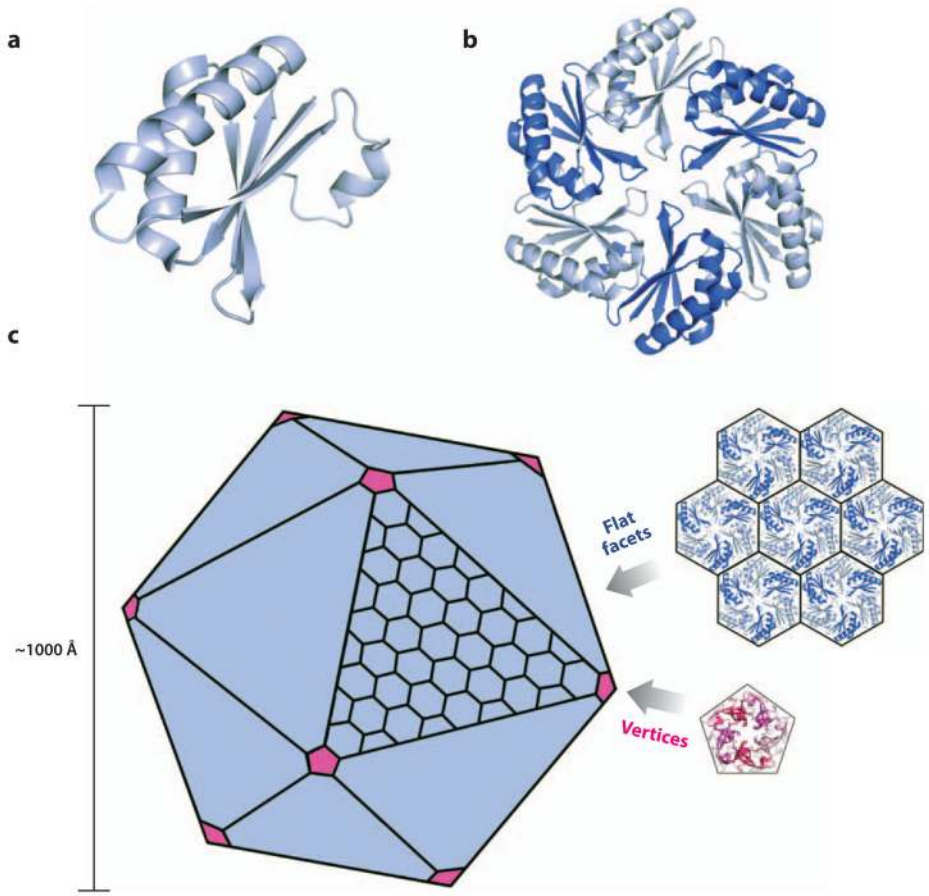

Figure 3.

Idealized model for assembly of the carboxysome and related bacterial microcompartments. (*a*) A ribbon diagram of a typical bacterial microcompartment (BMC) shell protein fold. (*b*) A hexameric assembly of a BMC protein in a ribbon diagram. (*c*) Hexameric building blocks of the BMC proteins can assemble into a molecular layer (*right*), which forms flat facets of the polyhedral shells of various bacterial microcompartments. The pentameric proteins (CcmL or CsoS4A) from the carboxysome (*bottom, right*) have been argued to form vertices of the icosahedral carboxysome (*left*) (68). The Pdu and Eut microcompartments are less geometrically regular than the carboxysome and are potentially more complex.

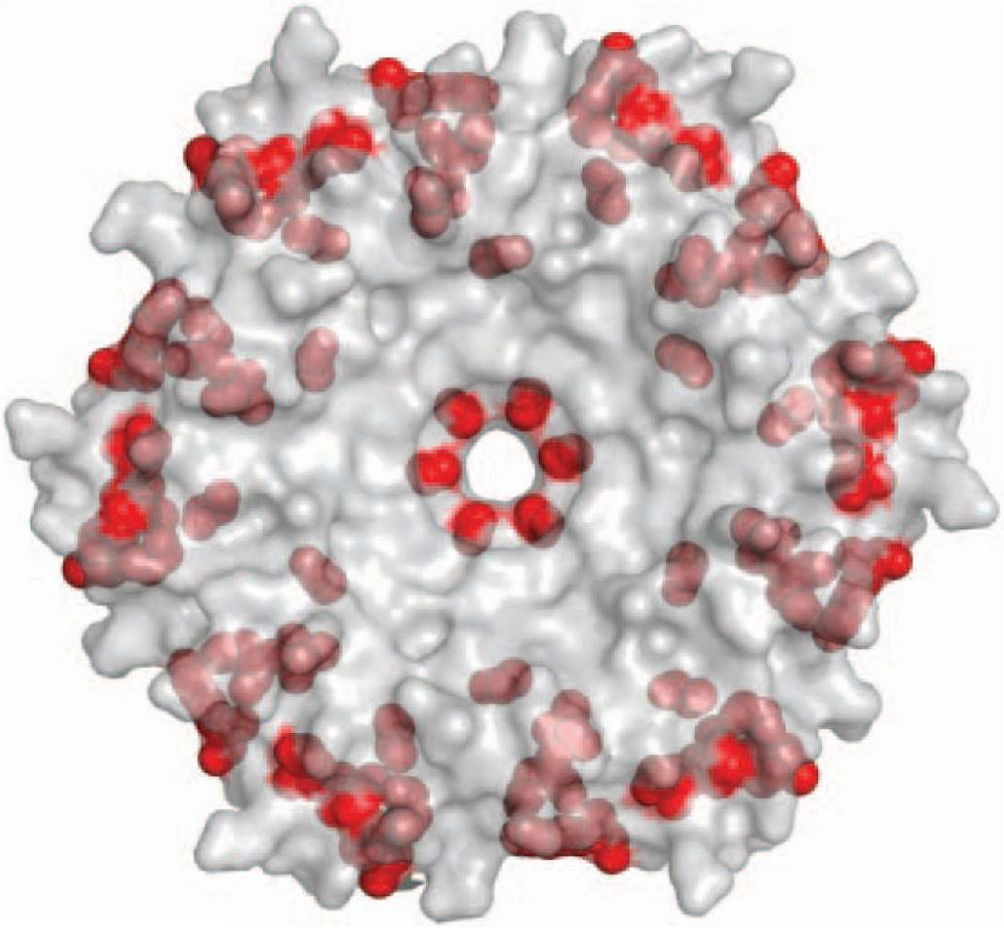

Figure 4.
Sequence conservation among diverse bacterial microcompartment (BMC) shell proteins. Conserved amino acid positions (*red*) were defined as those having sequence identity above 80% in an alignment of 2174 BMC sequences. Positions of high conservation occur mainly at the perimeter, where hexamers meet. In the CcmK1 protein, these residues include A4, G6, A19, D21, K25, V29, G38, G48, V50, V53, and G70; the conserved residue in the pore is glycine G38.

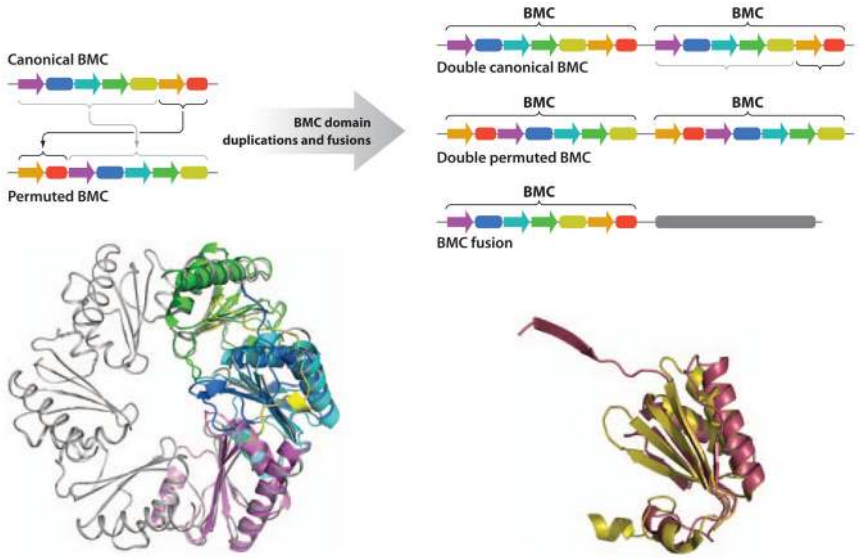

Figure 5.

Variations on the bacterial microcompartment (BMC) protein fold. (*Top*) Secondary structure schematics of BMC proteins in their various arrangements. Individual secondary structure elements are colored. The canonical BMCs include CcmK1/2/3/4, CsoS1A/B/C, PduA/J, and EutM. CcmO likely encodes tandem canonical BMC domains. The permuted, single-domain BMC proteins include PduU and EutS. (*Bottom, right*) The cores of canonical and permuted BMC proteins are in close agreement, as shown by a superposition of PduU (*salmon*) over CcmK2 (*yellow*). (*Bottom, left*) Both EutL and CsoS1D have permuted BMC domains in tandem, but their tertiary arrangements differ. Individual BMC domains are colored separately. When the N-terminal BMC domains (*blue*) of the two proteins are superimposed, their C-terminal domains (CsoS1D in *magenta* and EutL in *green*) adopt different positions in the hexamer. The linker regions between domains are colored yellow.

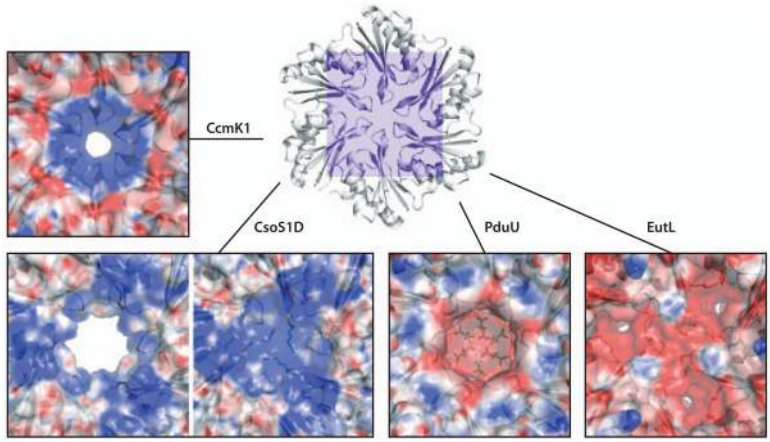

Figure 6.

Various pores of bacterial microcompartment (BMC) proteins. Top views of central pores from some representative BMC proteins colored by electrostatic potential (positive: *blue*; negative: *red*). Canonical BMC proteins (e.g., CcmK1 from the β-carboxysome) have a small pore at the sixfold, which has a diameter of 4 to 6 Å (39, 69, 72). A circularly permutated tandem BMC protein from the α-carboxysome, CsoS1D, adopts alternative conformations with open and closed pores (40). A circularly permutated BMC protein from the *pdu* microcompartment, PduU, revealed a totally closed pore (19). Whether this pore opens for transport is unknown. Another circularly permutated tandem BMC protein, EutL, has been observed in two distinct conformations, open (70) (not shown) and closed (40, 70) (shown). See Supplemental Figure 2 for additional views of the pores.

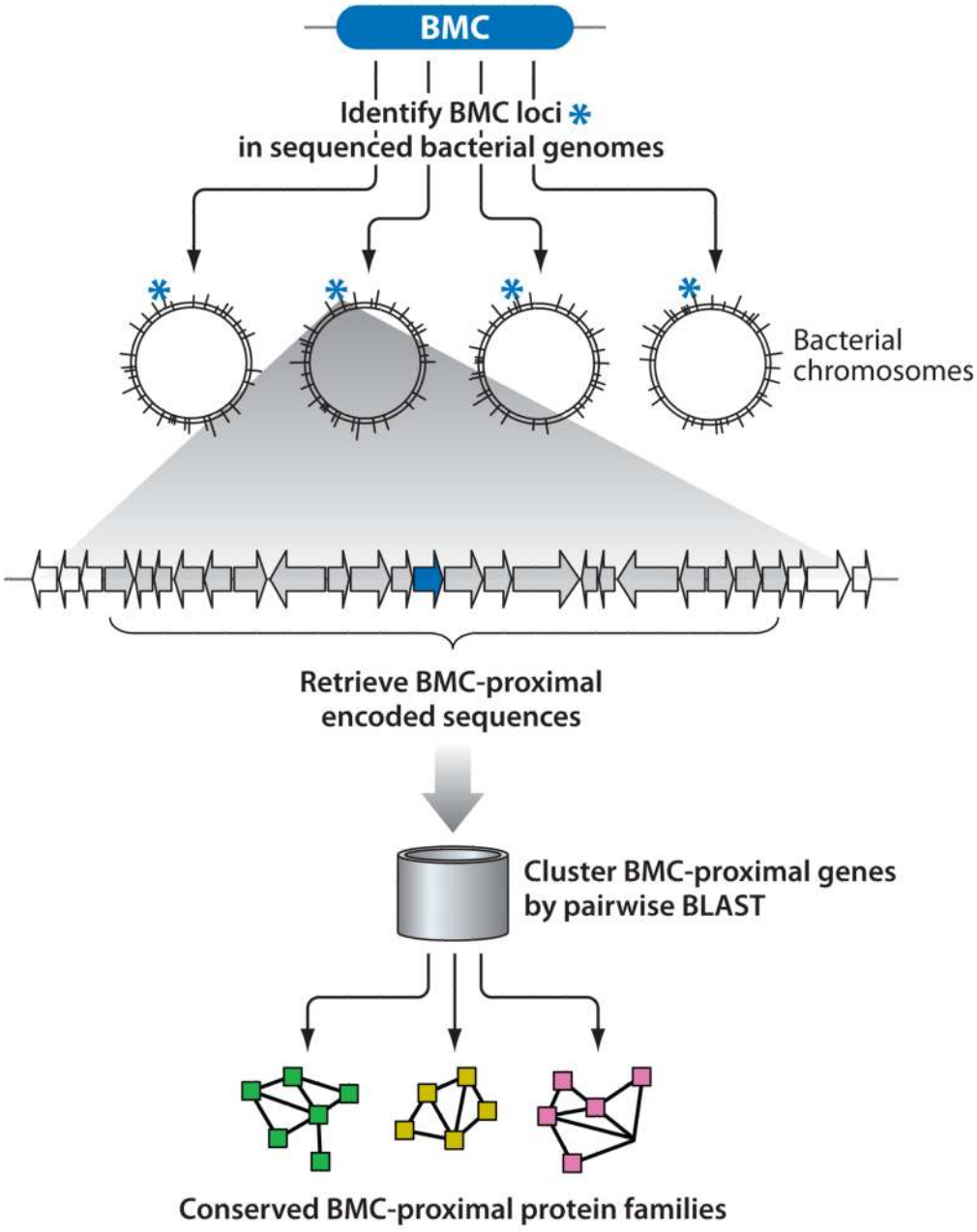

Figure 7.

Schematic for identifying conserved bacterial microcompartment (BMC)-proximal protein families. First, BMC homologues were retrieved from the NCBI database (CDD ID cl01982). Their respective chromosomal positions were located in fully sequenced bacterial genomes deposited into the EBI Integr8 database. For each chromosomal position identified, 10 BMC-proximal open reading frames in both the 5′ and 3′ direction were retrieved as candidate microcompartment-associated genes and subsequently assigned to homologous sequence clusters by performing a full pairwise BLAST analysis. The dominant clusters suggest likely microcompartment-associated functions (Supplemental Tables 2 and 3).