Engineering the Bacterial Microcompartment Domain for Molecular Scaffolding Applications

Eric J. Young1,2, Rodney Burton2, Jyoti P. Mahalik3,4, Bobby G. Sumpter3,4, Miguel Fuentes-Cabrera3,4, Cheryl A. Kerfeld1,2,5 and Daniel C. Ducat1,2,*

As synthetic biology advances the intricacy of engineered biological systems, the importance of spatial organization within the cellular environment must not be marginalized. Increasingly, biological engineers are investigating means to control spatial organization within the cell, mimicking strategies used by natural pathways to increase flux and reduce cross-talk. A modular platform for constructing a diverse set of defined, programmable architectures would greatly assist in improving yields from introduced metabolic pathways and increasing insulation of other heterologous systems. Here, we review recent research on the shell proteins of bacterial microcompartments and discuss their potential application as “building blocks” for a range of customized intracellular scaffolds. We summarize the state of knowledge on the self-assembly of BMC shell proteins and discuss future avenues of research that will be important to realize the potential of BMC shell proteins as predictively assembling and programmable biological materials for bioengineering.

Keywords: scaffold, synthetic biology, bacterial microcompartment, shell proteins, BMC, spatial organization, metabolic engineering, self-assembly

INTRODUCTION

With the advent of synthetic biology and recent advances in protein engineering, designing, constructing, and controlling biomolecule-based materials at the nanoscale is a rapidly developing field. Currently, there is a lack of modular building blocks for predictably fabricating custom sub-cellular architectures that can be subsequently programmed with precise functions (Figure 1A). Because of their self-assembly properties, proteins containing the pfamdomain 00936 (pfam0936) are promising building blocks that can be repurposed to design novel protein scaffolds in vivo, distinct from their natural propensity to assemble bacterial microcompartments (BMCs). In this perspective, we discuss current research on pfam0936 proteins in the context of their potential as a biological material for the construction of custom nano-architectures and intracellular scaffolds.

Organizing interrelated cellular components in time and space is crucial to increase efficiency of diverse cellular processes, including processes in metabolism, signaling, and division (Pawson and Scott, 1997; Agapakis et al., 2012). Typically, cells colocalize components of a shared pathway, conferring a host of benefits that include increased enzymatic intermediate flux and limited
pathway cross-talk (Good et al., 2011; Agapakis et al., 2012). Biological engineers have increasingly explored a variety of rational colocalization strategies to capitalize on such benefits. These engineered systems range in complexity from simple fusion proteins to dynamic artificial scaffolds (Conrado et al., 2008; Horn and Sticht, 2015; Myhrvold and Silver, 2015) or compartments (Giessen and Silver, 2016). As biologists move toward increasingly complex cellular engineering goals (Bashor et al., 2010), one challenge is designing sophisticated subcellular colocalization approaches that recapitulate the elegance of natural systems (Good et al., 2011). We focus here on molecular scaffold construction.

Polymerizing biomolecules represent ideal building blocks because they can self-assemble into higher-order arrangements in vivo. To date, DNA hybridization nanotechnology (e.g., DNA origami) is perhaps the best developed molecular building platform (Pinheiro et al., 2011). DNA architectures are especially flexible under non-physiological conditions where a nearly limitless array of architectures can be predictively constructed and controlled at scales approaching the sub-nanometer (Wilner et al., 2009; Fu et al., 2012; Funke and Dietz, 2016). Yet translating this technology to intracellular application has been partially constrained because the concentration of single-stranded nucleic acid building blocks and environmental properties important for nucleic acid folding (e.g., temperature, ions) are not easily manipulated in vivo (Pinheiro et al., 2011). While recent studies continue to advance the capability of nucleic acid assemblies achieved within the cell (Conrado et al., 2012; Myhrvold and Silver, 2015; Siu et al., 2015; Elbaz et al., 2016), proteins may offer another viable, naturally inspired solution. One early example of a synthetically designed scaffold was comprised of a string of protein–protein interaction domains that were used to recruit three cognate enzymes involved in the conversion of acetyl-CoA to mevalonate (Dueber et al., 2009). Co-recruitment of these enzymes substantially increased the mevalonate yield in vivo, yet only marginal improvements were reported when this approach was used for other metabolic pathways (Horn and Sticht, 2015). One proposed reason that this strategy is not widely successful is that this design lacks an inherent organized structure and may aggregate in unpredictable ways, hindering a rational design process (Lee et al., 2012). A genuinely modular protein-based scaffold would be composed of defined subunits which self-assemble into a concrete structure, which is dependent on the given application.

Toward this goal, engineering naturally found proteins which self-assemble into defined, nano to macromolecular architectures offers a powerful base to approach artificial scaffold construction (Howorka, 2011); the components of BMCs are particularly promising in this regard (Kerfeld and Erbilgin, 2015). In their native context, BMCs encapsulate related enzymes within a unique self-assembled protein shell (Kerfeld et al., 2010; Yeates et al., 2010; Axen et al., 2014; Kerfeld and Erbilgin, 2015). Many recent efforts have emphasized the engineering of BMCs to encapsulate new pathways for improved function of heterologous metabolic production (Bonacci et al., 2012; Choudhary et al., 2012; Lawrence et al., 2014; Cai et al., 2015; Gonzalez-Esquer et al., 2015; Lee et al., 2016; Quin et al., 2016; Baumgart et al., 2017; Liang et al., 2017; Slininger et al., 2017; Wagner et al., 2017; Young et al., 2017). While BMCs hold much promise as defined, engineered compartments, their pfam0936 domain containing shell proteins are unique on their own and possess the capacity to self-assemble into a variety of higher-order structures when expressed in isolation (Havemann et al., 2002; Kerfeld et al., 2005; Parsons et al., 2008, 2010; Dryden et al., 2009; Pitts et al., 2012; Lassila et al., 2014; Pang et al., 2014; Noël et al., 2015; Sutter et al., 2015; Held et al., 2016). Loci encoding BMC-domain proteins are found in at least 23 bacterial phyla, while each instance having a minimal of three unique pfam0936 domain containing proteins (Axen et al., 2014). This diversity likely includes many new “building blocks” for constructing a multitude of novel, programmable architectures, but unlocking the true potential of the pfam0936 domain will require a deeper understanding of the fundamentals governing self-assembly. We propose that the establishment of design principles—rules which result in a defined, predictable assembly—for the pfam0936 domain will provide the foundation for creating an array of nano to macromolecular structures. These designer structures may then be functionalized to cater to their individual application (Figure 1A). We discuss the promise and potential limitations of this strategy below.

### STRUCTURAL CHARACTERISTICS AND SELF-ASSEMBLY OF BMC-H PROTEINS

Numerous crystal structures of pfam0936-containing proteins have contributed to a detailed structural understanding of BMC shell proteins and models of how they “tile” into the facets of BMCs (Figure 1B) [summarized in Kerfeld and Erbilgin (2015)]. The signature domain of BMCs has little structural variation across the multitude of functionally distinct and distantly related BMCs, indicating a pivotal role in assembling the BMC shell (Crowley et al., 2010; Kinney et al., 2011). The main constituent of BMC shells are typically small (~100 amino acids) proteins containing the BMC domain (BMC-H) which form a ~70 Å hexagonal disk with distinct faces and a circular pore in the center (Figure 1B; Kerfeld et al., 2010; Yeates et al., 2010) other components of BMC shells are BMC-T (containing a tandem fused copy of pfam0936) and BMC-P (pfam03319) proteins (Kerfeld and Erbilgin, 2015), but are not a focus of this perspective. The concave side of BMC-H proteins features a surface depression that can harbor both flexible extensions of the N and C protein termini, whereas the convex side has varied electrostatic properties across homologs (Figure 1B; Kerfeld et al., 2010; Yeates et al., 2010).

In many crystal structures, a subset of residues found along the edge periphery mediate an inter-hexamer hydrogen bond network, permitting tiled assembly of conjoined arrays (Figure 1C; Kerfeld et al., 2005; Tanaka et al., 2008; Klein et al., 2009; Crowley et al., 2010; Takenoya et al., 2010; Kinney et al., 2011; Pitts et al., 2012; Cai et al., 2015). These edge residues (DxxK, RPH) are widely conserved throughout BMC-H proteins and thus, imply a crucial role in maintaining
FIGURE 1 | BMC-H attributes and potential as modular building blocks. (A) Cartoon schematic depicting distinct BMC hexamers (red, blue, and green) assembling into modular intracellular architectures that can recruit and concentrate cytosolic proteins (yellow and orange). (B) General features of BMC hexamers are highlighted through the example protein, PduA. A cross-section of a hexamer (right) illustrates the conserved shape and pore, while a hexamer is shown as part of a larger facet (left) that assembles through hexamer-hexamer contacts (box). (C) An expanded view of the interface between two BMC-H protein (PduA, PDB:3NGK), highlighting electrostatic interactions mediated by key residues (blue = positive). (D) Transmission electron microscopy of different assemblies of heterologously expressed BMC-H proteins in E. coli [PduA: Nanotubes (PDB:3NGK), MicH: Rosettes (PDB:5DJB), RmmH: Nanotubes (PDB:5L38), CcmK2: Lack of structure (PDB:4OX7)]. Inlet: CcmK4. Scale bar 250 nm. (E) Multiple sequence alignment of representative BMC-H proteins. Asterisks indicate key residues positioned at the hexamer-hexamer interface.
hexamer–hexamer interactions (Kerfeld and Erbilgin, 2015). For example, in the crystal structure of a PduA lattice—a canonical example—the antiparallel association of two adjacent lysine residues mediates the bulk of the inter-hexamer association (Figure 1C; Crowley et al., 2010; Pang et al., 2014; Sinha et al., 2014). Although, the buried interaction surface area at the hexamer-hexamer interface is typically less than other protein-protein interfaces, it is likely that the multiplicative nature of the interaction (1 hexamer surrounded by six others) provides sufficient cooperativity to permit higher-order arrays (Crowley et al., 2010). Since tiling behavior with consistent inter-hexamer distances has been observed in BMC-H sheets by high-resolution microscopy techniques (Dryden et al., 2009; Sutter et al., 2015), the interface observed in crystals is likely physiologically relevant.

The flexibility of high-order formation of BMC-H homologs begins to take shape outside the confining context of a crystalline array. Many distinct architectures can be formed by purified BMC-H proteins in vitro, including: 100 nm spheroids (Kerfeld et al., 2005; Keeling et al., 2014), extended nanotubes (Noël et al., 2015), and honeycombed tiles (Lassila et al., 2014). The macromolecular assembly behavior and the formation of such high-order structures are influenced by pH and ionic strength (Dryden et al., 2009; Noël et al., 2015; Jorda et al., 2016). Similarly, overexpression of BMC-H proteins in vivo leads to the self-assembly of a myriad of higher-order structures inside the cells, including: tubes (Parsons et al., 2010; Pang et al., 2014; Noël et al., 2015), filaments (Havemann et al., 2002; Parsons et al., 2008; Heldt et al., 2009; Pang et al., 2014), and other structures (Parsons et al., 2008, 2010; Pitts et al., 2012; Lin et al., 2014; Sutter et al., 2015; Held et al., 2016).

Because the methodology used to express BMC-H proteins varies among labs and studies (e.g., host, promoter strength, protein concentration, growth condition, sample preparation), it is not always clear if the distinct intracellular structures generated by BMC-H homologs in separate reports are due to intrinsic self-assembly properties, or the specific experimental conditions. Nonetheless, multiple lines of evidence suggest that properties of BMC-H proteins predispose them toward specific higher-order architectures (Pang et al., 2014; Sinha et al., 2014). To illustrate this point, we heterologously expressed a panel of BMC-H homologs from distinct BMCs under identical conditions in E. coli (Supplemental Material). We find that expression of BMC-H homologs PduA, MicH, RmmH, and CcmK2 form varied macromolecular assemblies in vivo (Figure 1D), generally in agreement with prior reports. PduA and RmmH form nanotube-like structures (Pang et al., 2014; Noël et al., 2015) and MicH (5815 BMC-H) forms “swiss roll” rosettes thought to be an extended sheet of rolled up protein (Sutter et al., 2015). Despite orderly tiling in CcmK2 and CcmK4 crystal structures (Synechococcus elongatus PCC 7942), over-expressing these homologs in E. coli does not lead to the formation of prominent macromolecular structures (Figure 1D). It is unclear if the absence of visible structures via transmission electron microscopy (TEM) thin section represents a lack of higher-order self-assembly, or if smaller assemblies are formed in the cytoplasm which are insufficiently discriminated from other cytoplasmic elements; as was previously proposed for other smaller BMC assemblies (Lassila et al., 2014). Collectively, it appears heterologously expressed BMC-H proteins form an assortment of in vivo assemblies, but exactly how the intrinsic features of each homolog (differences in primary structure) contribute to differences in self-assembly are currently unknown.

UNDERSTANDING THE DETERMINANTS OF MACROMOLECULAR ASSEMBLY OF BMC-H PROTEINS

As evident by the diversity of structures formed by BMC-H proteins (Figure 1D), there must be subtle primary structure differences that dictate changes in higher-order assembly. One region anticipated to influence assembly dynamics surrounds residues at the inter-hexamer junction; although there is strict conservation of some sidechains at this interface, some positions exhibit variance across homologs (Figure 1E; Cai et al., 2015). For example, both CcmK2 and CcmK4 contain arginine in comparison to the asparagine residues of PduA, RmmH or lysine of MicH (Figure 1E, red text). Supporting this hypothesis, experimental evidence generated by targeted amino acid substitutions at interface residues of PduA indicate they influence macromolecular assembly in vivo (Pang et al., 2014). Other studies of BMC-H proteins with modified hexamer-hexamer interface residues show they alter the formation of isolated BMCs (Sinha et al., 2014), size of tiled arrays in vitro (Sutter et al., 2015), and disrupt crystal packing contacts and orientation (Pang et al., 2014; Sinha et al., 2014).

In addition to residues at the hexamer-hexamer interface, other hexamer features may dictate the self-assembly behavior of BMC-H homologs. It has been well documented that the overall electrostatic surface profiles varies significantly among homologs (Figure 1D; Kerfeld et al., 2010; Kinney et al., 2011). Electrostatic differences—known to effect the self-assembly of proteins (Keskin et al., 2008)—could influence the preferred interaction orientation between BMC-H proteins, predisposing them to a particular assembly architecture. Besides the overall electrostatic profile, other unique regions in primary structure of BMC-H homologs could manipulate self-assembly; one such region is the variable C-terminal region (Figure 1E, boxed). Longer C-terminal extensions (Figure 1E, CcmK4) originally were hypothesized to interfere with lateral molecular tiling through steric clash, a hypothesis partially supported by the observation of that truncation mutants form hexamers which pack more tightly in crystal lattices (Tanaka et al., 2009). Some crystal forms of CcmK2 orthologs appear to have hexamers which are stacked upon each other (dodecamer) interacting through the C-terminal extensions (Tanaka et al., 2009; Samborska and Kimber, 2012). Although the physiological significance of these crystal contacts is uncertain, in vitro analysis of the molecular weight (Tanaka et al., 2009) and FRET interaction (Samborska and Kimber, 2012) of CcmK2 with modified C-termini provide some supporting evidence of a functional role. However, it
should be noted that C-terminal truncation does not disrupt the formation of heterologous BMC shell in the presence of other BMC shell components (Cai et al., 2016).

To establish detailed models for oligomeric BMC-H self-assembly, dynamic techniques that can interrogate the differences in the nucleation and expansion of shell protein arrays are required. High-speed atomic force microscopy (HS-AFM) is one emergent technique because of the high spatial and temporal resolution it affords. HS-AFM was used to capture the individual changes in BMC-H proteins association/dissociation rates into larger sheets (Sutter et al., 2015). Dynamic light scattering, and complimentary biophysical techniques can also quickly assess particle size (from single molecules to large assemblies) based on changes in the optical properties creating a more high-throughput pipeline to evaluate factors controlling assembly; as recently employed in a reengineered PduA variant (Jorda et al., 2016).

Computational frameworks and molecular dynamics offer another potentially powerful tool for predicting and understanding the behavior of BMC shell proteins. One recent study simulated the steps of BMC assembly by utilizing a computational model that varied the strength of hexamer-hexamer and hexamer-cargo affinity (Perlmutter et al., 2016). From this, two classes of BMC assembly emerged that proceeded through distinct hierarchies (Perlmutter et al., 2016). Another recent example used Monte Carlo simulations with a coarse-grained potential to study the 2D self-assembly of CcmK2 (Synechocystis sp PCC 6803) (Mahalik et al., 2016). In these simulations, 2D sheets were found to form rapidly after the association of an initial clustering of four hexamers, suggesting that self-assembly is rate-limited by a nucleation event (Mahalik et al., 2016). In turn, nucleation rates strongly depended on the concentration of hexamers and their relative 2D orientation upon collision (Mahalik et al., 2016).

To illustrate of how subtle differences in primary structure influences hexamers’ self-assembly, we performed preliminary simulations of the initial steps of the 3D self-assembly with two different BMC-H homologs (Figures 2A–C). Employing the Thomas-Dill (Thomas and Dill, 1996) coarse-grained potential—found to best approximate the fully atomic potential in Mahalik et al. (2016)—to compute the angular dependence in the potential of mean force (PMF) between two RmmH hexamers (Mycobacterium smegmatis) or two CcmK2 hexamers (S. elongatus 7942). The PMF was calculated as a function of the distance between the center of mass of each hexamer tile and the angle θ (defined in Figure 2A). CcmK2 and RmmH structures were obtained by relaxing the corresponding crystal structure in an aqueous environment (Supplemental Material). The angular dependence of PMF for CcmK2 and RmmH pairs is shown in Figure 2B: the PMF of RmmH has a clear minimum at θ ∼−45°, whereas CcmK2 is, by comparison, flatter. This minimum association angle of RmmH is consistent with the formation of regular repeating curved surface akin to the model created of in vivo/in vitro RmmH nanotubes (Noël et al., 2015). In contrast, in vitro macromolecular assemblies of CcmK2 orthologs (Keeling et al., 2014), depict flexibility in the overall macromolecular structure supporting a lack of defined interaction orientations. The differences in angular dependence of the PMF is further demonstrated if one uses different initial structures of the same BMC-H protein. This is illustrated in Figure 2C for CcmK2. In this figure, the PMFs of CcmK2 pairs formed from either the crystal or the relaxed structure are shown. The PMF of the crystalline pair has a clear angular minimum (θ ∼−80°), where the PMF of the relaxed pair lacks any clear minimum. This is notable because the backbone root mean-square deviation between the two structures only differs by 0.67 Å. The distinct PMFs are caused by differences in side-chain rotamers altered in the relaxation procedure. These small structural changes are enough to cause a difference of almost 1 K_BT, indicating that subtle structural changes which may occur between crystallization and in situ conditions can have a profound impact on self-assembly; also the overall lower PMF values for CcmK2 suggests a net weaker hexamer-hexamer interaction, consistent with other predictions (Perlmutter et al., 2016).

Although these observations of hexamer behavior in computational frameworks are preliminary, they are illustrative of the potential for in silico techniques to aid experimental methods in understanding and predicting behavior across several scales (from single molecules to large arrays). It is also noteworthy to mention the synergy between crystallography and electron microscopy for solving the macromolecular interactions in whole BMC assemblies (Sutter et al., 2017).

**FUNCTIONALIZATION OF EXISTING MACROMOLECULAR STRUCTURES**

The engineering of scaffolding platforms based on self-assembling protein modules can be considered a two-faced challenge: one side is predictably making a discrete structure, and the other is functionalizing the structure for a specific purpose. Ideally, functionalizing the surface of BMC-H protein architectures should be modular in itself, so that scaffold structures could be “repurposed” for new enzymes and pathways with minimal redesign. There are two ways in which functional proteins could be organized to BMC-H assemblies: through natural or synthetic motifs.

One approach to functionalizing heterologous assemblies is to use binding motifs BMCS natively employed to recruit cargo (Fan et al., 2012; Aussignargues et al., 2015). Frequently, BMC core proteins contain small peptides (~20 amino acids) as extensions of the N- or C-termini that are necessary for encapsulation (Choudhary et al., 2012; Kinney et al., 2012; Kim and Tullman-Ercek, 2014; Lawrence et al., 2014; Lin et al., 2014; Gonzalez-Esquer et al., 2015; Jakobson et al., 2015; Cai et al., 2016; Held et al., 2016; Quin et al., 2016; Wagner et al., 2017). Collectively known as “encapsulation peptides” (EPs), modeling (Fan et al., 2012; Kinney et al., 2012) and solution structures (Lawrence et al., 2014) adopt an alpha-helical conformation with an amphipathic charge distribution (Aussignargues et al., 2015). Although EPs vary widely in primary structure, it has been demonstrated that non-native EPs can interact with non-cognate BMCS (Jakobson et al., 2015) and amino acid
substitutions can alter the affinity for BMC encapsulation (Kim and Tullman-Ercek, 2014), likely through conservation of specific amphipathic characteristics. While promising, general use of EP motifs for predictive recruitment is currently impaired by uncertainty in the interface location and affinity EP-BMC component binding (Fan et al., 2012; Lawrence et al., 2014; Aussignargues et al., 2015).

An alternative strategy is appending natural or synthetically derived protein–protein interaction domains—known as adaptor domains—to BMC-H proteins. In this way, virtually any protein encoded with the cognate adaptor domain could be post-translationally concentrated, conferring specified enzymatic functions to the designer architecture. While this approach is potentially powerful, it must be determined if fusion of adaptor domains to BMC-H proteins will alter higher-order assembly. In published work, fusions to BMC-H proteins have ranged from small affinity tags (Havemann et al., 2002; Kerfeld et al., 2005; Dryden et al., 2009; Samborska and Kimber, 2012) to fluorescent proteins (~26 kDa; ~2X the size of a single pfam0936 domain), and these modified BMC-H proteins still incorporate into BMCs (Parsons et al., 2008, 2010; Savage et al., 2010; Cai et al., 2013, 2015; Cameron et al., 2013; Sun et al., 2016). Yet, fusions of certain characteristics (e.g., size, charge) could disrupt association through steric clash or compromised electrostatics, thereby changing higher-order behavior. Indeed, although a fluorescent protein fusion to major shell protein CcmK2 incorporates into functional BMCs (Cameron et al., 2013), an unmodified copy of the protein must also be present, as the fusion is unable to solely complement a full CcmK2 background; this is supported also by fluorescent protein fusions to CcmK2 aggregates and does not interact with other shell protein components (Lin et al., 2014). In contrast, it is notable that some shell proteins can still assemble functional BMCs without a native copy present (Parsons et al., 2010; Sun et al., 2016), suggesting that some fusions can be tolerated or other shell protein components can relieve the compromised function of the fusion (Chowdhury et al., 2016). So, it remains to be fully elucidated how specific fusions and how certain properties (e.g., fusion orientation — N- vs. C-termini — fusion size/charge) will alter self-assembly behavior.

FUTURE APPLICATIONS AND CLOSING REMARKS

The potential to form a spectrum of unique, functionalized architectures from a single, modulated building block can be anticipated to be a powerful tool for the future of synthetic biology and nanotechnology. While alternative approaches to build synthetic scaffolds have been proposed (see Introduction), often they do not possess defined geometry (Lee et al., 2012) or are not yet fully compatible for in vivo applications (Pinheiro et al., 2011). In a larger assembly, individual pfam0936 domains...
are structurally well-defined, which facilitates prediction of the relative position of the recruited proteins. This enables a design-test-build model of reiterative engineering, while also increasing the chance that successful designs from one pathway can be reimplemented for others.

Furthermore, the ability to construct different pfam0936 domain architectures raises the possibility that specific geometries can be chosen to tailor them for a given scaffolding application (Figure 2D). In the simplest case, scaffolding proteins to single hexamers or very small tiled arrays would be expected to concentrate up to 6 proteins per hexamer into cytosolic microdomains, akin to the earliest synthetic scaffold designs (Dueber et al., 2009). Stepping up in size, structures with 1D or 2D geometries—filaments or large planar sheets—could be more appropriate for the colocalization of signal-transduction or redox pathways; likened to a macromolecular switchboard (Good et al., 2011). More complex, sheltered architectures (e.g., nanotubes) would likely allow for increased efficiencies for metabolic pathways by enabling substrate channeling between a series of enzymes—an effective, but not yet fully understood, approach to increase pathway flux and insulation (Bauler et al., 2010; Idan and Hess, 2013; Castellana et al., 2014; Wheeldon et al., 2016). It is worthwhile to note that tubes, or other partially enclosed geometries, can realize roles akin to pathway encapsulation within a compartment, but would not require the engineering of small molecule-specific transporters or pores because of their access to the cytosol. As an added degree of utilization, it is feasible to imagine scaffolding orthogonal pathways on opposing faces of a given architecture (e.g., both on the outside and lumen of nanotubes). As the field determines the elements responsible for positioning BMCs inside cells (Savage et al., 2010), the ability to transpose these components to spatially localize entire pfam0936 domain assemblies arises, further increasing the engineering specificity for application.

Broadly speaking, we have outlined the promise and hurdles inherent to the use of pfam0936 domain containing proteins as building blocks for designer scaffold assemblies. One strength of using this domain to build ordered protein assemblies rests on the potential to construct a multitude of different intracellular assemblies from one common, protein domain. Even though the majority of published examples of higher-order pfam0936 assemblies are in prokaryotes, recent reports have documented engineering of several distinct higher-order structures in plant chloroplasts (Lin et al., 2014) indicating the potential for this approach across a wide diversity of organisms. Once design principles can be established, ways to further modify and cater the self-assembly process can be applied (Luo et al., 2016). Coupling this with more general knowledge on the principles of multiprotein complex formation (Ahnert et al., 2015; Murugan et al., 2015; Glover and Clark, 2016) the pfam0936 domain holds promise to aid in ushering in an era of true in vivo nanometer scale molecular engineering for the design of programmable architectures.

**AUTHOR CONTRIBUTIONS**

EY conceived the project ideas, conducted experiments and analyzed data, and wrote and edited the manuscript. RB, JM, BS, and MF-C contributed project ideas and experiments related to the computational work, and wrote and edited the manuscript. CK and DD conceived project ideas and wrote and edited the manuscript.

**FUNDING**

A portion of this work, i.e., simulations and writing of the manuscript, was conducted at the Center for Nanophase Materials Sciences, which is a DOE Office of Science User Facility. The rest of the work was performed at MSU-DOE PRL, funded through the Department of Energy (Grant: DE-FG02-91ER20021).

**ACKNOWLEDGMENT**

The authors would like to thank Dr. Alicia Withrow of the MSU Center for Advanced Microscopy for support with TEM.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2017.01441/full#supplementary-material

### REFERENCES

Agapakis, C. M., Boyle, P. M., and Silver, P. A. (2012). Natural strategies for the spatial optimization of metabolism in synthetic biology. *Nat. Chem. Biol.* 8, 527–535. doi: 10.1038/nchembio.975

Ahnert, S. E., Marsh, J. A., Hernández, H., Robinson, C. V., and Teichmann, S. A. (2015). Principles of assembly reveal a periodic table of protein complexes. *Science* 350:aaa2245. doi: 10.1126/science.aaa2245

Aussignargues, C., Paasch, B. C., Gonzalez-Esquer, R., Erbilgin, O., and Kerfeld, C. A. (2015). Bacterial microcompartment assembly: The key role of encapsulation peptides. *Commun. Integr. Biol.* 8:e1039755. doi: 10.1080/19420889.2015.1039755

Axen, S. D., Erbilgin, O., and Kerfeld, C. A. (2014). A taxonomy of bacterial microcompartment loci constructed by a novel scoring method. *PLoS Comput. Biol.* 10:e1003898. doi: 10.1371/journal.pcbi.1003898

Bashor, C. J., Horwitz, A. A., Peisajovich, S. G., and Lim, W. A. (2010). Rewiring cells: synthetic biology as a tool to interrogate the organizational principles of living systems. *Annu. Rev. Biophys.* 39, 515–537. doi: 10.1146/annurev.biophys.050708.133652

Bauler, P., Huber, G., Leyh, T., and McCammon, J. A. (2010). Channeling by proximity: the catalytic advantages of active site colocalization using brownian dynamics. *J. Phys. Chem. Lett.* 1, 1332–1335.

Baumgart, M., Huber, I., Abdollahzadeh, I., Gensch, T., and Frunzke, J. (2017). Bacterial microcompartment loci constructed by a novel scoring method. *Front. Microbiol.* [Epub ahead of print].

Dueber, J. E., Frunzke, J., Abdollahzadeh, I., and Gensch, T. (2009). The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2017.01441/full#supplementary-material
Good, M. C., Zalatan, J. G., and Lim, W. A. (2011). Scaffold proteins: hubs for controlling the flow of cellular information. Science 323, 680–686. doi: 10.1126/science.1198701

Havemann, G. D., Sampson, E. M., and Bobik, T. A. (2002). PduA is a shell protein of polyhedral organelles involved in coenzyme B12-dependent degradation of 1, 2-propanediol in Salmonella enterica serovar Typhimurium LT2. J. Bacteriol. 184, 1253–1261.

Held, M., Kolb, A., Perdue, S., Hsu, S. Y., Bloch, S. E., Quin, M. B., et al. (2016). Engineering formation of multiple recombinant Eut protein nanocompartments in E. coli. Sci. Rep. 6:24359. doi: 10.1038/srep24359

Heldt, H., Frank, S., Seyedarabi, A., Ladiakis, D., Parsons, J. B., Warren, J. M., et al. (2009). Structure of a trimeric bacterial microcompartment shell protein, EutB, associated with ethanol utilization in Clostridium kluveri. Biochem. J. 423, 199–207. doi: 10.1042/B20090780

Horn, A., and Stich, H. (2015). Synthetic protein scaffolds based on peptide motifs and cognate adaptor domains for improving metabolic productivity. Front. Bioeng. Biotechnol. 3:1. doi: 10.3389/fbioe.2015.00191

Howorka, S. (2011). Rationally engineering natural protein assemblies in nanobiotechnology. Curr. Opin. Biotechnol. 22, 485–491. doi: 10.1016/j.copbio.2011.05.003

Idan, O., and Hess, H. (2013). Engineering enzymatic cascades on nanoscale scaffolds. Curr. Opin. Biotechnol. 24, 606–611. doi: 10.1016/j.copbio.2013.01.003

Jakobson, C. M., Kim, E. Y., Slivinger, M. F., Chien, A., and Tullman-Ercek, D. (2015). 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. 290, 24519–24533. doi: 10.1074/jbc.M115.651919

Jorda, J., Leibly, D. J., Thompson, M. C., and Yeateset, T. O. (2016). Structure of a novel 13 nm dodecahedral nanocage assembled from a redesigned bacterial microcompartment shell protein. Chem. Commun. 52, 5041–5044. doi: 10.1039/c6cc0851h

Keeling, T. J., Samborska, B., Demers, R. W., and Kimber, M. S. (2014). Interactions and structural variability of β-carboxysomal shell protein CcmL. Photosynth. Res. 121, 125–133.

Kerfeld, C. A., and Erbilgin, O. (2015). Bacterial microcompartments and the modular construction of microbial metabolism. Trends Microbiol. 23, 22–34. doi: 10.1016/j.tim.2014.10.003

Kerfeld, C. A., Heinhorst, S., and Cannon, G. C. (2010). Bacterial microcompartments. Annu. Rev. Microbiol. 64, 391–408. doi: 10.1146/annurev.micro.112008.111423

Kerfeld, C. A., Sawaya, M. R., Tanaka, S., Nguyen, C. V., Phillips, M., Beeby, M., et al. (2005). Protein structures forming the shell of primitive bacterial organelles. Science 309, 936–938. doi: 10.1126/science.1113397

Keskin, O., Gursoy, A., Ma, B., and Nussinov, R. (2008). Principles of protein–protein interactions: what are the preferred ways for proteins to interact? Chem. Rev. 108, 1225–1244. doi: 10.1021/cr070409x

Kim, E. Y., and Tullman-Ercek, D. (2014). A rapid flow cytometry assay for the relative quantification of protein encapsulation into bacterial microcompartments. Biotechnol. J. 9, 348–354. doi: 10.1002/biot.201300391

Kinney, J. N., Axen, S. D., and Kerfeld, C. A. (2011). Comparative analysis of carboxysome shell proteins. Photosynth. Res. 109, 21–32. doi: 10.1007/s11120-011-9624-6

Kinney, J. N., Salmeen, A., Cai, F., and Kerfeld, C. A. (2012). Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly. J. Biol. Chem. 287, 17729–17736. doi: 10.1074/jbc.M112.355305

Klein, M. G., Zwart, P., Bagby, S. C., Cai, F., Chisholm, S. W., Heinhorst, S., et al. (2009). Identification and structural analysis of a novel carboxysomal shell protein with implications for metabolite transport. J. Mol. Biol. 392, 319–333. doi: 10.1016/j.jmb.2009.03.056

Lassila, J. K., Bernstein, S. L., Kinney, J. N., Axen, S. D., and Kerfeld, C. A. (2014). Assembly of robust bacterial microcompartment shells using building blocks from an organelle of unknown function. J. Mol. Biol. 426, 2217–2228. doi: 10.1016/j.jmb.2014.02.025

Lawrence, A. D., Frank, S., Newnham, S., Lee, M. J., Brown, I. R., Xue, W. F., et al. (2014). Solution structure of a bacterial microcompartment targeting peptide and its application in the construction of an ethanol bioreactor. ACS Synth. Biol. 3, 454–465. doi: 10.1021/sb4001118

Lee, H., DeLoache, W. C., and Dueber, J. E. (2012). Spatial organization of enzymes for metabolic engineering. Metab. Eng. 14, 242–251. doi: 10.1016/j.ymben.2011.09.003
Lee, M., DeLoache, W. C., and Dueber, J. E. (2016). Employing bacterial microcompartment technology to engineer a shell-free enzyme aggregate for enhance 1-2-propanediol production in Escherichia coli. Metab. Eng. 36, 48–56. doi: 10.1016/j.ymben.2016.02.007

Liang, M., Frank, S., Lünsdorf, H., Warren, M. J., and Prentice, M. B. (2017). Bacterial microcompartment-directed polyphosphate kinase promotes stable polyphosphate accumulation. E. coli. Biotechnol. J. 12.

Lin, M. T., Occhialini, A., Andralojc, P. J., Devonshire, J., Hines, K. M., Parry, M. A., et al. (2014). β-Carboxysomal proteins assemble into highly organized structures in Nicotianachloroplasts. Plant J. 79, 1–12. doi: 10.1111/tpj.12536

Luo, Q., Hou, C., Bai, Y., Wang, R., and Liu, J. (2016). Protein assembly: versatile approaches to construct highly ordered nanostructures. Chem. Rev. 116, 13571–13632. doi: 10.1021/acs.chemrev.6b00228

Mahalik, J. P., Brown, K. A., Cheng, X., and Fuentes-Cabrera, M. (2016). Theoretical study of the initial stages of self assembly of a carboxysome’s facet. ACS Nano 10, 5751–5758. doi: 10.1021/acsnano.5b07805

Murugan, A., Zou, J., and Brenner, M. P. (2015). Undesired usage and the robust self-assembly of heterogeneous structures. Nat. Commun. 6,6203. doi: 10.1038/ncomms7203

Myhrvold, C., and Silver, P. A. (2015). Using synthetic RNAs as scaffolds and regulators. Nat. Publ. Group 22, 8–10. doi: 10.1038/nsmb.2944

Noel, C. R., Cai, F., and Kerfeld, C. A. (2015). Purification and characterization of protein nanotubes assembled from a single bacterial microcompartment shell subunit. Adv. Mater. Interfaces. 3:1500295.

Pang, A., Frank, S., Brown, L., Warren, M. J., and Pickersgill, R. W. (2014). Structural insights into higher order assembly and function of the bacterial microcompartment protein PduA. J. Biol. Chem. 289, 23277–23284. doi: 10.1074/jbc.M114.592895

Parsons, J. B., Dinesh, S. D., Deery, E., Leech, H. K., Brindley, A. A., Heldt, D., et al. (2008). Biochemical and structural insights into bacterial organelle form and biogenesis. J. Biol. Chem. 283, 14366–14375. doi: 10.1074/jbc.M709214200

Parsons, J. B., Frank, S., Bhella, D., Liang, M., Prentice, M. B., Mulvihill, D. P., et al. (2010). Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. Mol. Cell 38, 305–315. doi: 10.1016/j.molcel.2010.04.008

Pawson, T., and Scott, J. D. (1997). Signaling through scaffold, anchoring, and adaptor proteins. Science 278, 2075–2080.

Perlmutter, J. D., Mohajerani, F., and Hagan, M. F. (2016). Many-molecule encapsulation by anicosahedral shell. Elife 5:14078. doi: 10.7554/eLife.14078

Pinheiro, A. V., Han, D., Shih, W. M., and Yan, H. (2011). Challenges and opportunities for structural DNA nanotechnology. Nat. Nanotechnol. 6, 763–772. doi: 10.1038/nnano.2011.187

Pitts, A. C., Tuck, L. R., Faulds-Pain, A., Lewis, R. J., and Marles-Wright, J. (2012). Structural insight into the Clostridium difficile ethanolamine utilisation microcompartment. PLoS ONE 7:e48360. doi: 10.1371/journal.pone.0048360

Quin, M. B., Perdue, S. A., Hsu, S. Y., and Schmidt-Dannert, C. (2016). Encapsulation of multiple cargo proteins within recombinant Eut nanocompartments. Appl. Microb. Biotechnol. 100, 9187–9200. doi: 10.1007/s00253-016-7737-8

Samborska, B., and Kimber, M. S. (2012). A dodecameric CcmK2 structure suggests β-carboxysomal shell facets have a double-layered organization. Structure 20, 1353–1362. doi: 10.1016/j.str.2012.05.013

Savage, D. F., Afonso, B., Chen, A. H., and Silver, P. A. (2010). Spatially ordered dynamics of the bacterial carbon fixation machinery. Science 327, 1258–1261. doi: 10.1126/science.1180690

Sinha, S., Cheng, S., Sung, Y. W., McNamara, D. E., Sawaya, M. R., Yeates, T. O., et al. (2014). Alanine scanning mutagenesis identifies an asparagine–arginine–lysine triad essential to assembly of the shell of the Pdu microcompartment. J. Mol. Biol. 426, 2328–2345.