

Lawrence Berkeley National Laboratory

Environ Genomics & Systems Bio

Title

In Vitro Assembly of Diverse Bacterial Microcompartment Shell Architectures

Permalink

<https://escholarship.org/uc/item/1x73m6fb>

Journal

Nano Letters, 18(11)

ISSN

1530-6984

Authors

Hagen, Andrew R
Plegaria, Jefferson S
Sloan, Nancy
[et al.](#)

Publication Date

2018-11-14

DOI

10.1021/acs.nanolett.8b02991

Peer reviewed



Published in final edited form as:

Nano Lett. 2018 November 14; 18(11): 7030–7037. doi:10.1021/acs.nanolett.8b02991.

In vitro assembly of diverse bacterial microcompartment shell architectures

Andrew R. Hagen^a, Jefferson S. Plegaria^b, Nancy Sloan^a, Bryan Ferlez^b, Clement Aussignargues^b, and Cheryl A. Kerfeld^{*a,b,c}

^aEnvironmental Genomics and Systems Biology and Molecular Biophysics and Integrated Bioimaging Divisions, Lawrence Berkeley National Laboratory, 1 Cyclotron Road Berkeley, CA 94720, USA

^bMSU-DOE Plant Research Laboratory, Michigan State University, 612 Wilson Road, East Lansing, MI 48824, USA

^cDepartment of Biochemistry and Molecular Biology, Michigan State University, 603 Wilson Road, East Lansing, MI 48824, USA

Abstract

Bacterial Microcompartments (BMCs) are organelles composed of a selectively permeable protein shell that encapsulates enzymes involved in CO₂ fixation (carboxysomes) or carbon catabolism (metabolosomes). Confinement of sequential reactions by the BMC shell presumably increases the efficiency of the pathway by reducing the crosstalk of metabolites, release of toxic intermediates, and accumulation of inhibitory products. Because BMCs are composed entirely of protein and self-assemble, they are an emerging platform for engineering nanoreactors and molecular scaffolds. However, testing designs for assembly and function through *in vivo* expression is labor-intensive and has limited the potential of BMCs in bioengineering. Here, we developed a new method for *in vitro* assembly of defined nanoscale BMC architectures: shells and nanotubes. By inserting a “protecting group”, a short ubiquitin-like modifier (SUMO) domain, self-assembly of

*Corresponding Author: ckerfeld@lbl.gov.

Author Contributions

AH and JSP designed and performed the experiments and wrote the paper. NS generated reagents, assisted with experiments, and performed electron microscopy. BF designed experiments, generated reagents, collected and analyzed TEM data, and helped write the paper. CA also generated reagents and collected and helped analyze TEM data. CAK supervised the project, contributed to experimental design, and co-wrote the paper. All authors discussed the results and contributed critical feedback during the preparation of the manuscript.

The authors declare no competing financial interests.

ASSOCIATED CONTENT

The following files are available free of charge: Additional details on materials and methods (including protein expression and purification details, *in vitro* assembly set up and purification, SDS-PAGE analysis, TEM analysis (including negative stain and thin section), fluorescence analysis, DLS analysis and details on the plasmid constructions). These are accompanied by additional supporting table and figures: Solution behavior of SUMOBMC-HHO (Figure S1), SDS-PAGE analysis of purified SUMOBMC-H proteins (Figure S2), SDS-PAGE analysis of protein preps used for IV assemblies (Figure S3), Size exclusion chromatography analysis and molecular model of SUMOBMC-H proteins (Figure S4), TEM analysis of a neat *in vitro* mHO shell assembly without MBP-Ulp (Figure S5), SDS-Page analysis of SUMO cleavage of Halo hexamers (Figure S6), TEM analysis of *in vitro* assembled Halo shells (Figure S7), structural characterization of CAP-purified *in vitro* assembled mHO shells (Figure S8), uncropped SDS-PAGE gels of Figure 4 from main text (Figure S9), TEM images of purified *in vitro* assembled mHO shells (Figure S10), DLS analysis of various *in vitro* assembled mHO shell eluates (Figure S11), fluorescence analysis of various *in vitro* assembled mHO shell eluates (Figure S12), TEM analysis of mHO copurified with 5 nm Au-COOH particles (Figure S13), Representative TEM images of mHO⁺ copurified with 5 nm Au-COOH particles (Figure S14), and charge distribution map of three BMC-H hexamers (Figure S15).

shell proteins *in vivo* was thwarted, enabling preparation of concentrates of shell building blocks. Addition of the cognate protease removes the SUMO domain and subsequent mixing of the constituent shell proteins *in vitro* results in the self-assembly of three types of supramolecular architectures: a metabolosome shell, a carboxysome shell, and a BMC protein-based nanotube. We next applied our method to generate a metabolosome shell engineered with a hyper-basic luminal surface, allowing for the encapsulation of biotic or abiotic cargos functionalized with an acidic accessory group. This is the first demonstration of using charge complementarity to encapsulate diverse cargos in BMC shells. Collectively, our work provides a generally applicable method for *in vitro* assembly of natural and engineered BMC-based architectures.

Keywords

Bacterial microcompartments; structural biology; protein engineering; *in vitro* supramolecular self-assembly; shells; nanotubes; electrostatic-based encapsulation

Spatial organization is a hallmark of life, distinguishing “a living cell from a soup of the chemicals of which it is composed”.¹ While sub-cellular organization in the form of membrane-delimited compartments is ubiquitous in Eukarya, only recently has it been recognized that there are specialized compartments in prokaryotes. However, unlike eukaryotic organelles which are (with a few exceptions)² bound by lipid bilayers, the membranes of many prokaryotic organelles such as encapsulins^{3, 4} and bacterial microcompartments^{5–7} (BMCs) are proteinaceous.

BMC shells are composed of hexagonal- and pentagonal-shaped oligomers that typically assemble into polyhedral bodies of approximately twenty-five to several hundred nanometers.⁵ Typically, small pores formed at the cyclic symmetry axis of the hexagonal subunits mediate the selective passage of metabolites across this protein membrane, while the shell’s luminal surface interacts with the enzymatic cargo. The combined properties of selective permeability and specific enzyme encapsulation allow for the compartmentalization of segments of diverse metabolic pathways including carbon fixation or catabolic reactions that generate volatile or toxic intermediates. Evidence of rampant horizontal gene transfer of BMC loci attests to their role as metabolic modules that expand bacterial ecophysiological diversity.⁶

Recently, BMCs have begun to be exploited by metabolic engineers to increase pathway efficiency; this is in keeping with their naturally evolved function as enzyme-loaded nanoreactors.^{8–10} Indeed, numerous bioengineering challenges are universally confronted when installing heterologous pathways such as inefficient substrate channeling, enzyme instability, formation of toxic intermediates, and reaction cross-talk; all of these may be mitigated by confining key reactions to microcompartment shells. Moreover, with the elucidation of an intact shell structure¹¹ and methodological advances towards their isolation and loading with heterologous cargo,¹² these particles are now poised for applications including medical imaging, drug delivery and polyvalent antigen display, which have historically been the purview of other nanoassemblies like ferritin and virus-like particles.
13–15

A necessary step in the creation of bespoke nanoreactors or devices based on BMC shell architectures is understanding the rules for their assembly and elucidating their permselectivity. All BMC shells are composed of apparently structurally redundant proteins; ^{5, 6} hexamers (BMC-H) and trimers (BMC-T_{single} and BMC-T_{double}) tile the facets of a BMC shell, but are functionally distinct in their permeability properties. In addition, understanding the extent of the compositional plasticity of shells is critical for tailoring a synthetic shell to a given application. Engineering aside, several studies have implicated BMCs in pathogenesis,^{6, 16} and therefore the development of methods to rapidly gain insight into shell assembly from different combinations of proteins and the precise determinants of their permeability properties could lead to novel pharmaceutical interventions.

Although much has been learned about BMC shell structure and function from physiological studies of hosts that naturally harbor BMCs (e.g., *Synechococcus PCC 7942*¹⁷ and *Salmonella typhimurium*¹⁸), the substantial breadth of inferred BMC functional diversity is not captured by the small number of genetically tractable organisms. Further complicating the matter, heterologous expression of individual shell proteins often leads to insoluble aggregates or unexpected architectures (e.g., rosettes and nanotubes) that are only able to be identified by time-intensive electron microscopy analysis of the host strain.^{19–23} Moreover, synthetic shell production from heterologous expression of BMC shell protein genes is contingent upon subunit selection and is sensitive to narrow tolerances in relative expression levels. These practical considerations have limited sampling of diverse shell proteins, which constitute a vast combinatorial space (as of July 2018 there are >5000 unique BMC-H protein sequences deposited in the UniProt database²⁴). To advance the analysis of natural shell properties and to realize the potential of BMCs in bioengineering, it would be ideal to have a method to study shell assembly and properties *in vitro*.

Here, we describe a method for *in vitro* assembly of BMC architectures that circumvents the limitations of *in vivo* assembly. Our method depends on purification of individual shell components that are prevented from forming supramolecular assemblies via a translationally-fused macromolecular protecting group, short ubiquitin-like modifier²⁵ (SUMO) (Figure 1). The method was applied to shells from the catabolic BMC of *Haliangium ochraceum*, a model BMC shell system^{11, 26} and shells derived from an anabolic BMC, the β -carboxysome.²⁷ The method is not limited to shell architectures; we use it to control assembly of protein nanotubes formed by a single hexamer subunit from a catabolic BMC model system.²¹ Moreover, we demonstrate that our *in vitro* assembly method can be used to generate an engineered shell from components modified to display a high net positive charge on its luminal surface to allow for electrostatics-based capture of complementarily charged cargo during assembly. This hyper-basic shell variant enabled the encapsulation of a protein tethered to a highly acidic accessory protein or incorporation of wholly abiotic cargo: carboxy-functionalized gold nanoparticles. Our findings provide a strategy for screening and controlling self-assembly of BMC-based architectures, which can be applied to characterize native shells or to construct synthetic shells encapsulating a range of cargos.

When expressed in the absence of their usual shell protein complements, many BMC-H proteins (proteins containing a single Pfam00936 domain that form homohexamers) readily

form a variety of supramolecular structures including sheets^{20, 28} and tubes.^{21, 29} Such assemblies are mediated by the co-planar edge-edge interactions between hexamer subunits observed in structural studies.^{11, 20, 23, 28} We hypothesized that genetically fusing a bulky domain to a BMC-H protein would sterically frustrate assembly of the hexamers into larger molecular assemblies. To test this idea, we fused a hexahistidine tagged SUMO domain to the N-terminus of the sole BMC-H protein ($\text{SUMO}^{\text{BMC-H}_{\text{HO}}}$) in the model synthetic shell system derived from the BMC found in *H. ochraceum* (HO).²⁶ The SUMO domain is known to promote solubility and can be proteolytically removed by its cognate protease, Ubl-specific protease 1²⁵ (SUMO protease or Ulp), to yield its fusion partner in the native state.

When wildtype BMC-H_{HO} is expressed in *E. coli* BL21(DE3), a significant amount is found in the insoluble fraction, reflective of its propensity to self-assemble into large rosettes *in vivo*.²⁰ In contrast, the $\text{SUMO}^{\text{BMC-H}_{\text{HO}}}$ fusion product comprised a significant portion of the proteome, and remained largely in the soluble fraction according to SDS-PAGE analysis (Figure 2a). Similarly, large rosettes were observed in thin sections of *E. coli* expressing wildtype BMC-H_{HO} but were absent from the $\text{SUMO}^{\text{BMC-H}_{\text{HO}}}$ expression strain when examined with transmission electron microscopy (TEM) (Figure 2c). Purification using nickel-based affinity chromatography resulted in good yields of high purity protein (~20 mg/l of culture). $\text{SUMO}^{\text{BMC-H}_{\text{HO}}}$ can be concentrated with molecular weight cut-off spin filters to the point of “oiling out” (>100 mg/ml) with no apparent aggregation, in contrast to the native BMC-H_{HO} protein which typically aggregates at concentrations above 1 mg/ml. Buffer exchange into low sodium chloride buffer results in apparent aggregation of $\text{SUMO}^{\text{BMC-H}_{\text{HO}}}$, as indicated by the rapid appearance of cloudiness after elution from a desalting column (Figure S1). This can be subsequently reversed by addition of sodium chloride, suggesting that the SUMOylated protein exists in a dynamic, quasi-supramolecular assembly that is sensitive to the ionic strength of the buffer. Consequently, concentrated $\text{SUMO}^{\text{BMC-H}_{\text{HO}}}$ preparations were maintained in buffer containing 500 mM NaCl. When co-incubated with the SUMO protease (MBP-Ulp) at 1:20 (w/w) loading, $\text{SUMO}^{\text{BMC-H}_{\text{HO}}}$ is almost completely cleaved into SUMO and native BMC-H_{HO} proteolytic fragments within 1 h (Figure S2a).

Next we tested the *in vitro* assembly method on the non-natural architecture formed by a single BMC-H protein ($\text{BMC-H}_{\text{Rmm}}$) from a catabolic BMC proposed to be involved in aminoacetone degradation in *Mycobacterium smegmatis*.³⁰ $\text{BMC-H}_{\text{Rmm}}$ subunits were previously shown to spontaneously assemble into nanotubes both *in vivo* and *in vitro*.²¹ Unlike wildtype $\text{BMC-H}_{\text{Rmm}}$, which forms nanotubes at concentrations above 3 mg/ml,²¹ SUMOylated $\text{BMC-H}_{\text{Rmm}}$ is highly soluble (Figure 2a). $\text{SUMO}^{\text{BMC-H}_{\text{Rmm}}}$ could be purified in good yield (~10 mg/l of culture) and concentrated to >50 mg/ml without apparent aggregation. Furthermore, the addition of the SUMO domain prevented the formation of $\text{BMC-H}_{\text{Rmm}}$ nanotubes *in vivo* (Figure 2b). Cleavage of the SUMO domain from $\text{SUMO}^{\text{BMC-H}_{\text{Rmm}}}$ (7 mg/ml sample) by MBP-Ulp is nearly complete within 1 h (Figure S2b) and allows self-assembly of free $\text{BMC-H}_{\text{Rmm}}$ into characteristic nanotubes (Figure 3a), recapitulating the behavior of the wildtype protein. Size exclusion chromatography analysis of the SUMOylated BMC-H proteins indicates the solution state of the fusions is primarily hexameric, which is consistent with modeling that suggests the SUMO domain should not interfere with hexamerization of the individual protomers (Figures S4a and S4b).

Collectively, we demonstrate that SUMOylation of two different BMC-H proteins prevents their self-assembly *in vivo* allowing for the purification of large amounts of soluble protein.

Laborious screening of protein combinations and expression conditions have led to successful production of empty shells from just a few BMC systems.^{26, 27, 31–33} We next tested if SUMO could facilitate screening conditions for formation of three dimensional shells from combinations of multiple shell proteins *in vitro*. Because SUMO-BMC-H_{HO} did not form higher order assemblies *in vivo*, we tested whether proteolysis with MBP-Ulp could lead to minimal HO (mHO) shell³⁴ formation *in vitro* in the presence of a trimer (3x BMC-T1_{HO} proteins) and pentamer (5x BMC-P_{HO} proteins) (Figure S3). These three HO shell proteins were mixed together with their relative subunit stoichiometries mirroring those in the HO shell crystal structure (60 hexamer : 20 trimer : 12 pentamer).¹¹ After addition of MBP-Ulp, reactions were incubated at room temperature for 20 h and neat reactions were spotted on grids for negative stain TEM analysis. Copious mHO shells were readily observed (Figure 3b), which is comparable to *in vivo* produced HO shells.²⁶ Surprisingly, when MBP-Ulp was excluded in the assembly reaction, shells were also observed by TEM (Figure S5), though apparently not in the same abundance as when the protease was included. These results show that the SUMO domain provides enough steric hindrance to prevent the BMC-H_{HO} subunits from forming homooligomeric supramolecular structures *in vivo* (and *in vitro*). However, it does not fully preclude the assembly of shells in the presence of its cognate shell complements *in vitro*.

To evaluate the generalizability of our method for *in vitro* shell assembly, we next SUMOylated the hexamer subunits from the cyanobacterial β -carboxysome, a BMC phylogenetically distant from the HO shell. We have previously shown that synthetic carboxysome shells can be assembled when a combination of four different *Halotheca* sp. PCC 7418 (Halo) shell proteins were expressed in *E. coli*.²⁷ We fused the SUMO domain to two BMC-H proteins from the Halo system, CcmK1 and CcmK2 (SUMO-BMC-H_{ccmK1} and SUMO-BMC-H_{ccmK2}). These BMC-H paralogs have been shown to assemble into sheets in different conditions (e.g., in the crystalline state³⁵ or applied on a carbon surface³⁶). The SUMOylated carboxysome hexamers can be readily purified (40–50 mg/l of culture) (Figure S3) and highly concentrated (>100 mg/ml of purified protein) without apparent aggregation, mirroring the impact of SUMOylation observed with BMC-H_{HO} and BMC-H_{Rmm}. MBP-Ulp was confirmed to be active on these fusions, though curiously the cleavage kinetics of SUMO-BMC-H_{ccmK2} are appreciably slower, requiring more than two hours to proteolyze the majority of the fusion protein, although cleavage is >95% complete by the 20 hour timepoint (corresponding to when assemblies are analyzed by TEM; Figure S6). Shell assembly reactions were set up using the SUMOylated hexamers along with the respective trimer and pentamer (using putative stoichiometries for the carboxysome shells) and incubated with MBP-Ulp as before. Empty carboxysome shells with morphology indistinguishable from *in vivo* produced shells are observed in these assembly reactions, though apparently not in the same abundance or with the same morphological homogeneity of the HO shells (Figure 3c and Figure S7). Overexpression of wildtype CcmK1 and CcmK2 proteins are not known to form supramolecular assemblies *in vivo* as analyzed by thin-section TEM (Eric Young and Daniel Ducat, personal correspondence). Nonetheless, we analyzed *in vitro* cleaved

SUMO-BMC-H_{ccmK1} and SUMO-BMC-H_{ccmK2} via TEM and find no evidence of regular supramolecular assemblies, in agreement with *in vivo* observations (data not shown).

We next sought to develop an assay for shell assembly efficiency. We chose to further optimize assembly efficiency and engineer the HO shell (i.e., mHO) because of the availability of detailed structural information.^{11, 20, 26} We define efficiency as the amount of pure, assembled shells obtained as a fraction of theoretical yield based on protein inputs and assumed stoichiometries. To isolate assembled shells from the complex reaction mixture, which includes free shell proteins, MBP-Ulp, and trace contaminants, we made use of complementation-based affinity purification (CAP)³⁴. In brief, a pentamer modified with the highly specific Strep-II tag³⁷ (BMC-P_{SIIHO}) is used as an affinity handle to pull-down intact shells. We then used analytical densitometry of the affinity purification fractions to estimate the percentage of input proteins recovered in the eluate.

We tested this strategy by setting up a HO shell *in vitro* assembly reaction at 200 μ l scale (10 μ M calculated final concentration of hexamer for a theoretical 167 nM concentration of ~6 MDa shells, see Supplementary Information section). After overnight cleavage and assembly, the reaction was applied to a 1 ml StrepTrap column and equivalent quantities of the assembly, flow-through, and eluate fractions were analyzed by SDS-PAGE. Undesired proteins, primarily MBP-Ulp, cleaved SUMO and unincorporated hexamer, are found in the flow-through fractions (Figure 4a, lane 2) while the banding pattern of the elution fraction suggests shells were isolated (Figure 4a, lane 3; the pentamer subunit band is likely below the detection limit). Analysis by dynamic light scattering (DLS) reveals that indeed, the eluate fraction contains highly pure and homogenous shells (Figure S8a) of the expected size, and this is corroborated by TEM analysis (Figure S8b). Together these data validate our strategy for shell recovery and assembly efficiency quantification and reveal that, in this condition, approximately 20% of the input hexamer is recovered as assembled shells.

The effect of different subunit stoichiometries on assembly efficiency was next interrogated by setting up assemblies using varying stoichiometric amounts of trimer and pentamer. Five-fold stoichiometric excess of pentamer resulted in no appreciable change in assembly efficiency (data not shown). In contrast, three-fold stoichiometric excess of trimer results in an approximately 60% increase in shell yield compared to 1x trimer (Figure 4b, compare lanes 2 and 3), with a modest additional increase obtained using a nine-fold superstoichiometry (lanes 3 and 4; approximately double the yield compared to 1x trimer). When the BMC-T proteins are omitted entirely (lane 1), no shells are recovered indicating that this subunit is essential for shell formation. In addition, this further demonstrates that the CAP method does not recover partially formed shells such as might be formed from an association of five hexamers abutting a pentamer. Three-fold stoichiometric excess of trimer was used for subsequent experiments as a compromise between efficient assembly and conservation of the BMC-T protein preparation. Moreover, we performed a timecourse analysis of shell assembly by isolating shells after timepoints ranging from zero to sixteen hours. Shell recovery is not apparent until the two-hour timepoint (Figure 4c, lane 3). Assembly of shells in native organisms is presumably faster, which could indicate that native cargo enzymes help to guide assembly or that intracellular molecular crowding effects accelerate assembly. After six hours, shells are readily detectable and are recovered at

approximately half the yield obtained after overnight assembly (Figure 4c compare lanes 4 and 5).

We next tested the *in vitro* assembly of a HO shell composed of hexamers modified to increase the positive charge of the lumen. We hypothesized that this construct would enhance the targeting of negatively charged cargos to the shell. BMC-H_{HO} is the major shell component of the HO BMC, and we increased its net positive charge by mutating two glutamate residues (E65 and E69) to arginines (Figure 5a) in the SUMO-BMC-H_{HO} construct, generating SUMO-BMC-H⁺_{HO}. These two glutamate residues orient toward the lumen of the HO shell,¹¹ which would theoretically provide 720 additional positive charges per shell at neutral pH. TEM and DLS analysis of a purified *in vitro* assembly reaction containing SUMO-BMC-H⁺_{HO}, BMC-T1_{HO}, and BMC-P_{SIIIHO} resulted in the formation of minimal HO⁺ shells (mHO⁺) (Figure S10a and S11a). These shells have an average diameter (determined from TEM images) of 34 (±2) nm, which is comparable to full HO²⁶ and mHO shells.³⁴ After confirming the formation of mHO⁺ shells, we then assayed the encapsulation of GFP³⁸ fused to the negatively charged RraB protein.^{39, 40} This protein, which carries a -42 net charge at neutral pH, was fused to the N-terminus of GFP (RraB-GFP).

We performed *in vitro* assembly reactions with mHO⁺ shell proteins and RraB-GFP, as well as three control reactions of combinations with mHO and GFP. Shells from these reactions were purified using StrepTactin XT spin columns, which allowed for smaller scale reactions and higher throughput purifications compared to the 1 ml StrepTrap column. TEM images of purified mHO⁺ + RraB-GFP (Figure 5b) and mHO + RraB-GFP (Figure S10b) showed distinct microcompartment shells with average diameters of 34 (±2) nm and 36 (±2) nm, respectively. A close inspection of the TEM images also reveals phenotypic difference between the modified and unmodified mHO shells—the mHO⁺ + RraB-GFP shells have thicker walls compared to empty mHO⁺ (Figure S10a) and mHO + RraB-GFP (Figure S10b) shells, which suggests that the RraB-GFP proteins are localized at the inner-surface of the mHO⁺ shells. Moreover, SDS-PAGE analysis of the purified mHO⁺ + RraB-GFP reaction shows the presence of all three shell components, as well as the RraB-GFP cargo (Figure 5c, lane 1), which is consistent with the qualitative differences in the electron micrographs and argues against the possibility of staining artifacts. Based on the intensity of the BMC-T1_{HO} and BMC-H_{HO} monomer bands, the assembly of mHO shells appears consistent across the different reactions, however the cleaved BMC-H⁺_{HO} and BMC-P_{SIIIHO} bands co-migrate, as well as BMC-T1_{HO} and GFP (Figure 5c and S12a) bands, complicating comparisons of yield by SDS-PAGE. Nevertheless, the mHO⁺ + RraB-GFP assembly shows higher RraB-GFP band intensity than the corresponding mHO sample (Figure 5c, compare lanes 1 and 2). The enhanced incorporation of RraB-GFP within mHO⁺ relative to mHO is demonstrated further in our fluorescence analysis. We repeated, in triplicate, the *in vitro* assembly encapsulation reactions and examined the fluorescence properties of the purified reactions. All eluates show emission spectra characteristic of GFP³⁸ although the intensity of the emission at 510 nm for mHO⁺ + RraB-GFP eluates is approximately four times more intense than the corresponding mHO eluates with RraB-GFP and the mHO eluates with GFP (i.e., mHO + GFP and mHO⁺ + GFP; Figure 5d). This increase in GFP signal is corroborated by visualization of the eluates with a UV-transilluminator (Figure S12b). These fluorescence

data further indicate that GFP is a trivial component of the putatively overlapping BMC-T1_{HO} and GFP bands in Figure 5c.

Finally, we sought to determine if mHO⁺ shells showed enhanced capture (relative to unmodified shells) of a negatively charged abiotic cargo during assembly. *In vitro* assembly reactions spiked with 5 nm diameter lipoic acid coated gold nanoparticles (AuCOOH) were purified using StrepTactin XT spin columns and shell assembly and encapsulation was evaluated by TEM. Purification of mHO + AuCOOH (Figure S13) and mHO⁺ + AuCOOH (Figure 5e and Figure S14) samples contain shells with an average diameter of 41 (\pm 2) nm and 39 (\pm 6) nm, respectively. These diameters are comparable to the values determined for their mHO only counterparts. TEM images of mHO + AuCOOH display mostly empty shells, with some gold particles either unassociated or abutting the outside surface of the shells (Figure S13). In contrast, the TEM images mHO⁺ + AuCOOH show numerous Au particles that are overlaid with the shells (Figure 5e, S14), with as many as eight gold particles co-localized with a single shell. While we cannot definitely say these gold particles are inside the shells (standard transmission electron microscopy cannot resolve in the z-dimension), the lack of such observations in the unmodified mHO control shells strongly suggests that electrostatic interactions between the modified shell lumen and complementarily charged AuCOOH nanoparticles mediate encapsulation of this abiotic cargo within the shells.

In this study, we have developed a generalizable *in vitro* assembly method to facilitate the study of BMC-based architectures (Figure 1). Addition of the SUMO domain prevents the spontaneous assembly of BMC-H shell proteins into extended sheets and nanotubes *in vivo*. The SUMO domain also enhances the solubility and yield of BMC-H shell proteins, which would otherwise aggregate at high concentrations. Purified BMC-H protein concentrates can then be used to assemble three-dimensional tubes or shells (with the addition of pentamer and trimer). This *in vitro* strategy can be applied to BMC shell proteins from functionally diverse BMCs, as demonstrated by the successful assembly of carboxysome and metabolosome shells, or tubular architectures derived from a single metabolosome shell hexamer (Figure 3). A high net-positive charge was subsequently engineered onto the luminal side of the major shell protein of the HO shell to enhance the capture of complementarily charged cargos. These two modifications may be regarded as (1) a conditional loss of function mutation and (2) a gain of function mutation, which enabled the first encapsulation of cargo in a microcompartment shell using electrostatics.

The ~15 kDa SUMO domain effectively prevents spontaneous assembly of hexamer subunits into extended sheets and tubes. Simple geometric considerations may explain this phenomenon. Hexamers can assemble into extended arrays (sheets and tubes) due to their ability to form ~0° edge-edge (co-planar) associations, whereas angled interactions with the trimer and pentamer introduce the curvature required to close into 3D icosahedra. This low-angle self-association and closely juxtaposed N-termini likely explain the fused SUMO domain's ability to disrupt assembly, though it is curious to note that SUMOylated hexamers aggregate in low salt conditions which may suggest that the SUMO domain does not completely ablate self-assembly.

The spontaneous formation of shells and nanotubes in the simple buffering system employed here demonstrates that no extrinsic assembly factors such as osmolytes or divalent cations are required for assembly, leaving room for further optimization. Likewise, the effects of additional shell proteins on the resulting architectures can be readily evaluated. Efficiency of assembly is, however, sensitive to the relative stoichiometries of the constituent shell proteins as demonstrated by the increase in shell yield when superstoichiometric amounts of trimer are used to assemble HO shells. This may be explained by a kinetic competition between sheet formation and shell formation, which can be biased towards the latter with higher trimer amounts. Suboptimal subunit stoichiometries may therefore account for the relatively less efficient assembly of the carboxysome shells which have the additional potentially confounding property of being comprised of multiple BMC-H paralogs. Nevertheless, demonstration of *in vitro* assembly of varied architectures from three functionally and phylogenetically distant BMCs suggests the method is broadly applicable.

To-date, most targeting efforts of non-native cargo^{8, 41, 42} to BMC shells involve encapsulation peptides⁴³ that are native to that BMC type. Efficiency of encapsulation by this method has been consistently low, so we devised a modified shell targeting mechanism based on charge. In the model HO shell system,¹¹ the external and internal surfaces have an overall negative charge (Figure S15). These electrostatic properties likely play an important role in the biogenesis of the native HO BMC but are not ideal for encapsulating non-natural cargo using electrostatics as both sides of the shell are similarly charged. Using the conditional loss of function modification of BMC-H_{HO} (resulting from fusion with SUMO), we identified a mutant capable of assembling into a minimal HO shell with the opposite (positive) charge (mHO⁺) on the luminal side for selective encapsulation of negatively charged cargo. This engineered gain of function mutation to the HO shell mimics the electrostatic properties of viral capsids that generally have an acidic exterior and basic interior for DNA or RNA packaging⁴⁴ and therefore may be suited for encapsulating nucleic acids in BMCs.

To complement our new mHO⁺ shell, we developed an electrostatic-based encapsulation module by repurposing the small hyper-acidic RraB domain, which had been previously used to enhance stability of recombinant proteins.⁴⁰ As proof of principle, we fused RraB to GFP and compared encapsulation in mHO and mHO⁺ shells, produced using our *in vitro* assembly method. Fluorescence analysis showed four times higher GFP signal from mHO⁺ shells copurified with RraB-GFP than in reactions with mHO and RraB-GFP (Figure 5d). These results suggest the combination of a positively charged lumen and a negatively charged domain fused to the cargo led to the targeted encapsulation of the biotic RraB-GFP cargo. Using the same engineered mHO⁺ construct, we also demonstrated the targeting of a negatively charged abiotic cargo to a BMC shell *in vitro*. TEM micrographs show that gold nanoparticles functionalized with lipoic acid were co-localized with the purified mHO⁺ shells (Figure 5e). This suggests that the Au particles are caged inside mHO⁺ shells due to the electrostatic interaction between the carboxylic acid groups on the surface of the Au particles and the positively charged mHO⁺ lumen. The *in vitro* assembly method enabled construction of an engineered mHO shell with a hyper-basic lumen to facilitate the encapsulation of acidic cargo, demonstrating an alternative encapsulation strategy that has not been performed before in BMC bioengineering.

A similar charge complementarity strategy has also been applied to lumazine synthase microcompartments.^{45–47} However, in contrast to lumazine synthase, which is made up of only one type of pentagonal subunit, BMC shells are composed of multiple types of structurally redundant, interchangeable subunits. This suggests BMCs are a more versatile system for bioengineering as the diversity of natural and synthetic shell components can be mixed-and-matched to create a wide array of functional architectures. Furthermore, the ability to produce shell components at high concentrations enables the rapid screening for *in vitro* assembly of shells, thereby providing a flexible platform to design and test the assembly of synthetic nanocompartments with individually customized building blocks.

Bioinformatic studies have revealed the widespread occurrence of BMCs across the Kingdom Bacteria, the majority of which remain to be functionally characterized.^{6, 48} All BMC shells are composed of structurally redundant hexamers and trimers (on average 3.5 BMC-H, 1.4 BMC-T per BMC locus⁶) typically with paralog-specific residues surrounding the pores. Structural modeling of these paralogs revealed differently sized and charged pores that presumably correspond to the specific metabolite permeability requirements of a given BMC shell.⁴⁸ Therefore, the ability to heterologously express and purify in high yields and control the stoichiometry of shell proteins (including multiple paralogs) afforded by our *in vitro* method will accelerate both the bioengineering and fundamental study of BMCs. In the former case, our method will allow us to build and characterize diverse BMC architectures assembled from shell proteins derived from different BMC systems. These chimeric BMC shells or scaffolds (tubes or sheets) can be imparted with properties that facilitate metabolite specificity and selectivity by varying the identity or quantity of shell proteins with known permeability or by including rationally designed mutants with the desired structural or physicochemical properties. Our approach could also be used to expand the range of cargos packaged within BMCs, and thus, the reactions that could be encapsulated, from native and engineered enzymes to metal clusters or inorganic reporters. Moreover, our *in vitro* method provides a way to study the assembly and permeability properties of all BMC shells, including those originating from pathogenic strains. Future work on these types BMCs will help in our understanding of their potential role(s) in human health and disease and help develop novel therapeutic strategies to disrupt metabolism in pathogens.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health, National Institute of Allergy and Infectious Diseases (NIAID) grant 1R01AI114975–01 (AH, NS, and CAK) with infrastructure support from the U.S. Department of Energy, Basic Energy Sciences, Contract DE-FG02–91ER20021 (JSP, BF, CA, and CAK). We thank Markus Sutter for helpful suggestions and assistance with figure creation, Cecilia Blikstad for the MBP-Ulp protein preparation, Rodney Burton for use of the pRB1 plasmid, Ciara Fromwiller for the pCF1 plasmid, Lauren Achtemeier for assistance with HO shell protein preparation, and Raul Gonzalez-Esquer for assistance with electron microscopy.

REFERENCES

1. Harold FM *Microbiol Mol Biol Rev* 2005, 69, (4), 544–64. [PubMed: 16339735]
2. Kedersha NL *The Journal of Cell Biology* 1986, 103, (3), 699–709. [PubMed: 2943744]

3. Sutter M; Boehringer D; Gutmann S; Gunther S; Prangishvili D; Loessner MJ; Stetter KO; Weber-Ban E; Ban N *Nat Struct Mol Biol* 2008, 15, (9), 939–47. [PubMed: 19172747]
4. Giessen TW; Silver PA *Nat Microbiol* 2017, 2, 17029. [PubMed: 28263314]
5. Kerfeld CA; Aussignargues C; Zarzycki J; Cai F; Sutter M *Nat Rev Microbiol* 2018, 16, (5), 277–290. [PubMed: 29503457]
6. Axen SD; Erbilgin O; Kerfeld CA *PLoS Comput Biol* 2014, 10, (10), e1003898. [PubMed: 25340524]
7. Kerfeld CA; Erbilgin O *Trends Microbiol* 2015, 23, (1), 22–34. [PubMed: 25455419]
8. Lawrence AD; Frank S; Newnham S; Lee MJ; Brown IR; Xue WF; Rowe ML; Mulvihill DP; Prentice MB; Howard MJ; Warren MJ *ACS Synth Biol* 2014, 3, (7), 454–65. [PubMed: 24933391]
9. Liang M; Frank S; Lunsdorf H; Warren MJ; Prentice MB *Biotechnol J* 2017, 12, (3).
10. Frank S; Lawrence AD; Prentice MB; Warren MJ *J Biotechnol* 2013, 163, (2), 273–9. [PubMed: 22982517]
11. Sutter M; Greber B; Aussignargues C; Kerfeld CA *Science* 2017, 356, (6344), 1293–1297. [PubMed: 28642439]
12. Hagen A; Sutter M; Sloan N; Kerfeld CA *Nat Commun* 2018, 9, (1), 2881. [PubMed: 30038362]
13. Schmidt N; Mishra A; Lai GH; Wong GC *FEBS Lett* 2010, 584, (9), 1806–13. [PubMed: 19925791]
14. Maity B; Fujita K; Ueno T *Curr Opin Chem Biol* 2015, 25, 88–97. [PubMed: 25579455]
15. Giessen TW *Curr Opin Chem Biol* 2016, 34, 1–10. [PubMed: 27232770]
16. Jakobson CM; Tullman-Ercek D *PLoS Pathog* 2016, 12, (5), e1005558. [PubMed: 27171216]
17. Long BM; Badger MR; Whitney SM; Price GD *J Biol Chem* 2007, 282, (40), 29323–35. [PubMed: 17675289]
18. Bobik TA; Havemann GD; Busch RJ; Williams DS; Aldrich HC *J Bacteriol* 1999, 181, 5967–5975.
19. Pitts AC; Tuck LR; Faulds-Pain A; Lewis RJ; Marles-Wright J *PLoS One* 2012, 7, (10), e48360. [PubMed: 23144756]
20. Sutter M; Faulkner M; Aussignargues C; Paasch BC; Barrett S; Kerfeld CA; Liu LN *Nano Lett* 2015.
21. Noël CR; Cai F; Kerfeld CA *Advanced Materials Interfaces* 2015, n/a–n/a.
22. Heldt D; Frank S; Seyedarabi A; Ladikis D; Parsons JB; Warren MJ; Pickersgill RW *Biochem J* 2009, 423, (2), 199–207. [PubMed: 19635047]
23. Pang A; Frank S; Brown I; Warren MJ; Pickersgill RW *J Biol Chem* 2014, 289, (32), 22377–84. [PubMed: 24873823]
24. The UniProt, C. *Nucleic Acids Res* 2017, 45, (D1), D158–D169. [PubMed: 27899622]
25. Li SJ; Hochstrasser M *Nature* 1999, 398, (6724), 246–51. [PubMed: 10094048]
26. Lassila JK; Bernstein SL; Kinney JN; Axen SD; Kerfeld CA *J Mol Biol* 2014, 426, (11), 2217–28. [PubMed: 24631000]
27. Cai F; Bernstein SL; Wilson SC; Kerfeld CA *Plant Physiol* 2016.
28. Young EJ; Burton R; Mahalik JP; Sumpter BG; Fuentes-Cabrera M; Kerfeld CA; Ducat DC *Front Microbiol* 2017, 8, 1441. [PubMed: 28824573]
29. Lee MJ; Mantell J; Hodgson L; Alibhai D; Fletcher JM; Brown IR; Frank S; Xue WF; Verkade P; Woolfson DN; Warren MJ *Nat Chem Biol* 2018, 14, (2), 142–147. [PubMed: 29227472]
30. Jorda J; Lopez D; Wheatley NM; Yeates TO *Protein Sci* 2013, 22, (2), 179–95. [PubMed: 23188745]
31. Parsons JB; Frank S; Bhella D; Liang M; Prentice MB; Mulvihill DP; Warren MJ *Mol Cell* 2010, 38, (2), 305–15. [PubMed: 20417607]
32. Sargent F; Davidson FA; Kelly CL; Binny R; Christodoulides N; Gibson D; Johansson E; Kozyrska K; Lado LL; Maccallum J; Montague R; Ortmann B; Owen R; Coulthurst SJ; Dupuy L; Prescott AR; Palmer T *Microbiology* 2013, 159, (Pt 11), 2427–36. [PubMed: 24014666]
33. Choudhary S; Quin MB; Sanders MA; Johnson ET; Schmidt-Dannert C *PLoS One* 2012, 7, (3), e33342. [PubMed: 22428024]
34. Plegaria JS; Kerfeld CA *Curr Opin Biotechnol* 2017, 51, 1–7. [PubMed: 29035760]

35. Kerfeld CA; Sawaya MR; Tanaka S; Nguyen CV; Phillips M; Beeby M; Yeates TO *Science* 2005, 309, (5736), 936–938. [PubMed: 16081736]
36. Garcia-Alles LF; Lesniewska E; Root K; Aubry N; Pocholle N; Mendoza CI; Bourillot E; Barylyuk K; Pompon D; Zenobi R; Reguera D; Truan G *PLoS One* 2017, 12, (9), e0185109. [PubMed: 28934279]
37. Schmidt TG; Koepke J; Frank R; Skerra A *J Mol Biol* 1996, 255, (5), 753–66. [PubMed: 8636976]
38. Pedelacq JD; Cabantous S; Tran T; Terwilliger TC; Waldo GS *Nat Biotechnol* 2006, 24, (1), 79–88. [PubMed: 16369541]
39. Zhou L; Zhao M; Wolf RZ; Graham DE; Georgiou G *J Bacteriol* 2009, 191, (21), 6665–74. [PubMed: 19717586]
40. Zou Z; Cao L; Zhou P; Su Y; Sun Y; Li W *J Biotechnol* 2008, 135, (4), 333–9. [PubMed: 18599143]
41. Quin MB; Perdue SA; Hsu SY; Schmidt-Dannert C *Appl Microbiol Biotechnol* 2016, 100, (21), 9187–9200. [PubMed: 27450681]
42. Jakobson CM; Slininger Lee MF; Tullman-Ercek D *Protein Sci* 2017, 26, (5), 1086–1092. [PubMed: 28241402]
43. Aussignargues C; Paasch BC; Gonzalez-Esquer R; Erbilgin O; Kerfeld CA *Commun Integr Biol* 2015, 8, (3), e1039755. [PubMed: 26478774]
44. Flenniken ML; Uchida M; Liepold LO; Kang S; Young MJ; Douglas T *Viruses and Nanotechnology* 2009, 327, 71–93.
45. Worsdorfer B; Pianowski Z; Hilvert D *J Am Chem Soc* 2012, 134, (2), 909–11. [PubMed: 22214519]
46. Worsdorfer B; Woycechowsky KJ; Hilvert D *Science* 2011, 331, (6017), 589–92. [PubMed: 21292977]
47. Azuma Y; Zschoche R; Tinzl M; Hilvert D *Angew Chem Int Ed Engl* 2016, 55, (4), 1531–4. [PubMed: 26695342]
48. Zarzycki J; Erbilgin O; Kerfeld CA *Appl Environ Microbiol* 2015, 81, (24), 8315–29. [PubMed: 26407889]

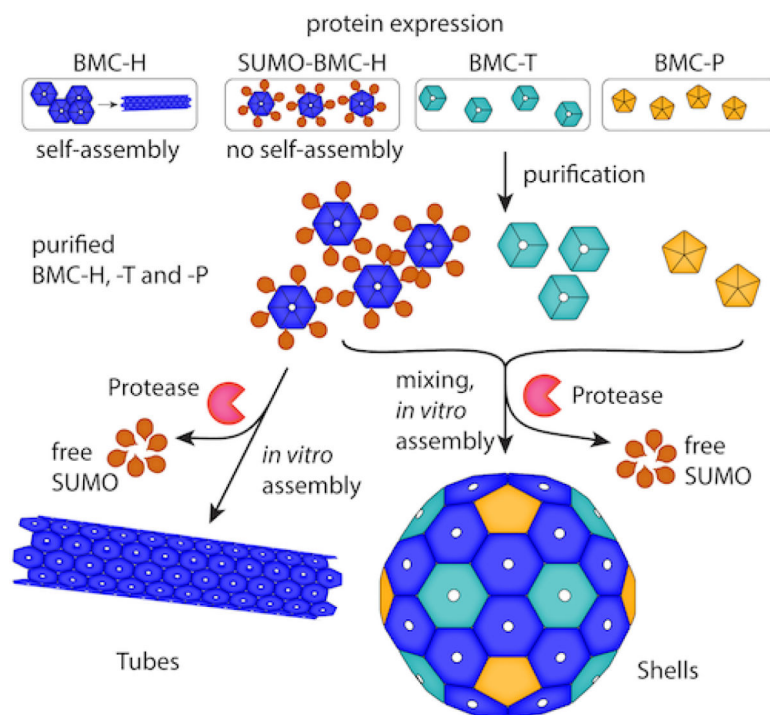


Figure 1. Schematic of the macromolecular protecting group strategy for in vitro assembly of BMC architectures.

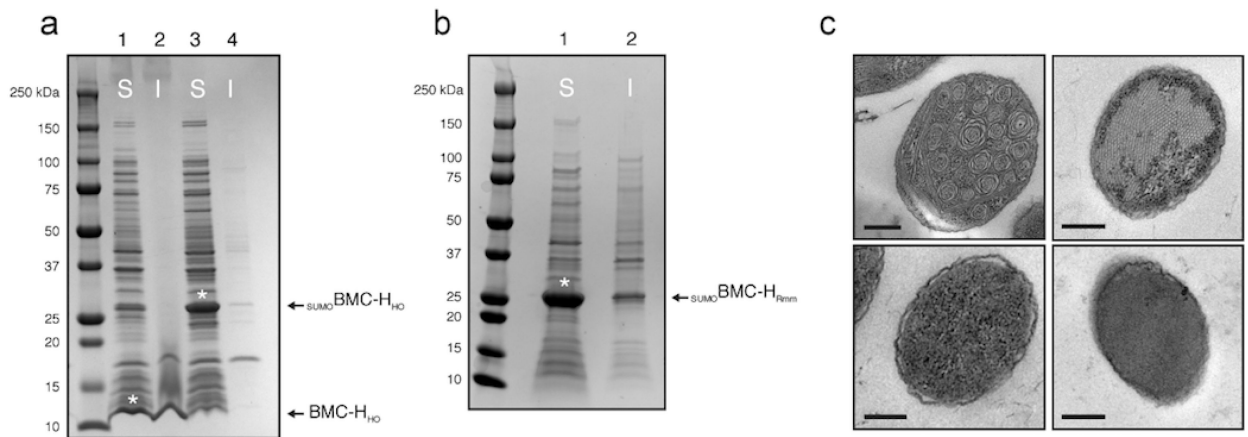


Figure 2.

Characterization of SUMOylated BMC-H proteins. (a) SDS-PAGE analysis of fractionated whole cell lysates of *Escherichia coli* strains expressing wildtype BMC-H_{HO} (lanes 1, 2) and SUMO-BMC-H_{HO} (lanes 3, 4). (b) SDS-PAGE analysis of SUMO-BMC-H_{Rmm}. For a. and b., soluble fractions are indicated with an “S” and insoluble fractions with “I,” and asterisks denote soluble protein of interest. (c) Transmission electron micrographs of thin sections of *E. coli* strains expressing wildtype BMC-H (upper left panel), SUMO-BMC-H_{HO} (lower left panel), WT BMC-H_{Rmm} (upper right panel), and SUMO-BMC-H_{Rmm} (lower right panel). Depicted results are representative of multiple fields/cells. Scale bars represent 200 nm.

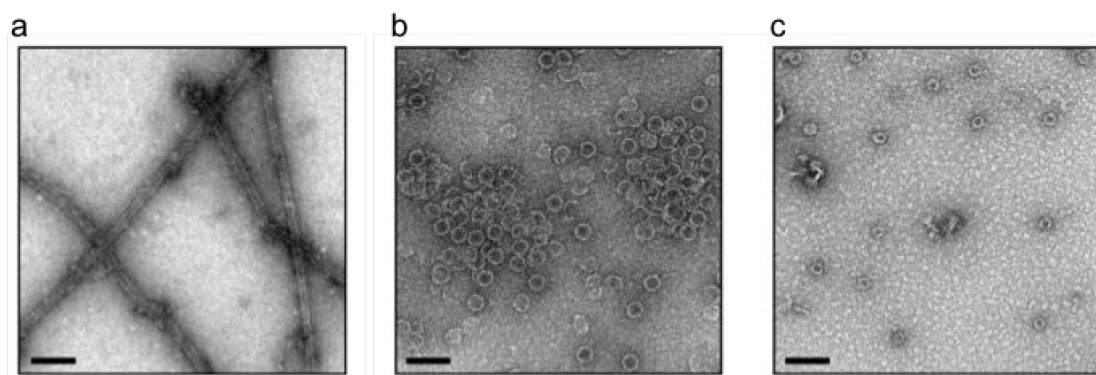


Figure 3. Transmission electron microscopy of various *in vitro* assembled nanoarchitectures. (a) *In vitro* assembled BMC- H_{Rmm} nanotubes. (b) *In vitro* assembled minimal HO shells. (c) *In vitro* assembled Halo shells (see also Figure S6). Scale bars = 100 nm; depicted results are representative of multiple fields.

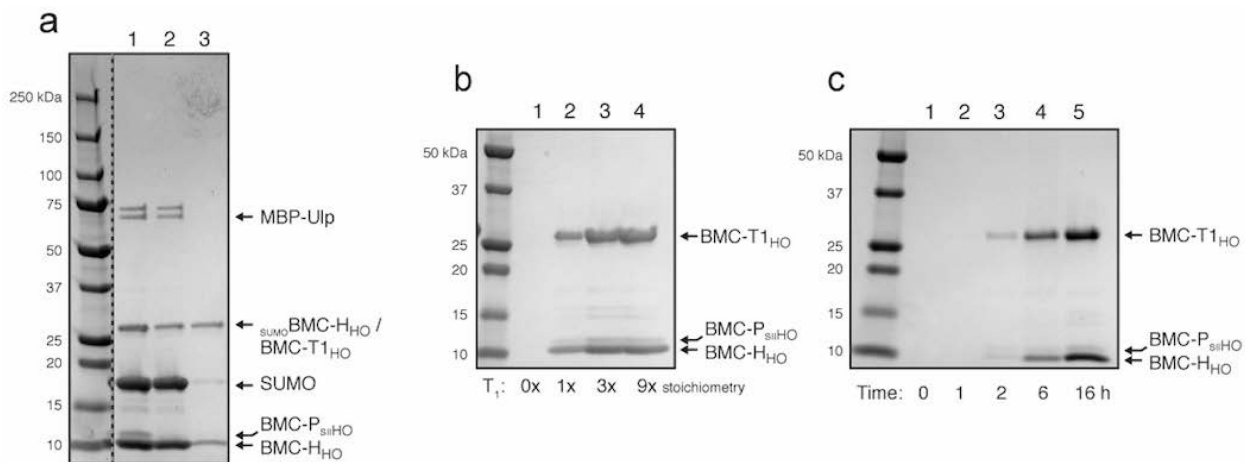


Figure 4.

SDS-PAGE analysis of *in vitro* assembly of minimal HO shells. (a) Shell assembly using CAP³⁴ (dashed line indicates cropped out lanes of irrelevant samples). Lane 1, 2 and 3 contains the neat assembly reaction, StrepTrap flow-through, and StrepTrap eluate, respectively. (b) BMC HO trimer sensitivity analysis showing five-fold concentrated eluates. Stoichiometry of BMC-T1_{HO} trimer used is shown beneath each lane. (c) Shell assembly time course (five-fold concentrated eluates; 3x BMC-T1_{HO} trimer stoichiometry used). Time between addition of MBP-Ulp protease and application of reaction to StrepTrap column is indicated below each lane. See Figure S9 for uncropped images of Figures 4b and 4c.

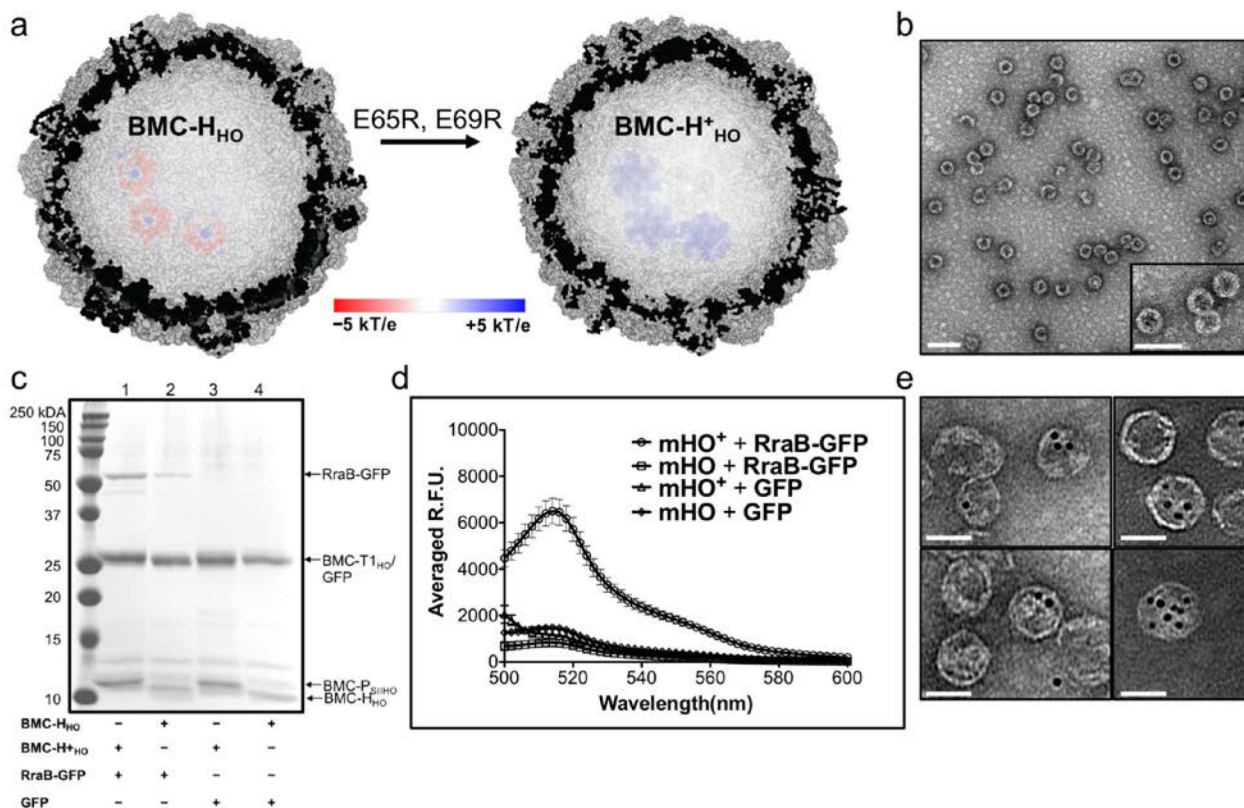


Figure 5.

Characterization of the electrostatic-based encapsulation of cargo. (a) Cutaway of the full HO shell structure (PDB ID: 5V74) incorporating the modelled electrostatic potential map of three BMC-H_{HO} (left) and BMC-H⁺_{HO} (right) hexamers. The electrostatic potential surface is colored from -5kT/e (red) to +5kT/e (blue). (b) Representative TEM image of purified mHO⁺ + RraB-GFP (scale bar = 100 nm). Inset is an enlarged view of the shells. See Figure S10c for a representative TEM micrograph of mHO⁺ + GFP (c) SDS-PAGE analysis of purified *in vitro* mHO/mHO⁺ shell assemblies with RraB-GFP or GFP. A comparable amount of protein was loaded in all the lanes, 1.2 – 1.9 μg; see Figure S11 for the DLS analysis of these assemblies. (d) Fluorescence emission spectra of purified of mHO⁽⁺⁾ shell assemblies with RraB-GFP or GFP. The error bars in the emission spectra plots are standard deviations from the average fluorescence emission at each wavelength measured for three independent purified reactions. (e) Various TEM images of mHO⁺ copurified with 5 nm Au-COOH particles (scale bars = 50 nm).