

## Supplementary Information

### A Designed Bacterial Microcompartment Shell with Tunable Composition and Precision Cargo Loading

Bryan Ferlez, Markus Sutter, Cheryl A. Kerfeld\*

**Table S1.** X-ray data collection and refinement statistics for CPH

<b>Data collection<sup>a</sup></b>	
Resolution range (Å)	43.8 – 1.61 (1.69-1.61)
Space group	P 6 <sub>3</sub> 2 2
Unit cell dimensions	73.5 73.5 119.6 Å 90 90 120 °
Total reflections	456,361 (64,426)
Unique reflections	25,615 (3,633)
Multiplicity	17.8 (17.7)
Completeness (%)	99.9 (99.2)
Mean I/sigma(I)	18.6 (6.5)
R-merge	0.09 (0.45)
R-meas	0.09 (0.46)
CC <sup>1/2</sup>	0.999 (0.973)
<b>Refinement<sup>a</sup></b>	
Resolution range (Å)	36.8-1.61 (1.67-1.61)
Number of reflections	25444 (2477)
Number of reflections used for R-free	1998 (196)
R-work (%)	20.9 (27.5)
R-free (%)	24.0 (31.0)
Number of non-hydrogen atoms	1538
macromolecules	1404
solvent	134
Protein residues	197
RMS (bonds, Å)	0.010
RMS (angles, °)	0.99
Ramachandran favored (%)	97.4
Ramachandran allowed (%)	2.6
Ramachandran outliers (%)	0
Clashscore	3.9
Average B-factor (Å <sup>2</sup> )	20.4

<sup>a</sup>Statistics for the highest-resolution shell are shown in parentheses.

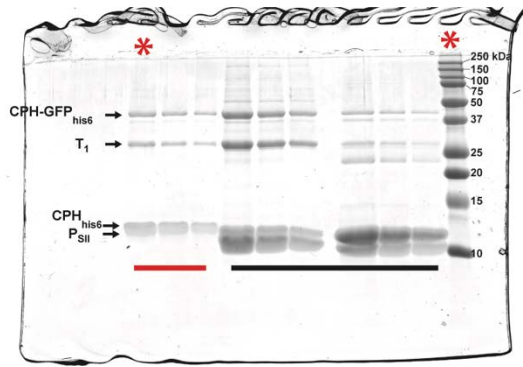
**Table S2.** Plasmids used in this study

Plasmid name	Vector	Cloning site 1 gene(s)	Cloning site 2 gene	Cloning site 1 promoter	Cloning site 2 promoter
pBF27	pET11b	CPH	N/A	T7	N/A
pBF53	pETDuet	CPH <sub>his6</sub> , T <sub>1</sub>	P <sub>SII</sub>	T7	T7
pBF64	pBbA2k	CPH-GFP <sub>his6</sub>	N/A	Tet	N/A
pBF71	pBbA2k	WTH-GFP <sub>his6</sub>	N/A	Tet	N/A

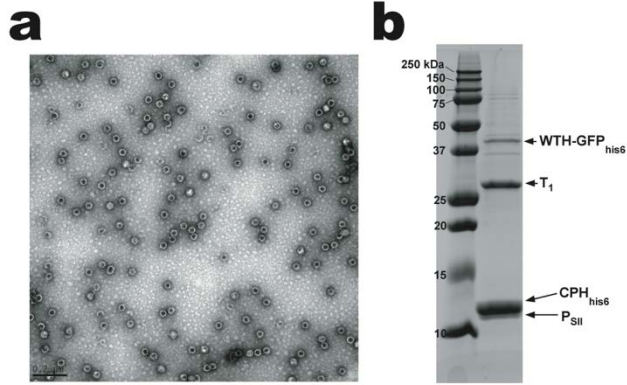
**Table S3.** Amino acid sequences of proteins used in this study

Protein	Amino acid sequence	MW (kDa) <sup>a</sup>	Note
CPH	MEVVAVHVIPRPHVNVDAALPLGRTPGMDKSAGSGSGSADALGMIE VRGFVGMVEAADAMVKA <del>AK</del> VELIGYEKTGGGYVTAVVRGDVA <del>AVK</del> AATEAGQRAAERVG	10.5	
CPH <sub>his6</sub>	(CPH)-GSHHHHHH	11.5	
T <sub>1</sub>	MDHAPERFDATPPAGEPDRPALGVLELTSIARGITVADAALKRAPSL LLMSRPVSSGKHLMMRGQVAEVEESMIAAREIAGAGSGALLDELE LPYAHEQLWRFLDAPVVADAW <del>EDTESVIIVETATVCAAIDSADAAL</del> KTAPVVLDRMRLAIGIAGKAFFLTGELADVEAAA <del>EVVRE</del> RRCGARLL ELACIARPVDELGRLLFF <u>YRIMLKSNRK</u>	23.2	The T <sub>1</sub> sequence contains an additional 10 amino acids (underlined) at the C-terminus due to a single base pair deletion in the stop codon in the pBF53 vector.
P <sub>SII</sub>	MVLGKVVGTVVASRKEPRIEGLSLLLVACDPDGTPTGGAVVCADA VGAGVGEVVL <del>YASGSSARQTEVTNNRPVDATIMAIVDLVEMGGDVR</del> FRKDG <del>SWSH</del> PQFEK	11.5	
CPH-GFP <sub>his6</sub>	(CPH)- GSGSGSKGEELFTGVVPILVELDGDVNGHKFSVRGE <del>GEDATNGKL</del> TLKFICTTGKLPVPWPTLVTTLT <del>YGVQCFARYPDHMKQHDFFSAM</del> PEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELK <del>GIDFKEDGNI</del> LGHKLEYNFNSHN <del>VYITADKQKNGIKANFKIRHNVEDGSVQLADHYQ</del> QNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRDH <del>MVLL</del> EFVTAAGIT HGMD <del>ELYKGS</del> HHHHHH	38.4	
WT-GFP <sub>his6</sub>	MADALGMIEVRGFVGMVEAADAMVKA <del>AK</del> VELIGYEKTGGGYVTAVV RGDVAAVKAATEAGQRAAERVGEVAVHVIPRPHVNVDAALPLGR TPGMDKSAGSGSGSKGEELFTGVVPILVELDGDVNGHKFSVRGE GEDATNGKLT <del>LKFICTTGKLPVPWPTLVTTLT</del> YGVQCFARYPDHMKQ HDFFSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELK GIDFKEDGNILGHKLEYNFNSHN <del>VYITADKQKNGIKANFKIRHNVEDGS</del> VQLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRDH <del>MVLL</del> EFVTAAGITHGMD <del>ELYKGS</del> HHHHHH	38.0	

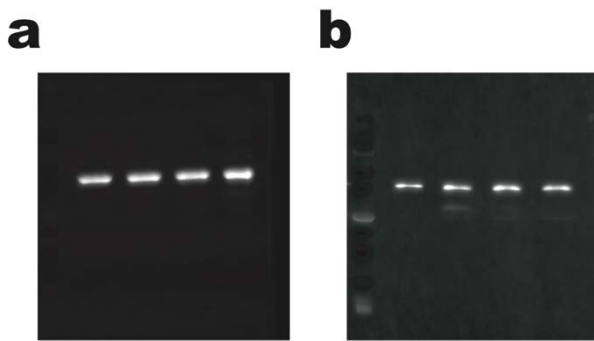
<sup>a</sup>) Calculated using the ExpASY web server ([web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/))



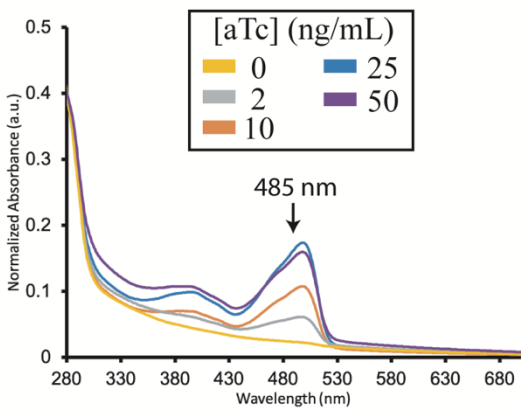
**Figure S1.** Uncropped SDS-PAGE used in Figure 2a. Red asterisks represent lanes displayed in Figure 2a. Red bar denotes different volumes loaded of the same CPH<sub>his6</sub>-T<sub>1</sub>-P<sub>SII</sub>/CPH-GFP<sub>his6</sub> sample. Black bar denotes samples not relevant to this study.



**Figure S2.** Characterization of CPH<sub>his6</sub>-T<sub>1</sub>-P<sub>SII</sub>/WTH-GFP<sub>his6</sub> shells used for CPA assays (Figure 3b, bottom). a) TEM micrograph and b) SDS-PAGE of purified shells. Scale bar is 0.2  $\mu$ m in (a).



**Figure S3.** Uncropped in-blot intrinsic GFP fluorescence images from a) Figure 3b, top and b) Figure 3b, bottom used to confirm equal loading across wells.



**Figure S4.** UV-vis spectra from purified CPH<sub>his6</sub>-T<sub>1</sub>-P<sub>SII</sub>/CPH-GFP<sub>his6</sub> shells. Spectra normalized to total protein content at 280 nm.