Title: Effect of metabolosome encapsulation peptides on enzyme activity, co-aggregation, incorporation and bacterial microcompartment formation

Authors: Rokas Juodeikis[1], Matthew J. Lee[1], Matthias Mayer[1], Judith Mantell[2,3], Ian R. Brown[1], Paul Verkade[2,3], Derek N. Woolfson[2,4], Michael B. Prentice[5], Stefanie Frank[6], Martin J. Warren[1]*

[1]Centre for industrial Biotechnology, School of Biosciences, University of Kent, Canterbury CT2 7NJ, UK

[2]School of Biochemistry, University of Bristol, Medical Sciences Building, University Walk, Bristol BS8 1TD, UK.

[3]Wolfson Bioimaging Facility, Medical Sciences Building, University Walk, Bristol BS8 1TD, UK.

[4]School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK.

[5]Department of Microbiology, University College Cork, Cork, Ireland

[6]Department of Biochemical Engineering, University College London, Bernard Katz Building, Gordon Street, London WC1E 6BT, UK.

* Correspondence to Martin J. Warren; m.j.warren@kent.ac.uk

Running Title: BMC encapsulation

Keywords: Bacterial organelles, protein aggregation, synthetic biology, cargo, targeting.

Summary

Metabolosomes, catabolic bacterial microcompartments, are proteinaceous organelles that are associated with the breakdown of metabolites such as propanediol and ethanolamine. They are composed of an outer multi-component protein shell that encases a specific metabolic pathway. Protein cargo found within BMCs is directed by the presence of an encapsulation peptide that appears to trigger aggregation prior to the formation of the outer shell. We investigated the effect of three distinct encapsulation peptides on foreign cargo in a recombinant BMC system. Our data demonstrate that these peptides cause variation with respect to enzyme activity and protein aggregation. We observed that the level of protein aggregation generally correlates with the size of metabolosomes, while in the absence of cargo BMCs self-assemble into smaller compartments. The results agree with a flexible model for BMC formation based around the ability of the BMC shell to associate with an aggregate formed due to the interaction of encapsulation peptides.

Introduction

Bacterial microcompartments (BMCs) are proteinaceous organelles found within a broad range of bacteria (Axen *et al.*, 2014, Bobik *et al.*, 2015, Kerfeld *et al.*, 2018). They consist of an outer semi-permeable protein shell that surrounds and sequesters a specific metabolic process. BMCs have a diameter of between 100-200 nm and can be broadly classified into two main groups on the basis of whether they promote anabolic or catabolic processes. Anabolic BMCs are associated with carbon fixation and house enzymes such as RuBisCO and carbonic anhydrase and are termed carboxysomes (Kerfeld *et al.*, 2010, Rae *et al.*, 2013, Yeates *et al.*, 2008). Catabolic BMCs are associated with the breakdown of metabolites such as propanediol, ethanolamine and choline, compounds often found in the gut mucosa (Bobik, 2006, Bobik *et al.*, 1999, Cheng *et al.*, 2008, Kerfeld *et al.*, 2018). These degradative BMCs are referred to as metabolosomes (Brinsmade *et al.*, 2005). In all cases studied so far, metabolosomes appear to produce an aldehyde intermediate that is then converted into its alcohol and acid derivatives. The need to encapsulate such metabolic processes appears to protect the cell from potentially toxic intermediates and at the same time promotes metabolic flux (Jakobson *et al.*, 2017, Penrod & Roth, 2006, Sampson & Bobik, 2008).

BMCs are composed of a number of different shell proteins that form either homogeneous hexameric or pentameric tiles, which link together to form the facets and vertices of the macromolecular structure respectively (Kerfeld *et al.*, 2005, Tanaka *et al.*, 2008). The structure of a small hollow recombinant BMC shell has recently been reported, providing strong evidence on the orientation of the tiles and how they associate to give the overall assembly (Sutter *et al.*, 2017). The shell protein tiles have a central pore that is thought to allow metabolites and cofactors into and out of the complex, possibly through a gated mechanism. Many of the metabolic cargo enzymes that are found within the lumen of the BMC are known to contain an N-terminal, or less frequently a C-terminal, encapsulation peptide that is around 18-20 amino acids in length, and which forms an amphipathic helix (Aussignargues *et al.*, 2015, Fan & Bobik, 2011, Fan *et al.*, 2010, Lawrence *et al.*, 2014, Kinney *et al.*, 2012). For

metabolosomes, studies have shown that the N-terminal 18 amino acids of PduD (D18) (Fan & Bobik, 2011) and PduP (P18) (Fan *et al.*, 2010), as well as the first twenty amino acids of PduL (L20) (Liu *et al.*, 2015), can act as targeting peptides that can direct foreign cargo, such as fluorescent proteins, to the BMC.

Research into the biogenesis of BMCs has largely focussed on the carboxysomes, which generally form larger and more symmetrical icosahedral structures than the metabolosomes (Kerfeld & Melnicki, 2016, Iancu *et al.*, 2007). Two types of carboxysomes are known, with the β-carboxysome structural proteins sharing high homology with catabolic BMC proteins. Imaging has shown that the protein shell of β-carboxysomes is able to form around an aggregate of the protein cargo, with the size and shape of the BMC defined by the architecture of the shell proteins (Cameron *et al.*, 2013, Chen *et al.*, 2013). This core-first assembly involves the formation of an initial aggregate that gains its shape from the interaction of the shell proteins, which as it closes pinches off the excess aggregate, allowing it to form another BMC. A different assembly process occurs within the α-carboxysome, where some key proteins help in this fabrication process, including CsoS2 (Cai *et al.*, 2015). This concomitant folding involves the assembly of shell and core proteins together with the chaperoning of core proteins. However, the mechanisms of BMC formation are likely not universal. Metabolosomes are much more irregular in both their size and shape in comparison to carboxysomes. Furthermore, through the recombinant production of BMC shell proteins derived from the propanediol utilisation system it has been shown that the shell proteins can self-assemble into a smaller BMC in the absence of any protein cargo (Parsons *et al.*, 2010, Mayer *et al.*, 2016). Thus, shell proteins have the ability to self-assemble into empty structures.

BMCs represent powerful metabolic units in the form of small bioreactors and as such they have caught the attention of metabolic engineers who see these organelles as a way to redesign specific aspects of cellular metabolism for the production of fine and commodity chemicals (Frank *et al.*, 2013, Kim & Tullman-Ercek, 2013, Lawrence *et al.*, 2014, Lee *et al.*, 2018b, Liang *et al.*, 2017, Mayer *et al.*, 2016). Key to the redesign of BMCs with alternative

metabolic functions is the ability to encapsulate new pathways into the structures. Evidence that BMCs can be reengineered has come from a proof-of-principle study whereby a pyruvate decarboxylase (PDC) and an alcohol dehydrogenase were directed to an empty BMC through the addition of encapsulation peptides (Lawrence *et al.*, 2014). The resulting BMC was found to have the ability to convert pyruvate into ethanol.

Although a broad range of potential encapsulation peptides have been identified, comparative studies on their ability to direct cargo to BMCs have been limited. Screens to report improved changes to the encapsulation peptide have been reported but these are based on the use of fluorescent proteins and their localisation to puncta in cells by fluorescent microscopy (Kim & Tullman-Ercek, 2014). However, as encapsulation peptides promote aggregation, which can be visualised also as fluorescent foci, the use of fluorescent puncta as a means to judge localisation does not differentiate between protein aggregation and BMC encapsulation (Lee *et al.*, 2018b).

Herein, we describe research that was undertaken to develop improved methods to quantitate the incorporation of cargo into BMCs. Initially, the ability of the *Citrobacter freundii* PduD, PduL and PduP to direct cargo to an empty recombinant *C. freundii* Pdu BMC was investigated. Thereafter, the *C. freundii* D18, L20 and P18 tags were investigated for their effect on the activity of a range of PDCs in order to inform on the choice of tag. The peptides were also studied for their ability to co-aggregate different proteins *in vivo*. We demonstrate that metallothionein can be used to follow aggregation and encapsulation as it is easily identifiable by TEM due to its increased electron density, which is most likely caused by its metal binding capacity. Finally, we present electron tomography studies of empty recombinant *C. freundii* Pdu BMCs and metallothionein targeted to these BMCs. Collectively, this work demonstrates a correlation between cargo-protein aggregation, the level of assimilation into BMCs and their resulting size.

Results

Targeting of native enzymes into recombinant Pdu BMCs

In order to understand the relative abilities of the encapsulation peptides found on *C. freundii* PduD (Fan & Bobik, 2011), PduL (Liu *et al.*, 2015) and PduP (Fan *et al.*, 2010) to mediate cargo encapsulation into a recombinant *C. freundii* Pdu BMC, we first sought to look at how well these proteins were integrated into an empty BMC. We therefore individually produced PduD, PduL and PduP in the presence and absence of empty Pdu BMCs (PduA-U) (Parsons *et al.*, 2010). When these BMC-associated enzymes were produced in the absence of compartments, large inclusion bodies were observed in a high proportion of cells (PduD: 55.8%; PduL: 67.0%; PduP 71.7%; **Figure A1 (Appendix 1)**) with no such structures present in an empty vector control strain (**Figure 1**). This observation is consistent with the prediction that these proteins self-aggregate.

Co-production of PduD or PduL with PduA-U resulted in a decrease in the observed aggregation (**Figure 1 and Figure A1 (Appendix 1)**). The inclusions observed in these cells were phenotypic of misfolded BMC shell proteins (**Figure A2 (Appendix 1)**), suggesting that the majority of PduD and PduL were encapsulated. However, co-production of PduP with BMCs only showed a slight decrease in the percentage of cells containing inclusion bodies (**Figure A1 (Appendix 1)**), although a high proportion of cells (58.1%) still contained inclusions, suggesting lower encapsulation efficiency or greater aggregation rate. Electron microscopy analysis was also carried out on purified BMCs from these strains (**Figure 1**). Areas of darker staining within the isolated BMCs indicate that the BMCs contain protein cargo and thus the number of BMCs containing darker staining can be used to gauge targeting to the lumen of intact microcompartments (**Figure A3 (Appendix 1)**). Darker electron density within the BMC is therefore indicative of protein encapsulation. These BMCs were compared to BMCs produced in the absence of cargo. Electron dense regions were observed in the lumen of BMCs isolated from the strains producing PduD, PduL or PduP, suggesting that all three proteins are incorporated into the recombinant BMCs. The efficiency of protein

incorporation displayed a degree of variation with both PduP and PduD showing a higher level of assimilation in comparison to PduL.

Effect of Pdu BMC encapsulation tags on three different pyruvate decarboxylases

Previously, we reported the effect of encapsulation tags on the activity of pathway enzymes associated with propanediol synthesis (Lee *et al.*, 2016). Herein, we investigated the effect of different encapsulation peptides on the activity of PDC, an enzyme we had previously shown could be targeted to a recombinant BMC (Lawrence *et al.*, 2014), and explored if the effects were identical on three homologous PDCs. The selected PDCs have a high level of protein sequence similarity (**Figure A4 (Appendix 1)**) as this would allow us to determine if the addition of encapsulation peptides have a reproducible effect on enzymatic activity.

In order to undertake this comparison we looked at the effect on enzyme activity after fusing three different BMC encapsulation peptides (D18, L20 and P18) onto PDCs from *Gluconacetobacter diazotrophicus* (GdPDC), *Zymomonas mobilis* (ZmPDC) and *Zymobacter palmae* (ZpPDC). The encapsulation peptides were fused to the N-terminus of the PDCs together with a hexa-histidine linker and a thrombin cleavage site. The control construct contained only a PDC with a hexa-histidine linker and a thrombin cleavage site. The activities of the purified tagged (D18, L20, P18) and untagged enzymes were followed in a linked assay. A comparison of the activities is shown in **Figure 2.**

The data show that despite a high level of sequence similarity between the enzymes (**Figure A4 (Appendix 1)**), particularly at the N-terminus where the encapsulation peptide is located, the different tags have a variable effect on activity. For instance, in comparison to the His-only control, the different tags did not greatly affect GdPDC activity, whereas the V_{max} of ZpPDC was reduced by around 50% when the protein was tagged with a D18 or P18 encapsulation peptide. Overall, the L20 tag was found to have the least disruptive effect on the activity of the enzymes. The activity results suggest that the effect of the encapsulation tags is not

dependent on sequence similarity, making it difficult to predict the behaviour of a tagged enzyme.

A comparative analysis of encapsulation efficiencies

Although mutagenesis approaches to improving the ability of encapsulation peptides to promote assimilation have been reported (Kim & Tullman-Ercek, 2014), no direct comparison has been undertaken to determine the relative targeting efficiencies of the various encapsulation peptides that have been used to direct non-natural cargo to recombinant BMCs. However, knowledge on the relative effectiveness of the different peptides would be a key facilitator of BMC technology in an industrial setting.

To investigate the encapsulation efficiencies of the D18, L20 and P18 encapsulation peptides we utilised a combination of confocal fluorescence microscopy and immuno-TEM. A yellow fluorescent protein (Citrine) was fused to the D18, L20 and P18 tags. The genetic constructs for these fused proteins were co-expressed with either recombinant BMCs containing an mCherry fused PduA (mA-U) (Parsons *et al.*, 2010) or an empty vector control (pLysS) and analysed (**Figure 3**).

Expression of either D18 or P18 tagged Citrine, in the absence of BMCs, resulted in the appearance of inclusion bodies, which were visible by both microscopy methods. In contrast, the L20 tagged protein appeared more soluble, but was observed to localise to the periphery of the cell. Analysis of a strain producing only mA-U revealed the presence of red puncta within the cytoplasm, an observation that has been attributed to the formation of BMCs within the cell (Parsons *et al.*, 2010). When co-produced with mA-U, both D18- and P18-tagged Citrine localised to these red punctate regions, which is indicative of co-localisation with BMCs.

In contrast, the L20-tagged protein did not appear to target specifically to the punctate regions of the cells, although there was a change in the observed phenotype with the signal more evenly dispersed throughout the cell. Quantification of the inclusion bodies present in these strains by TEM (**Figure A5 (Appendix 1)**) did not show a reduction when co-produced with

BMCs as has been observed previously with the full-length PduD, PduL and PduP proteins (**Figure A1 (Appendix 1)**). The difference may be explained by the modification of PduA in the mA-U construct as the addition of the mCherry label to PduA may alter BMC stability or result in increased shell protein aggregation.

Co-aggregation of tagged proteins

It would appear that the presence of D18 and P18 on proteins aids in the formation of cellular inclusion bodies. Previous work had also shown that the expression of 4 proteins involved in a 1,2-propanediol synthesis pathway tagged with encapsulation peptides results in the formation of a single large inclusion body, which is thought to contain all of the tagged components (Lee *et al.*, 2016). To investigate this aggregation phenotype more thoroughly we co-produced two fluorescent proteins (mCherry and Citrine) containing the various Pdu targeting tags to see if the fluorescent signals localise to the same region of the cell and form puncta indicative of protein aggregation (**Figure 4**).

Co-expression of His-mCherry with His-Citrine resulted in dispersed fluorescence throughout the cytoplasm (**Figure A6 (Appendix 1)**). However, the introduction of either a D18 or P18 tag onto the N-terminus of mCherry resulted in its aggregation (**Figures A7 and A8 (Appendix 1)**), whilst the His-Citrine fluorescent signal remained cytoplasmic. In contrast, the production of L20-mCherry resulted in mostly cytoplasmic fluorescence suggesting that this targeting peptide has a lower predisposition for aggregation (**Figure A9 (Appendix 1)**)

When the fluorescent proteins were tagged with either the D18 or P18 encapsulation peptides co-aggregation was observed, demonstrating that the tagged components are aggregated into a single inclusion body rather than separate bodies (**Figure 4**). Again, the L20 tagged fluorescent proteins appeared much more soluble than their P18 or D18 equivalents, although when the L20-fluorescent proteins were co-produced together with a different fluorescent protein tagged with either a D18 or P18 tag, the fluorescent signals appeared to co-localise with the punctate aggregates (**Figure 4 and Figure A9 (Appendix 1)**). Thus it would appear

that whilst the L20 encapsulation peptide is more soluble it interacts with D18- and P18-tags leading to co-aggregation.

Targeting metallothionein to recombinant Pdu BMCs

In order to study further the efficiency of targeting recombinant proteins into BMCs we utilised a metal binding protein, metallothionein, due to its propensity to bind a variety of metal ions (Kagi & Schaffer, 1988) making it comparatively easy to visualise within the cell by TEM when aggregated. The idea behind the metallothionein approach is that it should allow the opportunity to observe encapsulated aggregates in whole cell sections as opposed to relying on *in vitro* data obtained from purified compartments, allowing a comparison of their relative size, shape and volume. We therefore tagged a metallothionein from *Fucus vesiculosus* (Morris *et al.*, 1999) (fvMT) with the three different encapsulation peptides as well as a His-tag control. These constructs were co-produced with PduA-U and the resulting strains, as well as purified BMCs, were analysed by TEM (**Figure 5**). As expected, the use of fvMT did indeed allow the protein aggregates to be easily identified by TEM due to the increased electron density. Thin sections of cells from strains producing only the His-fvMT, D18-fvMT, L20-fvMT and P18-fvMT all showed the presence of a large aggregate in the absence of BMCs. When this was repeated with strains producing not only the tagged fvMT but also empty compartments (PduA-U), smaller bodies with sharp edges were observed, which is indicative of encapsulation into a BMC.

Quantification indicated that the presence of either the D18 or L20 peptides on the fvMT resulted in efficient encapsulation (99.3% and 96.1% respectively; **Figure A10 (Appendix 1)),** in that the vast majority fvMT was found associated with BMC rather than an inclusion body. By way of contrast, the P18 tagged fvMT shows only roughly half that level (54.0%) with an even distribution between aggregates and BMCs. The presence of an encapsulation peptide resulted in an overall decrease in the presence of non-encapsulated aggregation with the D18-tag proving to be the most effective (**Figure A11 (Appendix 1)**). We also observed a high proportion of empty BMCs present in the cells (Figure 5, third row), suggesting that the ratio

of cargo to shell production in a recombinant system requires further optimisation. Significantly, the BMCs observed with the tagged-fvMT all appear much larger than the BMCs isolated from the untagged fvMT. To demonstrate that the angular electron-dense bodies are BMCs the thin sections of the various strains were analysed by immuno-TEM using an anti-PduA antibody (**Figure A12 (Appendix 1)**). If the angular structures are BMCs then we would expect them to cross-react with the anti-PduA antibody, is what was observed and consistent with these aggregates being surrounded by a BMC shell.

The purified compartments produced in these various strains were also analysed by TEM (**Figure 5** and **Figure A12 (Appendix 1)**). Significantly, the BMCs co-produced with His-fvMT did not appear to have any cargo present after purification. The BMCs isolated from strains with the fvMT-tagged with the encapsulation peptides had various observed incorporation efficiencies (D18 – 24.7%; L20 – 16.5%; P18 – 30.5%). The purified BMCs co-produced with P18-fvMT also contained what appeared to be large inclusion bodies (**Figure A13 (Appendix 1)**). We suggest that these are large P18-fvMT inclusions that co-purify non-specifically with the compartments. Inclusion bodies like these are likely to pull down during BMC purification due to their large size, highlighting the importance of using TEM for this analysis as well in addition to standard western blot analysis to define encapsulation of proteins into BMCs.

The fvMT-producing strains were further analysed by electron tomography in order to gain a better visualisation of the three-dimensional topography of these structures (**Figure 5 and Supplementary Video 1**). This approach revealed the structures formed by fvMT and BMC shell proteins are remarkably varied in size, shape and volume. Furthermore, AMIRA software was used to gather quantitative data analysing enclosed structures in these tomograms (Weber *et al.*, 2012). This approach allowed us to gain further insight into the 3D structure (**Figure 6**) of empty recombinant BMCs (**Supplementary Video 2**) and recombinant BMCs containing L20-fvMT (**Supplementary Video 3**) allowing us to quantitate the volume (empty: 54900±11013 nm^3 (n=29); L20-fvMT: 336411±177722 nm^3 (n=60)) and the largest diameter (empty: 61.77±15.38 nm (n=29); L20-fvMT: 127.51±60.97 nm (n=60)) of these structures. This

data suggests an average volumetric expansion of around 6 times to accommodate L20 targeted fvMT. Modelling suggests that strong cargo-cargo interactions lead to an increase in compartment size, while strong shell-shell interactions lead to smaller, better defined structures (Mohajerani & Hagan, 2018, Perlmutter *et al.*, 2016). Our observations therefore are indicative of an assembly mechanism dominated by strong cargo-cargo interactions. These findings could also explain the previously observed variability in BMC size and shape (Mayer *et al.*, 2016).

Discussion

It had previously been observed that the removal of the N-terminal extension found on some of the diol dehydratases linked with BMC-catabolic processes improved their solubility, allowing for their structure determination (Fan *et al.*, 2010). Studies on the structure of the P18 peptide revealed that it forms an amphipathic helix that encourages self-association through a coiled-coil interaction (Lawrence *et al.*, 2014). In this way the encapsulation peptides promote protein aggregation. The protein aggregate must then be able to interact with the luminal side of the BMC shell. Indeed, the identification of P18 as an encapsulation peptide also led to suggestions that the peptide may interact with one particular component of the shell, PduA. Specifically, modelling studies suggested that the P18 peptide could interact with a helical region of PduA (Fan *et al.*, 2012). However, this region of PduA is on the concave side of the PduA hexamer. If the recent structure determination of a recombinant BMC from *Haliangium ochraceum* (Sutter *et al.*, 2017) is an accurate representation of a wild type compartment, and all the shell proteins have their concave side facing into the cytoplasm, then this encapsulation model will need to be modified to explain how the cargo protein becomes localised within the lumen.

By studying the production of PduD, PduL and PduP in the presence and absence of BMCs we have shown that, individually, PduD, PduL and PduP all form aggregates in the cell. However, in the presence of BMCs the majority of PduD and PduL become incorporated into the BMCs, as viewed by higher electron density within isolated BMCs and the reduction in the

observable intracellular aggregation. High levels of intracellular aggregation were still observed for PduP, although targeting was confirmed within BMCs. Using TEM for investigation of higher density within the BMCs represents an important technique to help in the validation that protein cargo is being localised within the BMC. This can be used to help support evaluation of localisation with fluorescent proteins and the presence of fluorescence puncta within cells.

It was interesting to observe the effect of the addition of encapsulation peptides to a range of PDCs from different organisms. We thought that there may be some common effects of these encapsulation peptides on the activities of the homologous enzymes. For instance, we would have predicted that L20 would have less of an effect on activity than the other tags. In fact, we could not observe any specific trend with the peptide. The tags had little to no effect on the activity of the PDC from *G. diazotrophicus*, whereas the P18 and D18 tags had a clear effect on the *Z. palmae* PDC. This means that predicting the effect of the addition of a targeting peptide to an enzyme is likely to be very challenging.

Attachment of the encapsulation peptides to fluorescent proteins allowed for a study of their co-localisation in the absence of BMCs. The work shows that fluorescent proteins containing targeting tags can co-aggregate together prior to encapsulation. These results suggest a possible assembly mechanism where cargo proteins co-aggregate together prior to encapsulation.

To gauge the effect of targeting to BMCs, one really needs to be able to see if protein has been internalised within the structure and also to be able to measure the size and shape of the compartment. We were able to achieve this through the use of metallothionein where aggregated protein is easily identifiable within the cell due to its extra density caused by the acquisition of metal ions. Using fvMT we were able to demonstrate that all the encapsulation tags, when attached to the protein, cause encapsulation within recombinant BMCs. These studies are all consistent with a model for BMC formation whereby shell proteins interact with an initial aggregate. If the aggregate forms too quickly then the BMC cannot keep pace with

the aggregate and the BMC does not encapsulate the cargo which results in the formation of a large inclusion within the cell as seen with P18-fvMT (**Figure 5** and **Figure A10 (Appendix 1)).** However, if enough shell protein is available the shell is able to form around the aggregate with high flexibility. The extent of fvMT aggregation was differentially influenced by the various tags. L20 encapsulated fvMT aggregates were smallest, followed by D18, while P18 fvMT produced largest structures, which was often observed to form un-encapsulated inclusions. The extent of aggregation caused by the tags seems to be conserved for all experiments, suggesting that P18 tag causes the largest amount of aggregation, followed by the D18 tag, whilst L20 tag does not cause large amounts of aggregation, but can still effectively target to compartments if the protein is able to aggregate by itself, as apparent in the fvMT experiments.

In the absence of any cargo the shell proteins are able to associate together and generate comparatively small structures. Overall, the size and shape of BMCs appears to be primarily dictated by cargo-cargo and shell-cargo interactions, explaining why catabolic BMCs have such a varied topology as predicted by computational and theoretical modeling (Mohajerani & Hagan, 2018).

Experimental Procedures

Molecular biology and bacterial strains

DNA encoding PduL20 was synthesised and cloned into the *BglII* and *Nde*I sites of pET14b (**Sequence A1(Appendix 3)**). Genomic DNA used for PCR reactions was supplied by DSMZ (**Table A1 (Appendix 2)**). Primers (**Table A2 (Appendix 2)**), plasmids (**Table A3 (Appendix 2)**) and bacterial strains (**Table A4 (Appendix 2)**) used are available in the appendix. Molecular biology was carried out in *E. coli* JM109 strain, while all other experiments were carried out in *E. coli* BL21 Star (DE3) strain.

Growth of strains

BL21 Star (DE3) competent cells were transformed with appropriate plasmids. LB supplemented with ampicillin (100 mg/L) and chloramphenicol (34 mg/L) in baffled flasks was inoculated from an overnight starter culture. Cells were grown at 37 °C with shaking to an OD600 ~ 0.4, protein production was induced by addition of IPTG to a final concentration of 400 µM. Cultures were incubated for 4 h (confocal) or overnight at 19 °C with shaking.

TEM analysis

Cells were harvested by centrifugation at 3000 x g for 10 minutes. The cell pellet was resuspended in 2 mL 2.5% (w/v) glutaraldehyde in 100 mM sodium cacodylate buffer, pH 7.2, (CAB) and fixed for 2 hours with gentle rotating (20 rpm). Cells were pelleted by centrifugation at 6000 x g for 2 minutes and were washed twice for 10 minutes with 100 mM CAB. Cells were post-fixed with 1% (w/v) osmium tetroxide in 100 mM CAB for 2 hours and subsequently washed twice with ddH2O. Cells were dehydrated by incubation in an ethanol gradient, 50% EtOH for 10 minutes, 70% EtOH overnight, 90% EtOH for 10 minutes followed by three 10 minute washes in 100% dry EtOH. Cells were then washed twice with propylene oxide for 15 minutes. Cell pellets were embedded by resuspension in 1 mL of a 1:1 mix of propylene oxide and Agar LV Resin and incubated for 30 minutes with rotation. Cell pellets were infiltrated twice in 100% Agar LV resin. The cell pellet was re- suspended in fresh resin and transferred to a 1 mL Beem embedding capsule, centrifuged for 5 minutes at 3000 x g to concentrate the cells to the tip of the mould and incubated for 20 hours at 60 °C to polymerize.

Samples were ultra-thin sectioned on a RMC MT-XL ultra-microtome with a diamond knife (diatome 45°). Sections (60 – 70 nm) were collected on un-coated 300 mesh copper grids. Grids were stained by incubation in 4.5% (w/v) uranyl acetate in 1% (v/v) acetic acid solution for 45 minutes followed by washing in a stream of ddH_2O. Grids were then stained with Reynolds lead citrate for 7 minutes followed by washing in a stream of ddH_2O.

Electron microscopy was performed using a JEOL-1230 transmission electron microscope equipped with a Gatan multiscan digital camera operated at an accelerating voltage of 80 kV.

Purification of BMCs

Cells were harvested by centrifugation at 2683 x g. A 1 g wet cell pellet was resuspended in 20 mL Yeast Protein Extraction Reagent (Thermo Scientific) supplemented with Protease Inhibitor Cocktail Tablets, EDTA-Free (Sigma-Aldrich) and 500 Units Benzonase® Nuclease (Merck) and incubated for 3 hours at room temperature with gentle shaking. Cell lysate was pelleted by centrifugation for 5 minutes at 11,300 X g, the pellet was resuspended in 2 mL of 20 mM Tris-HCl, pH 8.0, 20 mM NaCl. The suspension was centrifuged at 4 °C for 5 minutes at 11,000 X g and the supernatant was collected. The NaCl concentration was increased to 80 mM with 5 M NaCl, this was then centrifuged at 4 °C for 5 minutes at 11,000 X g. The pellet was resuspended in 1 mL of 20 mM Tris-HCl, pH 8.0, and was clarified by centrifugation at 4°C for 5 minutes at 11,000 X g. The supernatant contain microcompartments was collected for analysis.

Pyruvate decarboxylase activity assay

Protein purification was carried out as described previously (Lee *et al.*, 2018a). Purified protein was buffer exchanged using a PD10 column (GE Healthcare) into 50mM Na-phosphate pH 7.0, 5mM $MgSO_4$, 0.1mM thiamine pyrophosphate buffer. Enzyme concentration was estimated using absorbance at 280nm and diluted to a stock concentration of 0.1 mg/mL. PDC activity was measured using an alcohol dehydrogenase coupled assay (Gounaris *et al.*, 1971). Briefly, pyruvate is decarboxylated by PDC leading to the production of acetaldehyde which is subsequently reduced by alcohol dehydrogenase (ADH) using NADH. The oxidation of NADH is measured at 340nm and the rate of the reaction is calculated using Michaelis-Menten equation. Reactions contained 0.15mM NADH, 20U ADH (ADH from *S. cerevisiae*; Sigma), 1µg PDC, 50µM-10mM Pyruvate, which was added last. All measurements were carried out at 25^{o}C in 50mM Na-phosphate buffer, pH 7.0, containing 5mM MgSO4 and 0.1mM thiamine pyrophosphate.

Confocal imaging

Following growth and induction of protein expression 1 mL of cells was harvested by centrifugation at 3000 x g. The resulting cell pellet was washed 3 times in PBS before incubation for 15 minutes in 2% (w/v) formaldehyde in PBS, cells were then washed a further 3 times in PBS. Cells (10 µL) were pipetted onto a 1.5 thickness coverslip before being inverted onto a drop of ProLong Gold antifade mountant (Life Technologies) on a glass slide Slides were incubated at room temperature in the dark for 24 hours to cure.

Images were acquired on a Zeiss LSM 880 with Airyscan system. Excitation light (514 nm for mCitrine or 561 nm for mCherry) was provided by an argon lamp (514nm) or HeNe Laser (561nm). Images were acquired using a 100x 1.46 NA oil immersion objective lens.

Immuno TEM

Strains were cultured as described previously, cells were harvested by centrifugation for 10 min at 3000 x g. The cell pellet was resuspended in 2% formaldehyde and 0.5% glutaraldehyde in 100 mM sodium cacodylate, pH 7.2, and incubated for 2 h with gentle rotating. Cells were pelleted by centrifugation at 6000 x g for 2 min and were washed twice for 10 min with 100 mM sodium cacodylate, pH 7.2. This was followed by dehydration of the samples in an ethanol gradient, 50% EtOH for 10 min, 70% EtOH for 10 min, 90% EtOH for 10 min, followed by three 15 min washes in 100% EtOH. Cell pellets were then resuspended in 2 mL LR white resin and incubated overnight with rotation at room temperature after which the resin was changed and incubated for a further 6 h. Cell pellets were resuspended in fresh resin and transferred to 1 mL gelatine capsules and centrifuged at 4000 x g to pellet the cells at the tip. Samples were polymerised at 60 °C for 24 h. Samples were ultra-thin sectioned on a RMC MT-XL ultramicrotome with a diamond knife (diatome 45°) sections (60–70nm thick) were collected on 300 mesh gold grids.

Grids were equilibrated in one drop of TBST (20 mM Tris–HCl buffer, pH 7.2, containing 500 mM NaCl, 0.05% Tween 20 and 0.1% BSA) before being transferred into a drop of 2% BSA in TBST and incubated at room temperature for 30 min. Grids were then immediately

transferred into a 20 µL drop of relevant primary antibody (rabbit anti- PduA (Parsons *et al.*, 2010) or mouse anti-GFP, for detecting Citrine; Sigma Aldrich) and incubated for 1 h. Grids were washed in a fresh drop of TBST followed by washing for 10 s in a stream of TBST. Grids were equilibrated in a drop of relevant secondary antibody (Goat anti-rabbit IgG 10 nm or Goat anti-mouse IgG 10 nm gold (Agar Scientific)) then incubated for 30 min in a fresh drop. Excess antibody was removed by washing in two drops of TBST before washing in a stream of ddH_2O and dried.

Grids were stained for 15 minutes in 4.5% uranyl acetate in 1% acetic acid solution followed by 2 washes in dH2O. Grids were then stained with Reynolds lead citrate for 3 min followed by a wash in ddH2O. Electron microscopy was performed using a JEOL-1230 transmission electron microscope equipped with a Gatan multi-scan digital camera at an accelerating voltage of 80 kV.

Tomography

Sections (250 nm) were cut from the existing blocks as described above. Gold fiducials (15 nm, Aurion, TomoSol solution) were applied to both surfaces of the sections. The sections were imaged at 200 kV in a Tecnai 20 TEM (FEI, the Netherlands) and double tilt series images acquired between −67° to +69.5° (first axis) and −66° to +64.5° (second axis) with 1.5° (above 45°) and 2° increments (below 45°). The pixel size on the 4k by 4k FEI Eagle camera was 0.74 nm. The resulting tomograms were reconstructed and combined using IMOD software (Kremer *et al.*, 1996). The isolated microcompartments were segmented manually using the AMIRA software suite, as shown in **Supplementary Video 2**. Only structures fully covered by the tomogram section were analysed. AMIRA software animations were further used for visualizing and analysing the data.

Author contributions

Conceptualization: Rokas Juodeikis, Martin J. Warren

Formal Analysis: Rokas Juodeikis, Matthew J. Lee

Funding Acquisition: Paul Verkade, Derek N. Woolfson, Martin J. Warren

Investigation: Rokas Juodeikis, Matthew J. Lee, Matthias Mayer, Judith Mantell

Methodology: Rokas Juodeikis, Matthew J. Lee, Judith Mantell, Ian R. Brown

Project Administration: Paul Verkade, Derek N. Woolfson, Martin J. Warren

Resources: Derek N. Woolfson, Michael B. Prentice, Martin J. Warren

Supervision: Paul Verkade, Derek N. Woolfson, Stefanie Frank, Martin J. Warren

Visualization: Rokas Juodeikis, Matthew J. Lee

Writing – Original Draft Preparation: Rokas Juodeikis, Matthew J. Lee, Judith Mantell, Martin J. Warren

Writing – Review & Editing: Rokas Juodeikis, Matthew J. Lee, Judith Mantell, Martin J. Warren

Acknowledgements

We are grateful to the Biotechnology and Biological Sciences Research Council of the UK for a strategic LoLa Award to MJW, DNW and PV (BB/M002969/1).

Ethics statement

None required.

Conflict of interest

None declared.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request. Supplementary Videos are available on https://figshare.com/ with Supplementary Video 1, electron microscopy tilt series and tomographic reconstructions of cells producing recombinant Pdu BMCs with and without targeting peptide or control (His-) tagged fvMT available at DOI: 10.6084/m9.figshare.10252982, Supplementary Video 2, recombinant Pdu BMCs traced using Amira software available at DOI:

10.6084/m9.figshare.10252985 and Supplementary Video 3, recombinant Pdu BMCs containing L20-fvMT traced using Amira software available at DOI: 10.6084/m9.figshare.10252988

References

Aussignargues, C., Paasch, B.C., Gonzalez-Esquer, R., Erbilgin, O., and Kerfeld, C.A. (2015) Bacterial microcompartment assembly: The key role of encapsulation peptides. *Communicative & Integrative Biology* **8**: e1039755.

Axen, S.D., Erbilgin, O., and Kerfeld, C.A. (2014) A Taxonomy of bacterial microcompartment loci constructed by a novel scoring method. *PLoS Computational Biology* **10**: e1003898.

Bobik, T.A. (2006) Polyhedral organelles compartmenting bacterial metabolic processes. *Applied Microbiology and Biotechnology* **70**: 517-525.

Bobik, T.A., Havemann, G.D., Busch, R.J., Williams, D.S., and Aldrich, H.C. (1999) The propanediol utilization (*pdu*) operon of *Salmonella enterica* serovar Typhimurium LT2 includes genes necessary for formation of polyhedral organelles involved in coenzyme B_{12}-dependent 1, 2-propanediol degradation. *Journal of Bacteriology* **181**: 5967-5975.

Bobik, T.A., Lehman, B.P., and Yeates, T.O. (2015) Bacterial microcompartments: Widespread prokaryotic organelles for isolation and optimization of metabolic pathways. *Molecular Microbiology* **98**: 193-207.

Brinsmade, S.R., Paldon, T., and Escalante-Semerena, J.C. (2005) Minimal functions and physiological conditions required for growth of *Salmonella enterica* on ethanolamine in the absence of the metabolosome. *Journal of Bacteriology* **187**: 8039-8046.

Cai, F., Dou, Z., Bernstein, S.L., Leverenz, R., Williams, E.B., Heinhorst, S., Shively, J., Cannon, G.C., and Kerfeld, C.A. (2015) Advances in understanding carboxysome assembly in *Prochlorococcus* and *Synechococcus* implicate CsoS2 as a critical component. *Life (Basel)* **5**: 1141-1171.

Cameron, J.C., Wilson, S.C., Bernstein, S.L., and Kerfeld, C.A. (2013) Biogenesis of a bacterial organelle: The carboxysome assembly pathway. *Cell* **155**: 1131-1140.

Chen, A.H., Robinson-Mosher, A., Savage, D.F., Silver, P.A., and Polka, J.K. (2013) The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo. *PLoS One* **8**: e76127.

Cheng, S., Liu, Y., Crowley, C.S., Yeates, T.O., and Bobik, T.A. (2008) Bacterial microcompartments: their properties and paradoxes. *Bioessays* **30**: 1084-1095.

Fan, C., and Bobik, T.A. (2011) The N-terminal region of the medium subunit (PduD) packages adenosylcobalamin-dependent diol dehydratase (PduCDE) into the Pdu microcompartment. *Journal of Bacteriology* **193**: 5623-5628.

Fan, C., Cheng, S., Liu, Y., Escobar, C.M., Crowley, C.S., Jefferson, R.E., Yeates, T.O., and Bobik, T.A. (2010) Short N-terminal sequences package proteins into bacterial microcompartments. *Proceedings of the National Academy of Sciences* **107**: 7509-7514.

Fan, C., Cheng, S., Sinha, S., and Bobik, T.A. (2012) Interactions between the termini of lumen enzymes and shell proteins mediate enzyme encapsulation into bacterial microcompartments. *Proceedings of the National Academy of Sciences* **109**: 14995-15000.

Frank, S., Lawrence, A.D., Prentice, M.B., and Warren, M.J. (2013) Bacterial microcompartments moving into a synthetic biological world. *Journal of Biotechnology* **163**: 273-279.

Gounaris, A.D., Turkenkopf, I., Buckwald, S., and Young, A. (1971) Pyruvate decarboxylase. I. Protein dissociation into subunits under conditions in which thiamine pyrophosphate is released. *Journal of Biological Chemistry* **246**: 1302-1309.

Iancu, C.V., Ding, H.J., Morris, D.M., Dias, D.P., Gonzales, A.D., Martino, A., and Jensen, G.J. (2007) The structure of isolated *Synechococcus* strain WH8102 carboxysomes as revealed by electron cryotomography. *Journal of Molecular Biology* **372**: 764-773.

Jakobson, C.M., Tullman-Ercek, D., Slininger, M.F., and Mangan, N.M. (2017) A systems-level model reveals that 1,2-Propanediol utilization microcompartments enhance pathway flux through intermediate sequestration. *PLoS Computational Biology* **13**.

Kagi, J.H., and Schaffer, A. (1988) Biochemistry of metallothionein. *Biochemistry* **27**: 8509-8515.

Kerfeld, C.A., Aussignargues, C., Zarzycki, J., Cai, F., and Sutter, M. (2018) Bacterial microcompartments. *Nature Reviews Microbiology* **16**: 277-290.

Kerfeld, C.A., Heinhorst, S., and Cannon, G.C. (2010) Bacterial microcompartments. *Annual Reviews of Microbiology* **64**: 391-408.

Kerfeld, C.A., and Melnicki, M.R. (2016) Assembly, function and evolution of cyanobacterial carboxysomes. *Current Opinions in Plant Biology* **31**: 66-75.

Kerfeld, C.A., Sawaya, M.R., Tanaka, S., Nguyen, C.V., Phillips, M., Beeby, M., and Yeates, T.O. (2005) Protein structures forming the shell of primitive bacterial organelles. *Science* **309**: 936-938.

Kim, E.Y., and Tullman-Ercek, D. (2013) Engineering nanoscale protein compartments for synthetic organelles. *Current Opinion in Biotechnology* **24**: 627-632.

Kim, E.Y., and Tullman-Ercek, D. (2014) A rapid flow cytometry assay for the relative quantification of protein encapsulation into bacterial microcompartments. *Biotechnology Journal* **9**: 348-354.

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. *Journal of Biological Chemistry* **287**: 17729-17736.

Kremer, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996) Computer visualization of three-dimensional image data using IMOD. *Journal of Structural Biology* **116**: 71-76.

Lawrence, A.D., Frank, S., Newnham, S., Lee, M.J., Brown, I.R., Xue, W.-F., Rowe, M.L., Mulvihill, D.P., Prentice, M.B., and Howard, M.J. (2014) Solution structure of a bacterial microcompartment targeting peptide and its application in the construction of an ethanol bioreactor. *ACS Synthetic Biology* **3**: 454-465.

Lee, M.J., Brown, I.R., Juodeikis, R., Frank, S., and Warren, M.J. (2016) Employing bacterial microcompartment technology to engineer a shell-free enzyme-aggregate for enhanced 1,2-propanediol production in Escherichia coli. *Metabolic Engineering* **36**: 48-56.

Lee, M.J., Mantell, J., Hodgson, L., Alibhai, D., Fletcher, J.M., Brown, I.R., Frank, S., Xue, W.F., Verkade, P., Woolfson, D.N., and Warren, M.J. (2018a) Engineered synthetic scaffolds for organizing proteins within the bacterial cytoplasm. *Nature Chemical Biology* **14**: 142-147.

Lee, M.J., Palmer, D.J., and Warren, M.J. (2018b) Biotechnological advances in bacterial microcompartment technology. *Trends in Biotechnology* **37**: 325-336.

Liang, M., Frank, S., Lunsdorf, H., Warren, M.J., and Prentice, M.B. (2017) Bacterial microcompartment-directed polyphosphate kinase promotes stable polyphosphate accumulation in *E. coli*. *Biotechnol Journal* **12:** 1600415.

Liu, Y., Jorda, J., Yeates, T.O., and Bobik, T.A. (2015) The PduL phosphotransacylase is used to recycle coenzyme A within the Pdu microcompartment. *Journal of Bacteriology* **14**: 23922-23928.

Mayer, M.J., Juodeikis, R., Brown, I.R., Frank, S., Palmer, D.J., Deery, E., Beal, D.M., Xue, W.F., and Warren, M.J. (2016) Effect of bio-engineering on size, shape, composition and rigidity of bacterial microcompartments. *Scientific Reports* **6**: 36899.

Mohajerani, F., and Hagan, M.F. (2018) The role of the encapsulated cargo in microcompartment assembly. *PLoS Computational Biology* **14**: e1006351.

Morris, C.A., Nicolaus, B., Sampson, V., Harwood, J.L., and Kille, P. (1999) Identification and characterization of a recombinant metallothionein protein from a marine alga, *Fucus vesiculosus*. *Biochemical Journal* **338**: 553-560.

Parsons, J.B., Frank, S., Bhella, D., Liang, M., Prentice, M.B., Mulvihill, D.P., and Warren, M.J. (2010) Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. *Molecular Cell* **38**: 305-315.

Penrod, J.T., and Roth, J.R. (2006) Conserving a volatile metabolite: a role for carboxysome-like organelles in Salmonella enterica. *Journal of Bacteriology* **188**: 2865-2874.

Perlmutter, J.D., Mohajerani, F., and Hagan, M.F. (2016) Many-molecule encapsulation by an icosahedral shell. *eLife* **5**.

Rae, B.D., Long, B.M., Badger, M.R., and Price, G.D. (2013) Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO2 fixation in cyanobacteria and some proteobacteria. *Microbiology and Molecular Biology Reviews: MMBR* **77**: 357-379.

Sampson, E.M., and Bobik, T.A. (2008) Microcompartments for B_{12}-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate. *Journal of Bacteriology* **190**: 2966-2971.

Sutter, M., Greber, B., Aussignargues, C., and Kerfeld, C.A. (2017) Assembly principles and structure of a 6.5-MDa bacterial microcompartment shell. *Science* **356**: 1293-1297.

Tanaka, S., Kerfeld, C.A., Sawaya, M.R., Cai, F., Heinhorst, S., Cannon, G.C., and Yeates, T.O. (2008) Atomic-level models of the bacterial carboxysome shell. *Science* **319**: 1083-1086.

Weber, B., Greenan, G., Prohaska, S., Baum, D., Hege, H.C., Muller-Reichert, T., Hyman, A.A., and Verbavatz, J.M. (2012) Automated tracing of microtubules in electron tomograms of plastic embedded samples of Caenorhabditis elegans embryos. *Journal of Structural Biology* **178**: 129-138.

Yeates, T.O., Kerfeld, C.A., Heinhorst, S., Cannon, G.C., and Shively, J.M. (2008) Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nature Reviews Microbiology* **6**: 681-691.

Figure Legends:

Figure 1. Encapsulation of native proteins in recombinant Pdu BMCs. Electron micrographs of *Escherichia coli* cells producing PduD, PduL or PduP in the absence (top panel) and presence (middle panel) of a minimal BMC shell system (PduA-U). The BMCs extracted from strains in the middle panel are also shown. The control sample contains an empty vector. Arrows indicate areas of protein aggregation. All scale bars are 200 nm.

Figure 2. Comparison of the kinetic values of modified PDCs. Encapsulation peptides (D18, L20 and P18) were fused to three distinct PDCs. After recombinant production and purification the kinetic parameters of the various encapsulation-fused PDCs were measured in terms of Vmax (left scale bar) and Km (right scale bar) and expressed as a percentage of the activity of the PDC without the encapsulation tag. Light grey bar – V_{max}; dark grey bar – K_M. Assays were carried out in triplicate; error bars equal one standard deviation.

Figure 3. Targeting of fluorescent proteins to recombinant Pdu BMCs. Transmission electron micrographs (left two columns) and confocal microscopy (right two columns) images of cells expressing differentially tagged Citrine in the presence and absence of mCherry-tagged BMCs (mAU). Cells containing the empty pLysS vector are unable to produce BMCs whilst those containing mAU within the pLysS vector produce BMCs with an mCherry tag as evidenced by red puncta within the cytoplasm. The production of Citrine is visualised by the presence of yellow. Superimposition of red and yellow is indicative of localisation. Scale bars in TEM micrographs show 0.2 μm and in confocal images 2 μm.

Figure 4. Co-localisation of fluorescent proteins tagged with encapsulation peptides. Confocal microscopy was performed on encapsulation peptide-fused fluorescent proteins to determine if the encapsulation peptides interact with themselves or each other to co-aggregate. Citrine (yellow) and mCherry (red) fluorescent proteins tagged with or without the various encapsulation peptides (D18, L20 or P18) were co-produced and imaged as shown. All scale bars are 2 μm.

Figure 5. Production of metallothionein fused to various encapsulation peptides in the presence and absence of BMCs. TEM analysis of cells producing metallothionein tagged with the D18, L20 and P18 encapsulation peptides in the absence (top row) and presence (second row) of PduA-U BMCs. TEM of isolated BMCs from the strains in row 2 are shown in row 3. The bottom row shows TEM tomography of sections cells from row 2 - see **Supplementary Video 1** for full tomography data. All scale bars are 200 nm.

Figure 6. 3D reconstructions of Pdu microcompartments. Traced tomograms of PduA-U (left) and L20H-fMT PduA-U (right). All scale bars are 200nm.

Appendix 1.

Figure A1. Percentage of cells expressing the relevant proteins or an empty vector (control) containing intracellular aggregates. Black bars indicate no compartment control, white bars – coproduced with PduA-U. Error bars equal one standard deviation between three separate counts of a single biological repeat (n≥330).

Figure A2. Electron micrograph of PduD co-expressed with recombinant BMCs (PduA-U). Arrow showing aggregation which is phenotypic for misfolded BMCs. Scale bar is 0.2 μm.

Figure A3. Percentage of purified intact BMCs coproduced with the relevant proteins showing electron density within the lumen of the compartment indicative of successful targeting. Error bars equal one standard deviation between three separate counts of a single biological repeat (n≥440).

Figure A4. Protein sequence alignment of the analysed pyruvate decarboxylases from *Gluconacetobacter diazotrophicus* (Gd.PDC), *Zymobacter palmae* (Zp.PDC) and *Zymomonas mobilis* (Zm.PDC). Generated using the multalin tool (Corpet, 1988).

Figure A5. Quantification of aggregates observed in cells producing Citrine targeted to BMCs (n=150). Black bars indicate no compartment control, while white bars indicate co-production with mA-U. Only one count was carried out.

Figure A6. Co-expression of His-tagged mCherry with Citrine tagged with different BMC targeting peptides. Scale bars show are 2 μm.

Figure A7. Co-expression of D18-tagged mCherry with Citrine tagged with different BMC targeting peptides. Scale bars show are 2 μm.

Figure A8. Co-expression of P18-tagged mCherry with Citrine tagged with different BMC targeting peptides. Scale bars show are 2 μm.

Figure A9. Co-expression of L20-tagged mCherry with Citrine tagged with different BMC targeting peptides. Scale bars show are 2 μm.

Figure A10. Quantification of aggregates containing facets observed in cells producing fvMT targeted to BMCs. Black bars indicate no compartment control, white bars – coproduced with PduA-U. Error bars equal one standard deviation between three separate counts of a single biological repeat (n≥220).

Figure A11. Quantification of overall aggregation observed in cells producing fvMT targeted to BMCs. Black bars indicate no compartment control, white bars – coproduced with PduA-U. Error bars equal one standard deviation between three separate counts of a single biological repeat (n≥435).

Figure A12. Immuno-TEM of targeted fMT constructs produced with or without compartments showing the localisation of PduA. First row – cells with no PduA-U; second row – cells with protein coproduced with PduA-U. All scale bars are 200 nm.

Figure A13. Large aggregates observed in microcompartment purification. Scale bar is 0.2 µm.

Appendix 2

Table A1. Genomic DNA used in this study	
Organism	DSMZ Number
Gluconacetobacter diazotrophicus	DSM-5601
Zymobacter palmae	DSM-10491
Zymomonas mobilis	DSM-424

Table A2. Primers used in this study	
Primer Name	Sequence
FWGdPDCNde	GTACATATGACCTATACCGTTGGACGCTATCTC
RVGdPDCSpeSac	GATACTAGTTCAGAGCTCGCCCGCGCGCGGCTGGCGGGCGTTG
FWZmPDCNde	GTACATATGAGTTATACTGTCGGTACCTATTTAGCGGAG
RVZmPDCSpeSac	GTTACTAGTCTAGAGCTCGAGGAGCTTGTTAACAGGCTTACGGCTG
FWZpPDCNde	GATCATATGTATACCGTTGGTATGTACTTGGCAGAAC
RVZpPDCSpeSac	GTTACTAGTTTAGAGCTCCGCTTGTGGTTTGCGAGAGTTGGTAGCTG
Fv.fMT.FW.2	GTACATATGGCGGGCACTGGCTGCAAGATCTGGGAAGAC
Fv.fMT.RV	CATACTAGTCACTTGCCGCAGCCGCAGCAGTC

Table A3. Plasmids used in this study		
Plasmid name	Description	Source
pET14b	Overexpression vector containing N-terminal hexahistidine-tag, modified to include an *Spe*I site 5' of *Bam*HI	Novagen
pET14b-D18	Overexpression vector containing an N-terminal D18 targeting tag followed by a short amino acid linker (AMGSS) then a hexahistidine-tag	Lee *et al.,* 2016
pET14b-L20	Overexpression vector containing an N-terminal L20 targeting tag (first 20 amino acids of PduL from *Citrobacter freundii*) followed by a short amino acid linker (AMGSS) then a hexahistidine-tag Synthesised DNA sequence shown in **Sequence A1 (Appendix 3)**. *BglII/NdeI* ligated into *BglII/NdeI* site of pET14b vector.	This study
pET14b-P18	Overexpression vector containing an N-terminal P18 targeting tag followed by a short amino acid linker (PMGSS) then a hexahistidine-tag	Lee *et al.,* 2016
pLysS	Basal expression suppressor	Novagen
pLysS-PduABJKNU	pLysS containing genes required for the formation of empty BMCs	Parsons *et al*., 2010
pET3a-pduD	pET3a vector containing *pduD* from C*itrobacter freundii* ligated into *NdeI/SpeI* site	Parsons *et al*., 2010
pET3a-pduL	pET3a vector containing *pduL* from C*itrobacter freundii* ligated into *NdeI/SpeI* site	This study
pET3a-pduP	pET3a vector containing *pduP* from C*itrobacter freundii* ligated into *NdeI/SpeI* site	This study

pET3a-mCherryPduABB'JKNU	pET3a vector containing genes required for the formation of empty BMCs tagged with mCherry fluorescent protein	Parsons *et al.,* 2010
pET14b.GdPDC	PCR product of GdPDC ligated into *NdeI/SpeI* sites of pET14b	This study
pET.D18-GdPDC	*NdeI/SpeI* fragment from pET14b.GdPDC ligated into *NdeI/SpeI* site of pET14b-D18	This study
pET.L20-GdPDC	*NdeI/SpeI* fragment from pET14b.GdPDC ligated into *NdeI/SpeI* site of pET14b-L20	This study
pET.P18-GdPDC	*NdeI/SpeI* fragment from pET14b.GdPDC ligated into *NdeI/SpeI* site of pET14b-P18	This study
pET14b.ZmPDC	PCR product of ZmPDC ligated into *NdeI/SpeI* sites of pET14b	This study
pET.D18-ZmPDC	*NdeI/SpeI* fragment from pET14b.ZmPDC ligated into *NdeI/SpeI* site of pET14b-D18	This study
pET.L20-ZmPDC	*NdeI/SpeI* fragment from pET14b.ZmPDC ligated into *NdeI/SpeI* site of pET14b-L20	This study
pET.P18-ZmPDC	*NdeI/SpeI* fragment from pET14b.ZmPDC ligated into *NdeI/SpeI* site of pET14b-P18	This study
pET14b.ZpPDC	PCR product of ZpPDC ligated into *NdeI/SpeI* sites of pET14b	This study
pET.D18-ZpPDC	*NdeI/SpeI* fragment from pET14b.ZpPDC ligated into *NdeI/SpeI* site of pET14b-D18	This study
pET.L20-ZpPDC	*NdeI/SpeI* fragment from pET14b.ZpPDC ligated into *NdeI/SpeI* site of pET14b-L20	This study
pET.P18-ZpPDC	*NdeI/SpeI* fragment from pET14b.ZpPDC ligated into *NdeI/SpeI* site of pET14b-P18	This study
pET_CC_Di_A_Citrine	Plasmid containing *citrine* gene in the *NdeI/SpeI* site	Lee *et al.,* 2018
pET14b.Citrine	*NdeI/SpeI* fragment from pET_CC_Di_A_Citrine ligated into *NdeI/SpeI* site of pET14b	This study
pET.D18-Citrine	*NdeI/SpeI* fragment from pET14b.Citrine ligated into *NdeI/SpeI* site of pET14b-D18	This study
pET.L20-Citrine	*NdeI/SpeI* fragment from pET14b.Citrine ligated into *NdeI/SpeI* site of pET14b-L20	This study
pET.P18-Citrine	*NdeI/SpeI* fragment from pET14b.Citrine ligated into *NdeI/SpeI* site of pET14b-P18	This study
pET_CC_Di_A_mCherry	Plasmid containing *mCherry* gene in the *NdeI/SpeI* site	Lee *et al.,* 2018
pET14b.mCheery	*NdeI/SpeI* fragment from pET_CC_Di_A_mCherry ligated into *NdeI/SpeI* site of pET14b	This study
pET.D18-mCheery	*NdeI/SpeI* fragment from pET14b.mCheery ligated into *NdeI/SpeI* site of pET14b-D18	This study
pET.L20-mCheery	*NdeI/SpeI* fragment from pET14b.mCheery ligated into *NdeI/SpeI* site of pET14b-L20	This study
pET.P18-mCheery	*NdeI/SpeI* fragment from pET14b.mCheery ligated into *NdeI/SpeI* site of pET14b-P18	This study
pET14b.fvMT	PCR of the coding sequence of fvMT ligated into *NdeI/SpeI* site of pET14b	This study

pET.D18-fvMT	*NdeI/SpeI* fragment from pET14b.fvMT ligated into *NdeI/SpeI* site of pET14b-D18	This study
pET.L20-fvMT	*NdeI/SpeI* fragment from pET14b.fvMT ligated into *NdeI/SpeI* site of pET14b-L20	This study
pET.P18-fvMT	*NdeI/SpeI* fragment from pET14b.fvMT ligated into *NdeI/SpeI* site of pET14b-P18	This study

Table A4. Strains used in this study

Strain	Genotype	Source
JM109	endA1, recA1, gyrA96, thi, hsdR17 (rk^-, mk^+), relA1, supE44, Δ(lac-proAB), [F', traD36, proAB, laqIqZΔM15]	Promega
BL21 Star (DE3)	F^- ompT hsdSB (rB^- mB^-) gal dcm (DE3)	Novagen

Appendix 3

Sequence A1. Synthetic sequence used to construct pET14b-L20

>L20H

AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTG
TTTAACTTTAAGAAGGAGATATCATGGATAAACAGCAACTGGAGACAACGGTCACCAAAGTTCTGGATGAAA
TGCGTGAGCGCGCCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCC
ATATG