Title
Development of protein organelles for tunable metabolite diffusion and sequestration of multi-enzyme pathways

Author
Slininger, Marilyn F

Publication Date
2017

Peer reviewed|Thesis/dissertation

Development of protein organelles for tunable metabolite diffusion and sequestration of multi-enzyme pathways

By

Marilyn Faye Slininger

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:
Professor Danielle Tullman-Ercek, Co-Chair
Professor Jay Keasling, Co-Chair
Professor Wenjun Zhang
Professor David Savage

Spring 2017

ABSTRACT

Development of protein organelles for tunable metabolite diffusion and sequestration of multi-enzyme pathways

by

Marilyn Faye Slininger

Doctor of Philosophy in Chemical Engineering

University of California, Berkeley

Professor Jay Keasling and Professor Danielle Tullman-Ercek, Co-chairs

The use of bacterial production could improve commercial feasibility for biochemicals such as pharmaceuticals and natural products. However, many non-native reactions in bacterial hosts have unacceptably low yields due to pathway toxicity or competition with other cellular processes. Compartmentalization is a potential solution to these problems, but compartmentalized organelles were long thought to be found only in eukaryotic cells. Remarkably, bacterial microcompartments (MCPs) are protein-based organelles that perform this sequestration function. We aim to apply these MCPs for metabolic engineering and the enhancement of biosynthesis pathways that do not function optimally in the cytosol. MCPs may also serve as a scaffold for application outside of the cell.

To achieve this goal, we characterized the stability of MCPs when exposed to a range of pH and temperature to define the limitations of this complex structure. We next demonstrated that it is possible to form chimeric MCPs by combining shell proteins from different systems in nature. Importantly, we found that these changes to shell protein structure can enhance the function of the encapsulated pathway. We discovered that pathway performance can also be improved by a single point mutation to the pore of a shell protein, likely altering small molecule diffusion across the shell. We are investigating the role of shell protein pore chemistry in tuning small molecule diffusion. Finally, we test signal sequence peptides at the C-terminus of proteins and find that variation of signal peptide sequence results in altered expression level and loading within MCPs. In total, these tools improve our ability to design custom bacterial organelles to control the partitioning of chemical reactions within cells.

Dedicated to my parents, Patricia and Edward,
and my partner, Jason

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 4

LIST OF TABLES

LIST OF ABBREVIATIONS

MCP, microcompartment
Pdu, 1,2-propanediol utilization
Eut, ethanolamine utilization
Grp, glycyl radical-generating protein
Etu, ethanol utilization
TEM, transmission electron microscopy
WT, wild-type
GFP, green fluorescent protein
1,2-PD, 1,2-propanediol
PDV, prodeoxyviolacein
HPLC, high performance liquid chromatography
NCE, no carbon-E medium
TB, terrific broth
LB, lysogeny broth
aTc, anhydrotetracycline
MSA, multiple sequence alignment
RBS, ribosome binding site
RuBisCo, Ribulose-1,5-bisphosphate carboxylase/oxygenase
NAD+/H, nicotinamide adenine dinucleotide oxidized/reduced
BMC, bacterial microcompartment
SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis
OD, optical density
DLS, dynamic light scattering
NMR, nuclear magnetic resonance spectroscopy
PCR, polymerase chain reaction

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation (MCB1150567), the Army Research Office (W911NF-15-1-0144), and a gift from the ExxonMobil Corporation.

This work is the result of many contributions from the team including: undergraduate students Quan Tran, Matthew Aguirre, and Alyssa Barbosa; and graduate students Eddy Kim, and Chris Jakobson.

I also want to express extreme gratitude to the Tullman-Ercek Lab. My mentors Eddy Kim, Mike Fisher, Kevin Metcalf, Steph Lopez, Anum Azam, Jeff Glasgow, and Sergey Boyarskiy. You all helped to mold me into a competent researcher. I must also mention my other coworkers and friends Lisa, Han Teng, Emily, Mike A., Norma, Chris, Georg, Doug, Min, Taylor, Charlotte and Bon. Thank you for setting great examples, making the lab a fun and exciting place to work, and sitting through my qual practice about a zillion times.

Thank you to my PI, Danielle Tullman-Ercek, for being an incredible advisor and advocate. I have grown more than I could have imagined under your guidance. I credit you and the environment you built for my great graduate school experience.

And most of all, thank you to my family and my husband, Jason, who have been my greatest support.

Chapter 1: INTRODUCTION

1.1 Benefits of sequestration for enzymatic reactions

Spatial organization of related functional elements is a common theme in nature. In complex organisms, functions such as motility, energy harvesting, and information processing are divided among specialized organs and tissues. At the level of a single cell in eukaryotes, compartmentalization is achieved through the use of membrane-bound organelles. Initially, bacteria were thought to be simpler systems, lacking internal organization. However, research over time has uncovered more and more ways in which bacteria localize various functions. One example is the bacterial microcompartment (MCP).

The defining characteristic of the MCP is a protein shell encapsulating a series of enzymes, all participating in one pathway. Several catabolic pathways are found within MCPs including the metabolism of 1,2-propanediol, ethanolamine, L-fucose, L-rhamnose, fucoidan, and ethanol [1–5]. Another type of MCP called the carboxysome is found in cyanobacteria and encapsulates carbonic anhydrase and RuBisCo for carbon fixation [6–9]. At least four potential benefits may arise from encapsulating pathways and imposing a diffusion barrier for small molecules. First, sequestration can increase local concentrations of intermediates and enhance flux through "slow" downstream enzymes [10]. Second, the MCP shields the cell from toxic intermediates [11]. Third, the MCP may prevent loss of intermediate molecules that would otherwise participate in side reactions or escape from the cell [9,12]. Fourth, MCPs contain a private pool of necessary cofactors, driving pathway activity [13–15]. It is of interest to harness these advantages to express heterologous pathways in bacteria for the production of biochemicals. The ability to insulate pathways from the cytosol widens the spectrum of reactions that could be efficiently performed in the cell.

1.2 Characterization of MCP systems from nature

In the 1970's, the carboxysome, which contains the carbon-fixing enzyme, RuBisCo, was discovered in chemoautotrophs [6,7]. This was the first observation of MCPs, and evidence that bacteria had more intricate cellular organization than previously thought. Since then, genetic elements encoding similar protein organelle systems were discovered in up to 20% of sequenced bacterial species [16]. A number of pathways within MCPs have been experimentally characterized [1–9]. The sequestered reactions have a common theme: the enzymes performing the second reaction of each pathway need high concentrations of their substrate to reach an efficient kinetic regime. Also, intermediates of these pathways are often toxic and/or volatile.

The work described in this manuscript primarily focuses on two MCP systems: the 1,2-propanediol utilization (Pdu) and the ethanolamine utilization (Eut) MCPs from *Salmonella enterica* serovar Typhimurium LT2. In nature, *S. enterica* uses MCP systems to metabolize these niche carbon sources and proliferate in the human gut [17–19]. In both the Pdu and Eut pathways the first reaction is dependent on coenzyme B12 and forms a toxic aldehyde intermediate (Figure 1.1). The aldehyde is then converted to either an alcohol or a Coenzyme-A conjugated product which is consumed in downstream steps. Early work by Sampson and Bobik demonstrated that an *S. enterica* mutant that does not form a Pdu MCP shell releases toxic levels of

propionaldehyde to the growth medium [11]. In addition, there is evidence that NAD^+/NADH, coenzyme A, and coenzyme B12 may be recycled or reactivated within the MCP, allowing the formation of a private pool of cofactors [13–15]. My studies of these native systems has yielded methods for altering both the protein shell and the enzyme cargo of MCPs. We also gain insights which help to identify the biotechnologically relevant pathways most likely to benefit from encapsulation.

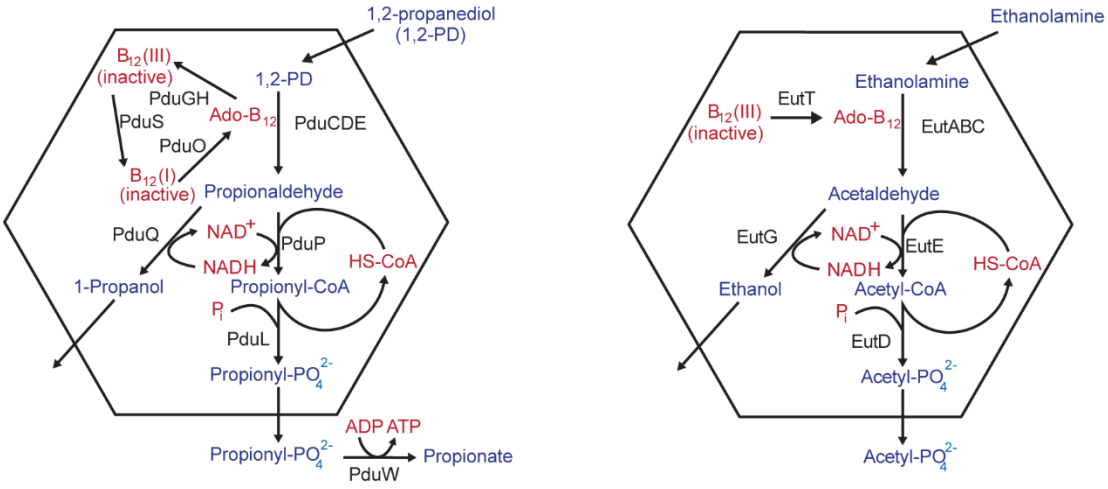


Figure 1.1 Pdu and Eut MCP pathway diagrams. Diagram illustrating the Pdu and Eut pathways including enzymes, cofactors, substrates, intermediates and products. The black hexagonal outline represents the boundary of the MCP. Enzymes outside of this boundary are not localized to the interior of the MCP.[20]

1.3 MCP shell and cargo assembly

The MCP shell is made of a homologous class of proteins called bacterial microcompartment (BMC) domains. These shell proteins form hundreds of hexamers or pentamers that tile together to form an enclosed shell between 50 and 200 nm in diameter (Figure 1.2**).** The packing symmetry of the monomers creates a central pore in each multimer about 0.5-1.5 nm in diameter. Each MCP system has between 1 and 10 different versions of shell protein making up the overall structure [5]. The Pdu compartment is made up of eight shell proteins (PduA, B, J, K, M, N, T and U)[21]. Similarly, the Eut MCP is composed of five shell proteins (EutS, M, N, L and K) that are homologous to the Pdu genes. The duplication of shell protein genes has allowed them to diverge in function. Each shell protein has unique structural properties that likely correspond with specialized roles [21–24].

The pores formed at the center of each shell protein multimer are possible routes for transport of small molecules across the shell boundary. Different shell proteins have various chemical environments in the pore that could modulate diffusion through the shell [25–28]. In the Pdu system, PduA is implicated in the transport of substrate molecule 1,2-PD [27]. Other shell protein pores have interesting characteristics that have yet to be explored fully. The crystal structure of EutL (similar in structure to PduB) was solved for two different conformations, hinting at a possible gated transport mechanism with the capability of allowing larger molecules such as cofactors to pass through the shell [23]. As mentioned above, cofactors like coenzyme

B_{12} and NADH are needed for pathway throughput. In contrast to substrate and product transport, once these factors are encompassed during shell formation, they may be recycled within the shell [13–15].

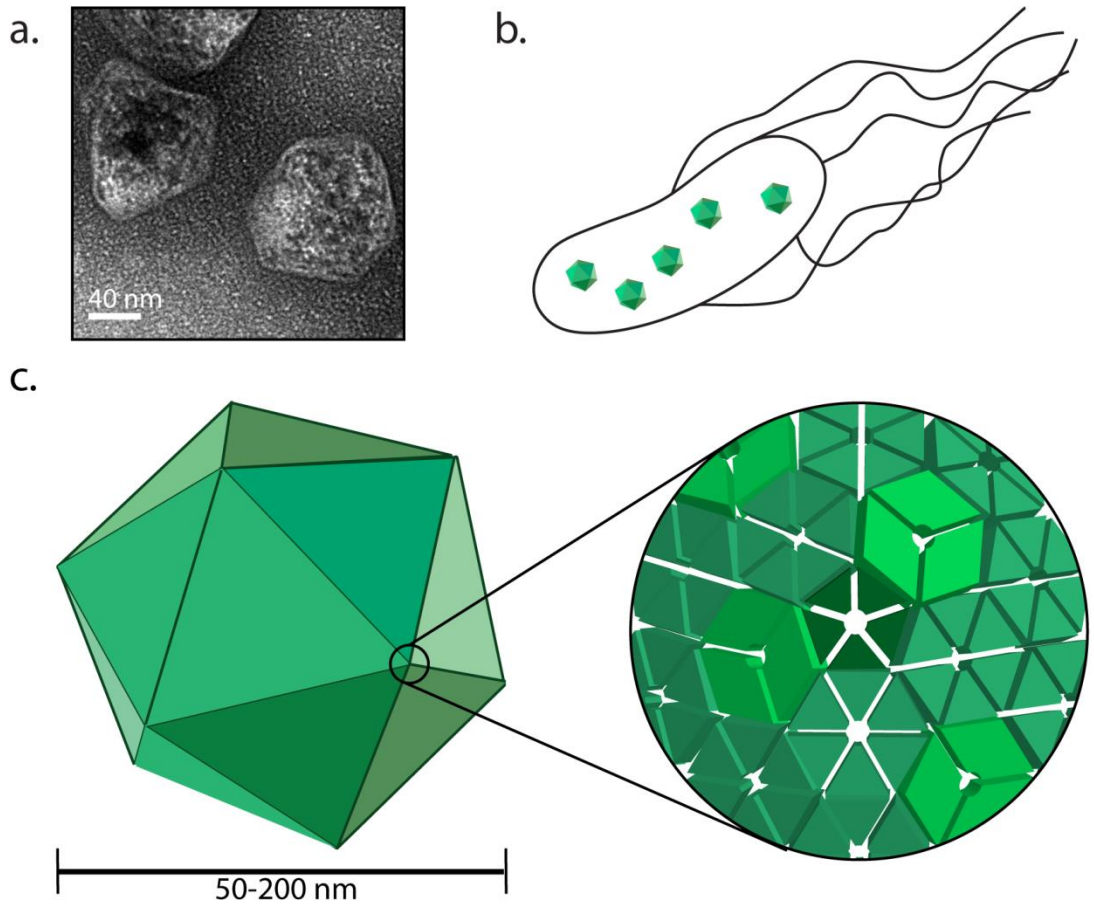


Figure 1.2 Overall architecture of the MCP. a) A transmission electron microscopy image of purified Pdu MCPs illustrating size scale and morphology. b) Illustration of several MCPs within *S. enterica*. d) Magnified illustration with cartoon model of hexamer and pentamer shell protein arrangement within the shell.

In addition to a variety of pore structures and possible transport phenotypes, nature has produced a range of MCP morphologies. MCPs vary in overall size, shape, and regularity. For example, an MCP of unknown function purified by expressing shell proteins from *Haliangium ochraceum* is round and less than 50 nm in diameter [29]. Pdu MCPs are irregularly faceted with a 150 nm diameter (Figure 1.2a). On the other hand, carboxysomes from cyanobacteria are very regular icosahedrons about 100 nm in diameter [30]. A better understanding of the relationships between MCP shell structure and functions will expedite engineering efforts to apply MCPs to encapsulate pathways of interest.

Bioinformatic analysis of MCP domain proteins across many species of bacteria shows a large degree of conservation in residues along the outer edge of all shell protein hexamers. According to binding models derived from observation of crystal structures, this outer edge is responsible for the protein interactions that hold multimers together in a shell structure. An

alanine scan of PduA resulted in a non-assembling mutant, $PduA^{K26A}$ that confirms this model of binding [31].

As the protein shell forms, enzyme cargo is recruited to the shell lumen. This is mediated by protein interactions between cargo and shell. These interactions are the subject of many recent investigations. For several well-studied MCP systems, peptides involved in interactions with the shell were identified. To illustrate, the Pdu MCP natively contains 7 enzymes, some of which are made up of multiple subunits (Figure 1.3a). Signal sequence peptides about 18-20 amino acids long that localize fused cargo to the MCP interior were identified at the N-termini of three Pdu enzymes: PduP, PduD, and PduL[20,32,33]. Structural characterization of the PduP signal sequence by NMR has revealed that it is alpha-helical [34]. Pull-down assays and protein interaction modeling predicts that the PduP N-terminus interacts with the C-terminus of shell protein PduA [32]. PduD is a subunit of the diol-dehydratase enzyme, and the other subunits (PduC and PduE) tag along by binding to PduD before or during MCP assembly (Figure 1.3a) [33]. The alchohol dehydrogenase, PduQ, is also encapsulated by binding to PduP [13]. The mechanism by which the other enzymes are recruited to the MCP is not yet clear, but it is likely that they form their own, undiscovered interactions with the MCP shell or with PduL, PduP, or PduCDE. For instance, in a recent study the shell protein PduB was implicated in binding cargo, though the PduB binding partner has not yet been identified [35]. Also, the enzyme PduS involved in coenzyme B_{12} recycling was found to interact with shell protein PduT, but the mechanism of this binding is unclear [36].

In a similar manner, N-terminal signal sequences were identified on the EutC and EutE enzymes from the Eut MCP, and the aldehyde dehydrogenases from a predicted glycyl radical-based MCP from *Clostridium beijerinckii* and from a MCP of unknown function from *H. ochraceum* [29,37–39]. Bioinformatics studies as well as analysis of a library of potential signal sequences have all shown that there are many different signal sequences possible, but successful encapsulation seems to be mediated by a common alpha-helical, hydrophobic motif [39–41]. In fact, signal sequences from the Eut system also function to encapsulate heterologous proteins within Pdu MCPs [39].

Recent studies on the assembly and organization of the beta carboxysome reveal a somewhat different mode of cargo encapsulation. In the case of the beta carboxysome of *Synechococcus elongatus*, auxiliary proteins CcmN and CcmM form a network of protein-protein interactions connecting all of the encapsulated enzymes and the shell [42–46] (Figure 1.3b). In this way, the core of cargo enzymes forms a dense, paracrystalline lattice [47,48]. Further, the enzyme core forms first and is then surrounded by the shell [42,49]. Although these complex networks exist in the native carboxysome, heterologous expression of the requisite shell proteins from a plasmid allows the formation of morphologically normal carboxysome structures that are empty of cargo [50]. In addition, fusions with the C-terminal domain of CcmN may be used to load heterologous cargo similarly to the signal sequences described for Pdu and Eut MCPs above [50].

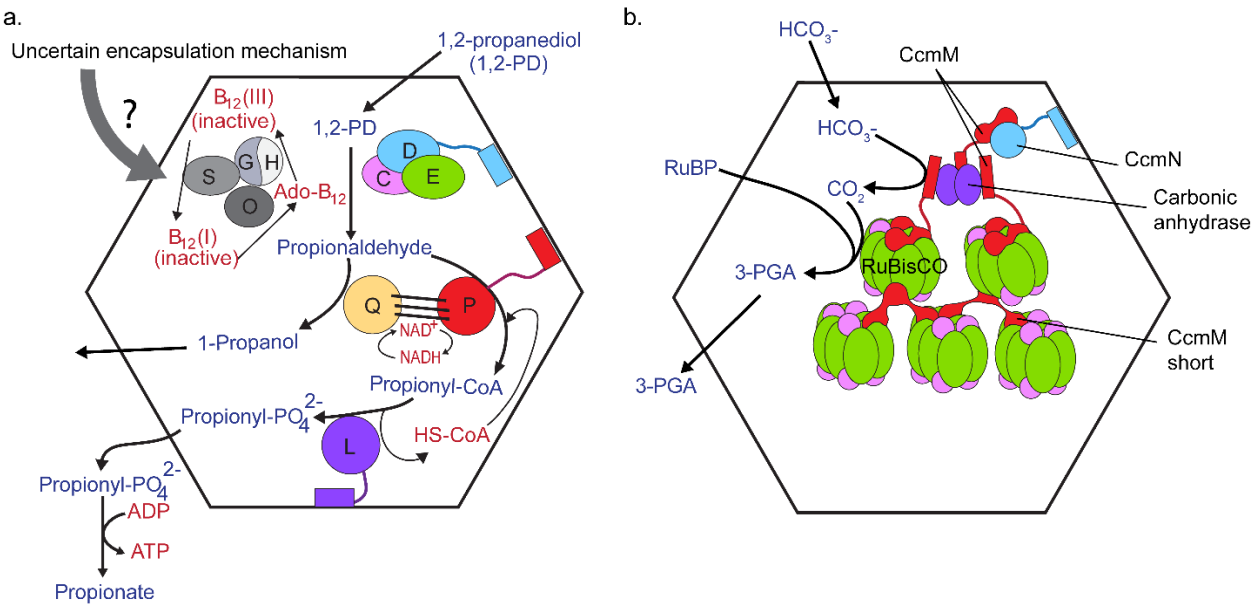


Figure 1.3 Native encapsulation methods. a) Pdu pathway b) β-carboxysome pathway [20,51]

1.4 Tools for MCP engineering

It is clear that the MCP is a complex part of cellular machinery. Engineering tools are necessary to apply the benefits of an MCP scaffold to non-native systems. To optimize a heterologous pathway within the MCP, it is desirable to gain control over several parameters: regulation of MCP production, localization of enzyme cargo, protein content and morphology of the shell, and diffusion of small molecules through the shell. There is significant progress towards these goals in recent work.

Producing MCPs on demand is the first step to enable both the application of MCPs and further study of their function. For Pdu and Eut MCP systems in *S. enterica*, MCPs are formed upon the addition of inducing molecules1,2-propanediol or ethanolamine, respectively. These molecules function in combination with allosteric transcriptional regulators PocR and EutR to activate transcription of the respective operons [52–59]. Normally the production of PocR is under the control of Crp and Arc global regulatory systems, which are affected by environmental conditions [57]. Overexpression of the transcriptional regulator PocR from a plasmid is sufficient to produce Pdu MCPs without 1,2-propanediol. Interestingly, the addition of 1,2-propanediol with PocR over-expressed results in increased loading of reporter proteins [60,61]. In recent work by Jakobson et al, the relative timing of expression of cargo and shell is shown to control amounts of total cargo loaded [60][61]. This indicates that the loading of heterologous cargo is not fixed, nor controlled by a sophisticated mechanism, but is simply proportional to the amount of that cargo present while a given MCP unit is assembling. Researchers may control the expression of the MCP shell to take advantage of this effect.

The discovery and development of signal sequences described above confers control over the MCP enzyme cargo. However, when heterologously over-expressed proteins are directed to the MCP, excess protein may still localize to the cytosol or form polar bodies. There are cases where this is undesirable. For example, an enzyme may produce a very toxic intermediate. The ssrA degradation tag may be used to eliminate cytosolic fluorescence of a Green Fluorescent

Protein (GFP) reporter, while the GFP localized to the MCP is protected [62]. When fused to the c-terminus of a protein, the ssrA tag directs the protein to be degraded by the native ClpXP machinery in the cytosol [63–65]. ClpXP is unable to access proteins within the MCP complex. The ssrA tag and other similar degradation tags are another tool to ensure specificity in protein localization within the cell.

Control over the regulation of MCP shell expression can also be achieved by heterologous expression of the MCP shell genes from a plasmid. This method can expand the utility of MCP systems to other host organisms. There are many examples, detailed below, in which well-formed MCP shells are expressed episomally. However, construction of MCPs omitting or over-expressing some shell proteins often results in mis-formed structures such as nanotubes, fibrils, sheets or rolls [24,42,66–68]. Developing an understanding of the factors influencing MCP assembly will aid in identifying the ratios of building blocks needed to assemble correctly into a functional structure. Alternatively, these other structures may also be studied to uncover new potential applications.

A number of MCP shells from different systems have been expressed from plasmids in *E. coli.* A plasmid expressing Pdu A, -B, -J, -K, -N, -T, and -U shell proteins from *Citrobacter freundii* produced morphologically normal Pdu MCP shells [66]. Likewise, the expression of Eut S, -M, -N, -L, and -K shell proteins from a plasmid produced assembled Eut MCPs [37]. Furthermore, carboxysomes were expressed in *E. coli* by placing the α-carboxysome operon from *H. neapolitanus* on a plasmid [69]. Interestingly, a method of identifying shell protein subtypes and constructing functional heterologous shell protein expression systems was demonstrated by the expression of seven shell proteins from a previously uncharacterized MCP in *H. ochraceum* [29]. The full set of shell proteins from a native operon is not always necessary to produce MCPs shells. For instance, EutS is able to form an MCP-like structure on its own. Heterologous expression platforms can simplify the MCP system and eliminate potential complications or competition caused by the presence of natively encapsulated enzymes. Future work is needed to identify robust ways to express well-formed MCPs.

The crystal structures of several shell proteins are available, including PduA, PduT, PduU, EutL, EutS, GrpN, and EtuB [70–75]. These crystal structures enable the current model of MCP assembly and small molecule transport, guiding genetic and biochemical studies. The central pores of MCP shell proteins are implicated in the transport of small molecule substrates and products across the shell.[26, 27] Mutations to the pore sequence of a shell protein have some effects connected to metabolite diffusion. For instance, Pdu MCP shell protein PduA is abundant in the shell and is suggested to be involved in the selective transport of 1,2-PD [27]. Specifically, occlusion of its pore by $PduA^{S40GSG}$, $PduA^{S40L}$, $PduA^{S40C}$, and $PduA^{S40Q}$ mutations lead to growth defects that are rescued by high 1,2-PD concentrations, while the $PduA^{S40A}$ mutation increased leakage of propionaldehyde to the growth medium [27]. Similarly, the pores of the carboxysome are hypothesized to control flux of bicarbonate and CO_2 across the shell [76]. All characterized carboxysome pore structures are positively charged at the narrowest point, which could enhance the flux of negatively charged bicarbonate relative to the escape of CO_2. Since different pore structures are postulated to lead to altered diffusivity of metabolites, engineering the pores could be the key to tailoring MCP properties for encapsulating non-native systems.

An alternative strategy for engineering MCP properties for non-native pathways is to use novel combinations of native shell proteins. Shell proteins from different MCP systems have the potential to interact to form chimeric assemblies if expressed together at appropriate levels. Cai

et al. demonstrated that an alpha-carboxysome shell protein complements the deletion of a beta-carboxysome shell protein [77]. Sturms et al. observed that co-expression of the Eut and Pdu operons in *S. enterica* from a synthetic promoter system resulted in co-purification of Eut and Pdu components, albeit with a disruption of Pdu MCP function [78]. Such observations are consistent with another finding: among many shell protein homologs, there are essential asparagine, arginine and lysine residues at the edge of each tile that are important for assembly, indicating a common interaction mechanism (For PduA these residues are K26, N29, and R79) [31,75]. In addition, shell assembly does not always depend on cargo loading. The protein shell can form when no cargo is expressed, and foreign proteins may also be localized to the interior of the shell [34,37,62,66]. Thus, it may be possible to bring shell proteins from different systems together to form chimeric MCP assemblies with new properties.

1.5 Impacts and Outlook

Harnessing the MCP as subcellular organization for metabolic engineering is a promising new approach to optimize microbial hosts for biosynthesis of industrially relevant compounds. Pathways engineered to function in bacteria often have unacceptably low yields due to several possible barriers. Engineered enzyme cascades may suffer from low substrate affinity or Km, toxic intermediates, undesirable side reactions, or imbalanced cofactors. All of these problems may be addressed through encapsulation to permit enzymatic reactions that might otherwise fail. This work represents several steps forward in control over MCP assembly, diffusion of small molecules through the shell, and encapsulation of custom cargos within MCP systems.

We first probe the sensitivity of MCP assembly by evaluating the stability of the MCP structure and identifying shell protein mutations that alter or disrupt the structure of Pdu MCPs. We determine that shell protein substitutions allow alteration of MCP shell content, and complement the deletion of shell protein PduA with the Eut shell protein EutM. We find that this substitution results in altered function of the encapsulated metabolism by affecting diffusion through the shell. Next we demonstrate that single amino acid mutations to the pore of the PduA shell protein are sufficient to engineer the performance of the encapsulated pathway, highlighting a new method to control MCP shell properties.

Finally, we report progress and outstanding challenges in the encapsulation of multi-enzyme cascades. We improve the flexibility of the signal sequence tagging by studying the violacein biosynthesis pathway enzymes as a model. We also find that expression level and cargo loading may be modulated by altering the sequence or configuration of the encapsulation peptide fusion. There are few examples of heterologous enzyme encapsulation, and only one in which two enzymes of an ethanol biosynthesis pathway are successfully localized to Pdu MCPs in *E. coli.* While these experiments are important proofs of concept, the best applications of MCP technology await discovery. With the tools developed thus far to enable MCP engineering, researchers now have the capability of applying MCPs to many more pathways. The road will be made easier by increased knowledge and simplification of design rules as well as identification of pathway targets most amenable to encapsulation.

Chapter 2: INVESTIGATION OF SHELL FORMATION AND THE ROLE OF SHELL PROTEINS IN NATIVE PATHWAY BEHAVIOR

A portion of the following is adapted from two previously published sources: Kim, Slininger, and Tullman-Ercek; Protein Science, 2014; and Slininger, Jakobson, and Tullman-Ercek; biorXiv, 2017 with permission.

2.1 Background

In the Tullman-Ercek lab, we sought to develop design rules for the proper assembly of Pdu MCPs and find ways in which the protein shell may be modified. We hypothesized that altering structure or protein content of the shell would have an effect on the function of an encapsulated pathway. To start with, we evaluated the stability of the MCPs *in vitro* to when exposed to a range of temperature and pH conditions. This measurement of the stability of protein interactions forming the shell provides a baseline to which mutant MCPs may be compared. Next, in seeking ways to modify the protein shell to alter binding interactions between shell and cargo, we stumbled upon several novel mutations resulting in aberrant shell structures. This analysis further defines regions of the MCP shell protein structure essential for function. Finally, we undertook and an in-depth comparison of shell proteins from Eut and Pdu MCP systems to discover whether small changes in the protein structure would result in different shell function and an altered performance of an encapsulated pathway.

2.2 Materials and methods

2.2.1 Strains and culture conditions

The bacterial strains used in this study are variants of *Salmonella enterica* serovar Typhimurium LT2 and *Escherichia coli* DH10B (Table 2.1).

For MCP purifications or microscopy experiments, single colonies were grown in 5 mL LB (Lysogeny broth)-Lennox medium at 30°C, 225 rpm, for 24 hours. Cell cultures were diluted from the primary culture 1:1000 into NCE supplemented with 42 mM sodium succinate, 1mM magnesium sulfate, and 50 μM ferric citrate. For cultures containing the plasmid expressing PduD$^{1\text{-}20}$-GFP-ssrA, half the usual amount of antibiotic was used (17 mg/mL chloramphenicol). Culture volumes were 400 mL of media in 2L flasks for MCP purification, and 5 mL of media in 24-well blocks (Analytical Sales and Services, Inc., cat. no. 24108) for microscopy. Cells were grown at 37˚C in an orbital shaker at 225 rpm. For cultures with PduD$^{1\text{-}20}$-GFP-ssrA expressed from a pBAD33 plasmid, 1.33 mM arabinose was added when subculture reached an OD_{600} of 0.4. After five additional hours of growth, samples were taken for fluorescence microscopy or Pdu MCP purification.

The plasmid expressing eight shell genes: *pduA, -B, -J, -K, -M, -N, -T,* and *–U,* was regulated by a pTET promoter with a psc101 origin of replication and a kanamycin resistance cassette. The plasmid was constructed by using biobrick assembly using *S. enterica* genomic DNA as a template. The regions upstream of each gene by at least 20 bp were preserved to maintain native ribosome binding affinities. Unless otherwise stated, this plasmid and its variants were induced with 10 ng/mL aTc. The truncation of *pduJ* was made by adding a stop codon and NotI restriction site in the middle of *pduJ* by round the horn PCR. The PduJ* mutation was made

in a similar fashion, but with a BsaI restriction site. The sequences of each plasmid were confirmed by Sanger sequencing. Plasmids are listed in Table 2.2.

For mutations to *pduA*, the gene was amplified from the *S. enterica* chromosome and inserted into a pBAD33, p15a, cm^R plasmid by golden gate assembly with BsaI restriction endonuclease [79]. This plasmid was transformed into *E. coli*. and was mutated as described above using PCR site-directed mutagenesis. EutM K24A was created in the same manner. Primers for each mutation are shown in Table 2.3.

Chromosomal modifications were made by recombineering as described by Court [80]. First, the native *pduA* gene was replaced with the *cat/sacb* cassette amplified from the TUC01 genome with primers mfsp105 and either mfsp106 or mfsp195 for *eutM* or *pduA* insertions, respectively (primer sequences listed in Table 2.3). This amplicon was used to create Δ*pduA::cat/sacb*. The e*utM* gene was amplified with primers mfsp199 and mfsp200 and used to make Δ*pduA::eutM*. To make mutations in *pduA,* the gene was first subcloned into a plasmid vector and mutated by site directed mutagenesis as described above. The inserts were then amplified from plasmids by PCR with primers mfsp197 and mfsp104 and used to make all chromosomal *pduA* mutants. For FLAG tag fusions, the non-polar Δ*pduA::cat/sacb* strain constructed with mfsp106 was used as a starting point to maintain the same RBS region for the subsequent *pduB* gene. The FLAG sequence was added by PCR to *pduA* and *eutM* sequences using primers mfsp304 and mfsp218, respectively before insertion by recombineering. Each strain was confirmed by Sanger sequencing of PCR amplicons from the PduA region of the *S. enterica* chromosome.

For growth of *S. enterica*, single colonies were picked from freshly streaked plates and were grown in 5 mL of TB (Terrific Broth) medium at 30°C, 225 rpm, for 24 hours. For growth curves with a 1,2-PD carbon source, cultures were diluted to an OD_{600} of 0.05 in 50 mL No-Carbon-E (NCE) minimal medium supplemented with (unless stated otherwise) 55 mM 1,2-PD, 150 nM coenzyme B_{12}, 1 mM magnesium sulfate, and 50 μM ferric citrate. Cultures were grown in 250 mL flasks at 37°C in an orbital shaker at 225 rpm. At every time point, 1.5 mL of cell culture sample was taken for OD_{600} measurement and HPLC analysis. The doubling time of each growth curve was calculated by plotting the log of the OD_{600} vs. time. The linear region of each individual curve was identified (including a minimum of three time points). The slope of the linear regression between time points was determined, and the doubling time was calculated with $t_D = \log(2)\,/slope$.

Strain	Organism	Genotype
MFSS044	*S. enterica* serovar Typhimurium	LT2 wild type
MFSS263	*S. enterica* serovar Typhimurium	LT2 $pduA^{K26A}$
MFSS236	*S. enterica* serovar Typhimurium	LT2 Δ*pduA::cat/sacb*
MFSS252	*S. enterica* serovar Typhimurium	LT2 Δ*pduA::eutM*
MFSS273	*S. enterica* serovar Typhimurium	LT2 Δ*pduA::eutM-FLAG*
MFSS392	*S. enterica* serovar Typhimurium	LT2 Δ*pduA:: pduA-FLAG*
MFSS337	*S. enterica* serovar Typhimurium	LT2 Δ*pduA::eutM* ^{K24A}
MFSS256	*S. enterica* serovar Typhimurium	LT2 $PduA^t$
MFSS275	*S. enterica* serovar Typhimurium	LT2 $PduJ^t$
TUC01 [80]	*E. coli*	DH10B *cat/sacB*

Table 2.1 Bacterial strains used in this study

Plasmid	Description
mfspl205	*pduD*$^{1\text{-}20}$*-gfp-ssrA*, pBAD, p15A, chloramphenicol^{R}
Psim6 [80]	λ-red, TS (30ºC), carbenicillin^{R}
mfspl090	*pduABJKMNTU*, pTET, pSC101, kanamycin^{R}
mfspl126	*pduAB J* ^{t}*KMNTU*, pTET, pSC101, kanamycin^{R}
mfspl131	*pduAB J*KMNTU*, pTET, pSC101, kanamycin^{R}

Table 2.2 Plasmids used in this study

Primer	Description	Sequence
mfsp105	*cat/sacB* ΔPduA forward	ttatagtcccaactatcggaacactccatgcgaggtcttttgtgacggaagatcacttcg
mfsp106	*cat/sacB* ΔPduA reverse (keep PduB RBS)	ttcccttcggtaagatttttttctacatcggtgtgagggcgatcaaagggaaaactgtccatat
mfsp195	complete ΔPduA sacb reverse (no PduB RBS)	caatcacctgcgccatgatctgttccaccagctcattgctgcatcaaagggaaaactgtccatat
mfsp197	PduA knock in reverse	caatcacctgcgccatgatctgttccaccagctcattgctgctcattggctaattccctcggta
mfsp104	PduA knock in forward	ttatagtcccaactatcggaacactccatgcgaggtctttatgcaacaagaagcactaggaatgg
mfsp199	knock in EutM forward	ttatagtcccaactatcggaacactccatgcgaggtctttatggaagcattaggaatgattgaaacc
mfsp200	knock in EutM reverse	ttcccttcggtaagatttttttctacatcggtgtgagggcgtcaaatgttgctgtcgccttg
mfsp218	PduA::EutM-FLAG knock in reverse	ttcccttcggtaagatttttttctacatcggtgtgagggcgtcacttgtcatcgtcatctttataatcaatgttgctgtcgcctttg
mfsp304	PduA::PduA-FLAG knockin reverse	ttcccttcggtaagatttttttctacatcggtgtgagggcgtcacttgtcatcgtcatctttataatcttggctaattcccttcggta
mfsp301	EutM K24A forward	cgatgcgatggtaGCGgccgcgcgcgtgaagctg
mfsp302	EutM K24A reverse	gcgcgcggcCGCtaccatcgcatcggaggcctcaatcagc
mfsp210	PduJ^{t} knockin reverse	tgctccttattcgccatcgattaggctgatttcggtatttattacggacgcgggataacg
mfsp204	PduJ^{t} knockin forward	ctttcgggatctccatgcttaatcacaggagaacggcagtatgaataacgcactgggact
mfsp103	PduA^{t} knockin reverse	ttcccttcggtaagatttttttctacatcggtgtgagggcgttattatgggatgacgtgtacggc
mfsp114	PduJ* reverse	ttatggtctcagtacttaggctgatttcggtaacgcggcctccgcatcgctcgccggacgcgggataacg
mfsp091	PduJ* forward	tattggtctcacatgaattcgagctcaactatcggaacactc
mfsp053	PduJ^{t} forward	tattctcgagatttacacagcgatgttgaggcc

mfsp054	PduJ^t reverse	tattctcgagttattacggacgcgggataacg

Table 2.3 Primers used in this study

2.2.2 Fluorescence microscopy

Bacteria were viewed using a Nikon Ni-U upright microscope with a 100x 1.45 n.a. plan apochromat objective. Images were captured using an Andor Clara-Lite digital camera and Nikon NIS Elements software. Fluorescence images were collected using a C-FL Endow GFP HYQ bandpass filter. All images were taken with 400 ms exposure and adjusted identically for contrast in Adobe Photoshop software.

2.2.3 MCP purification

5 mL cultures were grown in LB-Miller from a single colony for 24 hours at 30°C, then subcultured 1:1000 into 400 mL of NCE with 55 mM 1,2-PD and grown for 13-15 hours at 37°C. To express heterologous cargo, at OD600~0.4 the appropriate inducer was added at the concentrations indicated and the cultures were grown for a further 5.5 hours. MCPs were purified by sedimentation as previously described [81], with the following modification: in the place of BPER-II bacterial lysis solution (Thermo), a solution of 1% w/v octylthioglucoside (Santa Cruz Biotech) in 20 mM Tris (Fisher) pH 7.5 in water was used for cell lysis. They were stored at 4°C until analysis by SDS-PAGE or TEM. Total protein concentration of pure MCP sample was determined by Pierce BCA protein assay kit (Thermo Fisher).

2.2.4 SDS-PAGE and western blot

Denaturing protein electrophorese (SDS-PAGE) was performed on purified MCPs using a 12.5% polyacrylamide gel unless otherwise noted. A 130V potential was applied for 70 minutes. Unless otherwise noted, equal volumes (15 µL) of each MCP sample were loaded onto the gel, and densitometry quantifications were normalized by sample protein concentration. Gels were either stained with Coomassie dye or further processed for western blot. Western blotting was done with a PVDF membrane according to standard protocols. Blotting was done with a mouse anti-FLAG primary antibody (Sigma) 1:2000 dilution in 50 mM Tris 150 mM NaCl pH 7.6 with 0.05% Tween-20 (TBST) with 1% w/v dry milk. The secondary antibody was HRP-conjugated goat anti-mouse antibody (Thermo) diluted 1:1000 in TBST. Labeling was visualized with west-pico chemiluminescent substrate (Thermo) using a Bio-Rad ChemiDoc XRS+.

Densitometry analysis of Coomassie stained gels was performed using at least three different MCP purifications for each strain. Each purified sample was run on three different gels. Images were analyzed using the Image Lab software (Bio-Rad). The absolute signal measurement for each band was normalized by the signal for the band containing proteins PduT and PduE (band 7 in Figure 2.17). This band had the least variation in signal over all replicates.

2.2.5 Transmission electron microscopy

Samples were placed on 400 mesh formvar coated copper grids with a carbon film after the grids were treated by glow discharge. 10 μL of purified MCPs at a concentration of about 0.1 mg/mL was placed on each grid for 2 min. The grids were washed three times with deionized water before fixation. Then, 10 μL of 2% glutaraldehyde in water was placed on each grid for 1 min. Grids were then washed an additional three times with deionized water. Finally, samples

were stained in 1.6% aqueous uranyl acetate for 1 min. All solutions were centrifuged directly before use to avoid aggregate or particulate contact with grids. Samples were imaged with an REI Tecnai T12 transmission electron microscope and a Gatan Ultrascan 1000 camera (Gatan, Pleasanton, CA).

2.2.6 Immunostaining

Samples were placed on 400 mesh formvar coated copper grids with a carbon film after glow discharge treatment. 10 μL of purified MCPs at a concentration of about 0.1 mg/mL was placed on grids for 2 min. The grids were washed three times with deionized water. Blocking buffer was prepared with 25 mL of phosphate buffered salt (PBS) at pH 7.4, 0.2 g bovine serum albumin (BSA), and 25 μL of gelatin from cold water fish skin (Sigma). Then, grids were incubated with blocking buffer for 15 min. Grids were incubated with mouse anti-FLAG primary antibody (Sigma) diluted 1:1000 in the blocking buffer for 2 hours. Grids were washed four times with PBS for 2 min each. Grids were then incubated with anti-mouse IgG –gold secondary antibody produced in goat (Sigma) diluted 1:20 in blocking buffer for 1 hour. Grids were washed again four times with PBS for 2 min each time. Samples were fixed with 2% glutaraldehyde in PBS for 5 min. Samples were washed for 2 min in PBS, then four times in deionized water. Grids were stained with 1% uranyl acetate in water for 3 min before grid was allowed to dry.

2.2.7 Mass spectroscopy

Protein bands were excised manually and transferred to Eppendorf tubes (0.2 mL). Protein-containing gel pieces were washed with a series of washes that included alternating additions of 100 µL 0.1 M ammonium bicarbonate followed by 100 µL of 100% ACN in order to remove traces of coomassie staining. Once gel slices were thoroughly destained, they were dried using a speed-vac. Gel pieces were then subjected to 10 mM dithiothreitol for 20 minutes at 56 °C and subsequently alkylated with 20 mM iodoacetamide for 1 h at room temperature in the dark. Acetonitrile was then added to shrink the gel pieces and the pieces were thoroughly dried in the speed-vac prior to addition of 0.1 ml of a solution containing 12.5 ng/μl trypsin in 50 mM ammonium bicarbonate. Gel pieces were allowed to swell at 4C for 3 hours prior to overnight digestion at 37 °C. Liquid was collected following digestion and the gel pieces were bathed in 100% acetonitrile to adequately shrink the gel pieces. All liquid was collected into one tube and evaporated to dryness using a speed-vac. The peptide solution was reconstituted into 10 μl of 2% ACN with 0.1% TFA and assessed for peptide concentration based on the absorbance at 280 nm using a Nanodrop spectrophotometer. One microgram total peptide was loaded onto the nano-ACQUITY UPLC™ chromatographic system. Peptides were loaded and separated on a C18 Trizaic Naotile using a 60 min RP gradient at 450 nL/min (3–40% ACN over 40 min). The column temperature was set at 45 °C. Lock mass (Leucine enkephalin (556.2771 Da), 250 fmol/mL) was constantly infused by the NanoAcquity auxiliary pump at a constant flow rate of 1 μL/min with lockspray scans set at intervals of 45 seconds. The Xevo QTof ™ mass spectrometer (Waters) was programmed to switch between low (6 eV) and high (18–42 eV) energies in the collision cell, with a scan time of 1 s per function over a mass range of 50–2000 Da. LC-MS^E data were processed with ProteinLynx GlobalServer v2.3 (Waters) and searched in the associated *S. enterica* protein database (UniProtKB/SwissProt Protein Knowledge Base).

*2.2.7 Temperat*ure *incubation of purified MCPs*

50 µl of purified MCPs, at a concentration of 300 µg/ml, were incubated at the corresponding temperature for 15 minutes, then stored at 4°C for up to 2 h until samples were processed for transmission electron microscopy (TEM).

2.2.7 pH conditions

100 µl of purified MCPs, at a concentration of 300 µg/ml, were centrifuged at 4°C for 30 minutes at 21,000 g to pellet the MCPs. The supernatant was removed and the MCPs were resuspended in 100 µl of the appropriate buffer. The buffers used to adjust pH are listed Table 2.4. Samples were stored in buffer at 4°C for up to 2 h until samples were processed for TEM.

pH	Solute concentration
3	79.45 mM Citric acid, 41.1 mM sodium phosphate
4	61.45 mM Citric acid, 77.1 mM sodium phosphate
5	48.50 mM Citric acid, 103.0 mM sodium phosphate
6	36.85 mM Citric acid, 126.3 mM sodium phosphate
7	17.65 mM Citric acid, 164.7 mM sodium phosphate
8	50 mM Tris
9	50 mM Tris
10	100 mM CAPS
11	100 mM CAPS

Table 2.4: pH conditions. The listed buffers were used to achieve solutions of various pH for MCP sample suspensions. All buffers were supplemented with 150 mM NaCl.

2.2.8 Dynamic light scattering

100 µl of purified MCPs, at a concentration of 750 µg/ml, were placed in 8.5 mm disposable spectrophotometry cuvettes and analyzed on a Melvern Instruments Zen 3600. For temperature studies, the instrument was programmed to increase the temperature in increments of 5°C, equilibrate for five minutes, and record.

2.2.9 Thermal shift assay

25 µl of purified MCPs, at a concentration of 300 µg/ml, were mixed with SYPRO Orange at a final concentration of 15X (SYPRO Orange is proprietary dye and the absolute concentration is not disclosed). The temperature was increased from 30°C to 75 °C by 0.2°C increments held for 12 seconds, and fluorescence was measured in a Bio-Rad CFX96 Touch Real-Time PCR Detection System.

2.2.10 Size distribution analysis

TEM images were imported into ImageJ and the ellipse tool was used to manually draw an ellipse circumscribing each MCP. The measure function was used to determine the Feret's diameter for at least 100 MCPs per condition tested, obtained from random fields of view. The diameters were reported as a histogram and also a bar graph with the error bars representing one standard deviation from the mean Feret's diameter for MCPs at that condition.

2.3 Results

2.3.1 Stability of the wild-type Pdu MCP

The following is adapted from Kim, Slininger, and Tullman-Ercek; Protein Science, 2014 with permission. Edward Kim characterized MCPs enclosing GFP reporter, while I characterized native MCPs. The portion of work by Edward Kim is reported in his dissertation; some parts are included here for comparison.

To assess the stability of the wild-type Pdu MCP *in vitro* MCPs were purified from *S. enterica* by sedimentation. Cultures were grown in minimal media with 1,2-PD added to induce MCP formation. The purified protein samples were exposed to high temperatures, and buffers varying in pH.

The heat stability of MCP was characterized in two ways. The first method used was a thermal shift assay to measure the melting point of the MCP structure. The protein sample was incubated with the fluorescent dye SYPRO Orange. The fluorescent signal of this die is quenched in water. As MCPs disassemble, hydrophobic regions of the shell protein monomers are exposed to solvent. This allows the SYPRO dye to associate with the protein, decreasing the quenching effect of the water, and producing a fluorescent signal. This signal increases further when individual shell proteins denature at a much higher temperature. In the thermal shift assay, fluorescence of SYPRO Orange is measured as the temperature is increased at a rate of 1°C per minute. With this assay, we find that the melting point, or disassembly temperature, of native Pdu MCPs is 53°C (Figure 2.1a). Another student in the Tullman-Ercek lab, Edward Kim, found that Pdu MCPs encapsulating GFP with a fused localization signal (PduP$^{1-18}$-GFP) had a similar melting temperature of 52°C (Figure 2.1b).

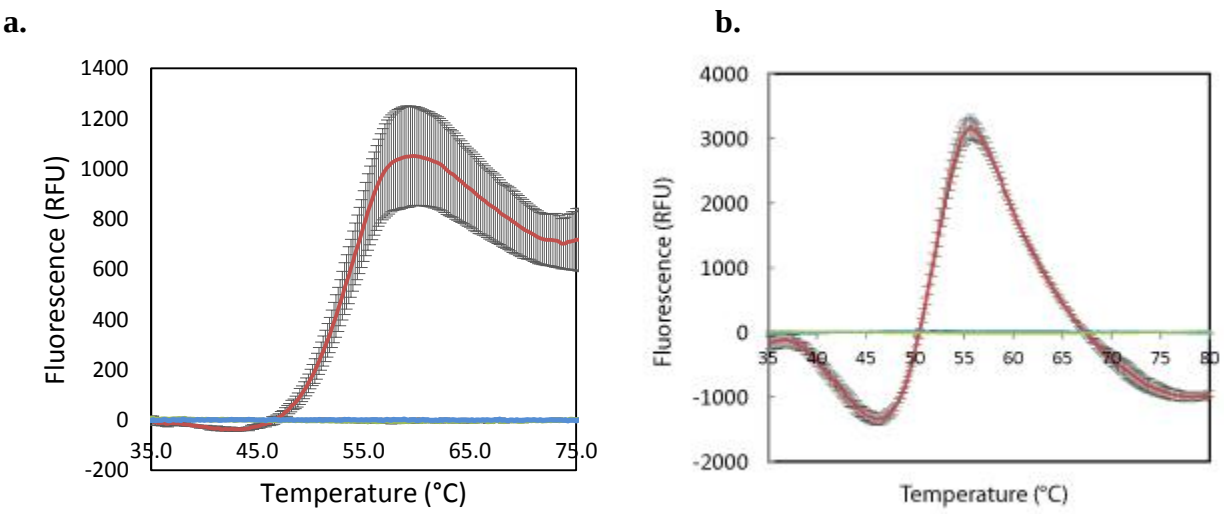


Figure 2.1 Thermal shift assay. Purified Pdu MCPs are incubated with SYPRO Orange as the temperature is increase by 1°C per minute (Red). Purified MCPs without SYPRO Orange (blue), and buffer (50 mM Tris, pH 8.0) with SYPRO Orange (green) are also shown. The average fluorescence from three replicates is shown. Error bars represent one standard deviation.

Another way to observe thermostability is by analyzing pure MCP samples using transmission electron microscopy (TEM) and dynamic light scattering (DLS) to detect how the morphology of the MCP structure changes with temperature. MCP samples were brought to an

elevated temperature for 15 min, and then cooled to 4°C for storage until TEM or DLS testing. By TEM, the MCP structures in both the native and the strain containing $PduP^{1-18}$-GFP appear to maintain normal MCP morphology up to 60°C, although clustering and aggregation is observed at 60°C and above (Figure 2.2 and Figure 2.3). This is in agreement with the observed thermal shift melting temperatures. At 70°C, the majority of MCPs are found within large aggregates. Interestingly, in some aggregates, the outlines of MCP shells are still visible. The size of each particle was derived from the TEM images using measurements of Feret's diameter. No differences in size of particle were observed across different temperatures where individual particles were measurable (Figure 2.4).

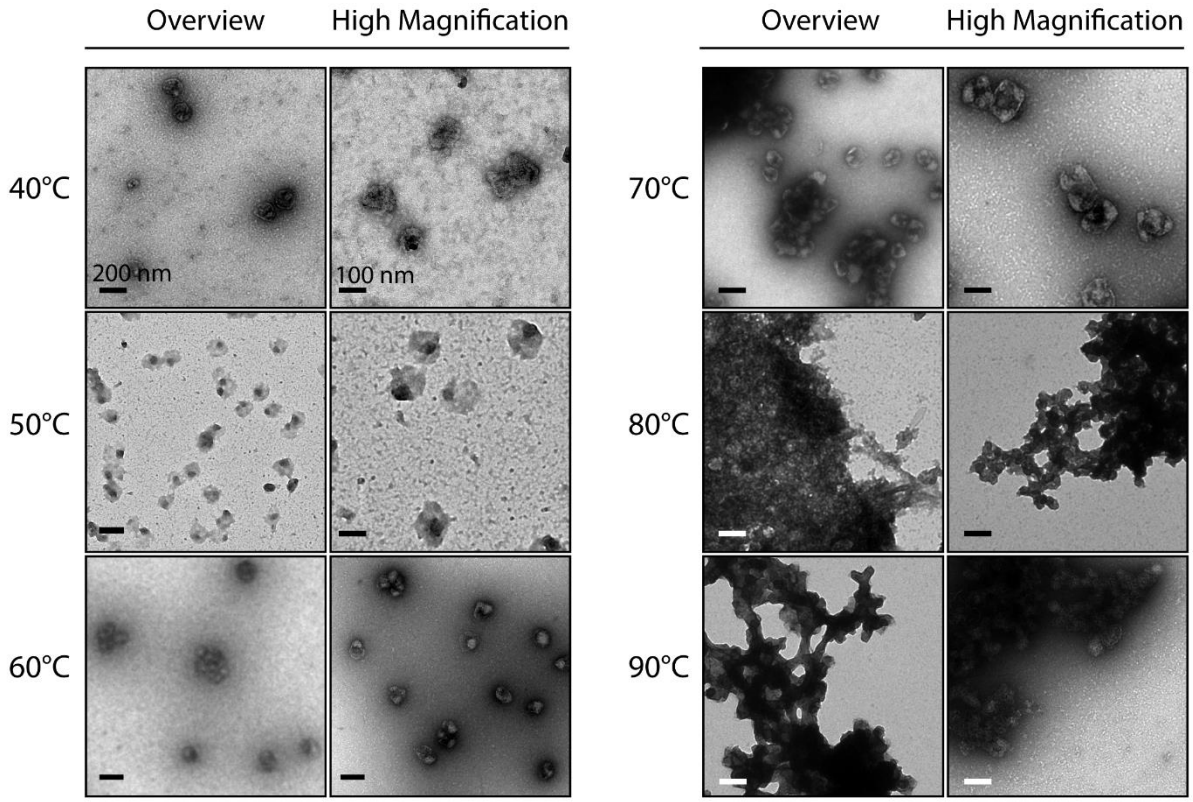


Figure 2.2 TEM images of heat-treated purified native MCPs. MCPs were incubated at elevated temperatures for 15 min then cooled to 4°C. Two representative images are shown for each temperature—one overview image (scale bar represents 200 nm, left image), and one higher magnification image (scale bar represents 100 nm, right image) to highlight the details of the MCPs.

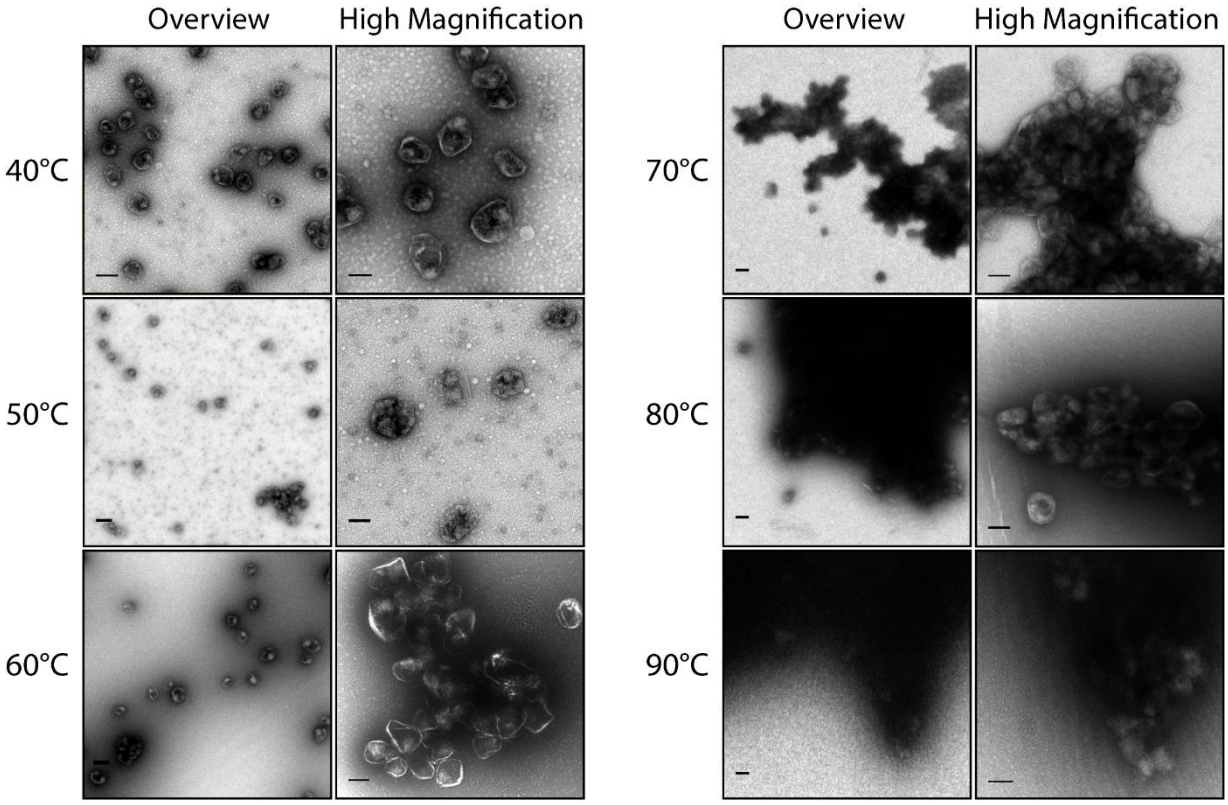


Figure 2.3 TEM images of heat-treated purified MCPs encapsulating PduP$^{1–18}$-GFP. MCPs were incubated at elevated temperatures for 15 min then cooled to 4°C. Two representative images are shown for each temperature—one overview image (scale bar represents 200 nm, left image), and one higher magnification image (scale bar represents 100 nm, right image) to highlight the details of the MCPs.

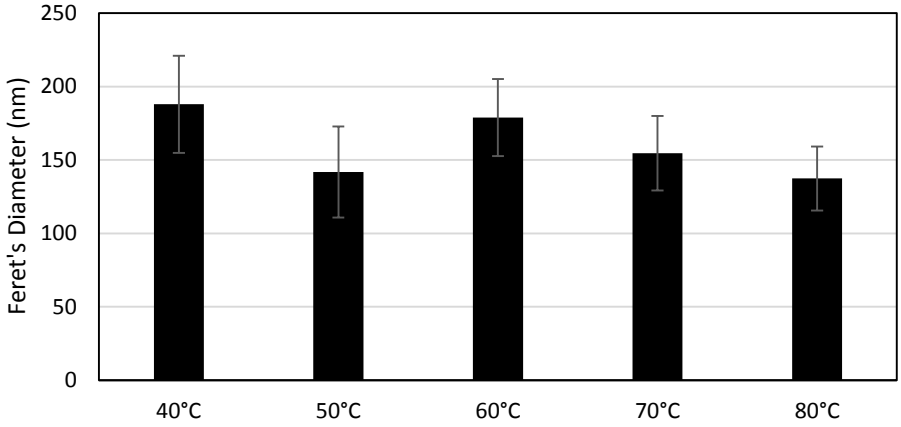


Figure 2.4 Size distribution analysis of TEM images of purified Pdu MCPs encapsulating PduP$^{1-18}$-GFP incubated at elevated temperatures. Size is represented by the Feret's diameter of an ellipse circumscribing each non-aggregated MCP particle, calculated using the ImageJ

software. Data is averaged over at least 100 measurements of MCPs from random frames of view. Error bars represent one standard deviation.

DLS further complemented the size measurements of Pdu MCPs *in vitro*. DLS measurements of untreated Pdu MCP encapsulating PduP$^{1-18}$-GFP showed a distribution of sizes that agree with TEM measurements (Figure 2.5). When these samples are exposed to increasing temperature, the distribution broadens at 50°C and becomes two peaks, with one peak representing large aggregate species (Figure 2.6). This is again in agreement with observations from TEM. At temperatures above 65°C the sample becomes too polydisperse for accurate DLS analysis.

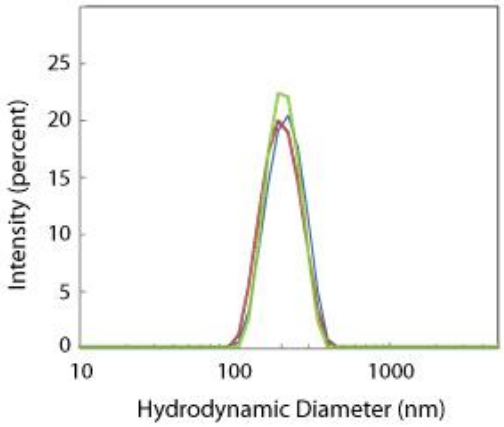


Figure 2.5 Size distribution profile of purified Pdu MCPs encapsulating PduP$^{1-18}$-GFP, using dynamic light scattering. The three colors represent data from three replicates.

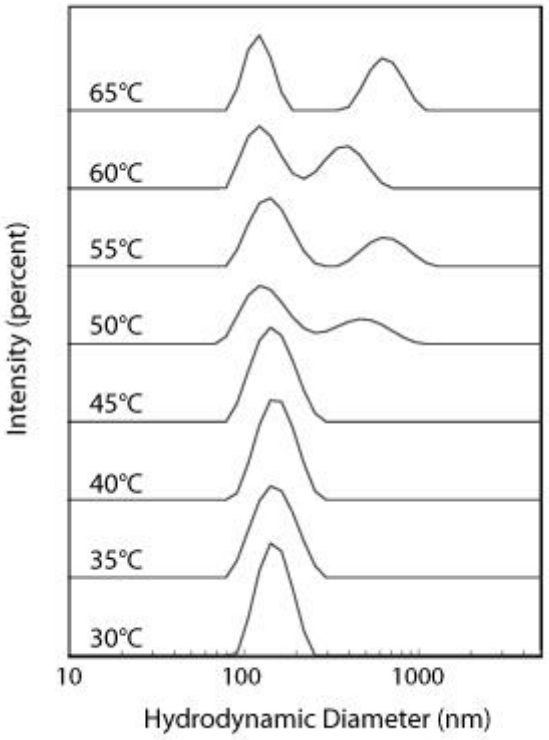


Figure 2.6 Hydrodynamic diameter of MCPs as measured by dynamic light scattering. Purified Pdu MCPs encapsulating PduP$^{1-18}$-GFP are incubated in a Malvern Instruments Zen 3600 which was programmed to increase temperature in increments of 5°C, equilibrate for 5 minutes, and record.

Another test of the stability of MCPs is to observe how they behave at different pH levels. After incubation in the buffers listed in Table 2.4, MCP morphology was observed by TEM. Samples retain normal MCP appearance between pH 6 and pH 10, although aggregates start to form at pH 9 and above or pH 4 and below (Figure 2.7). At pH 5, 10, and 11, MCPs appear smaller. Size analysis shows that the MCPs at pH 5, 10, and 11 average about 100 nm in diameter, while native MCPs at neutral pH are between 150 and 200 nm in diameter (Figure 2.8). MCPs could not be measured below pH 5 due to extensive aggregation.

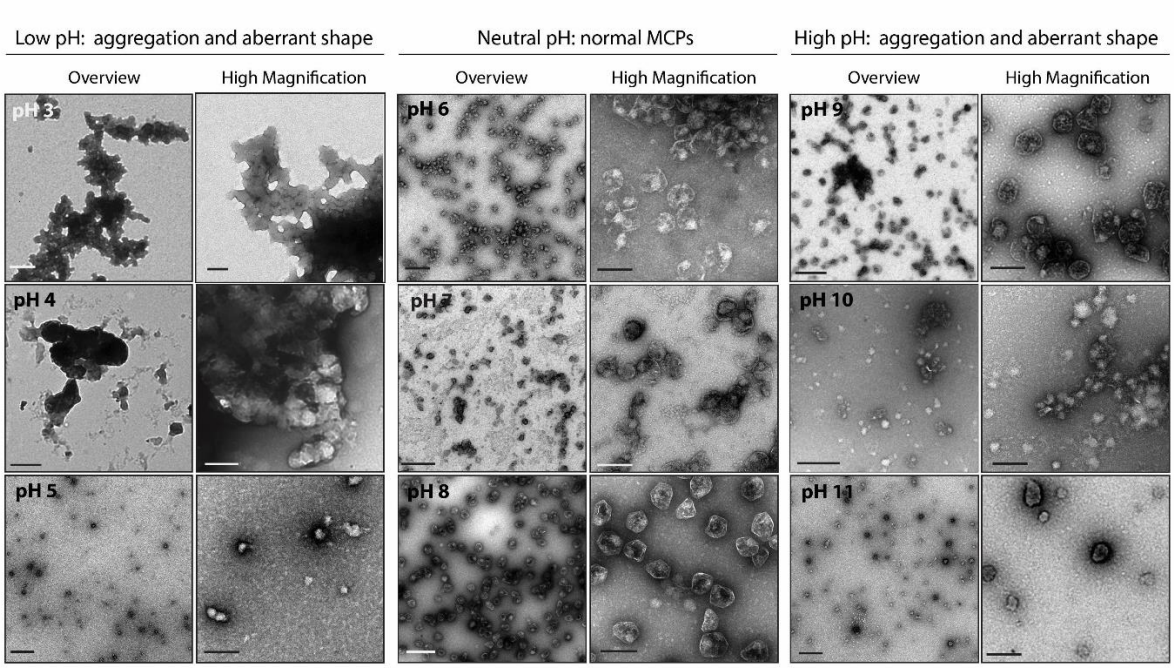


Figure 2.7 TEM images of purified native MCPs in buffers of varying pH. Two representative images are shown for each pH—one overview image (scale bar represents 500 nm, left image), and one higher magnification image (scale bar represents 200 nm, right image) to highlight the details of the MCPs.

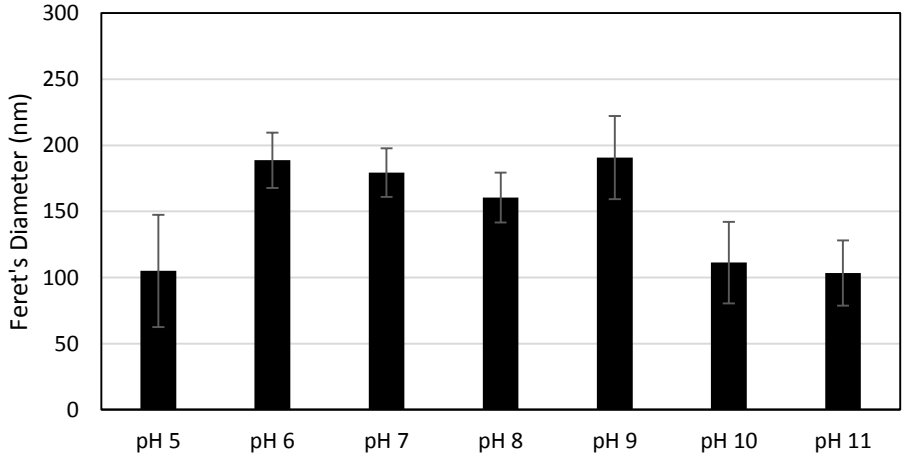


Figure 2.8 Size distribution analysis of TEM images of purified native Pdu MCPs over changing pH conditions. Size is represented by the Feret's diameter of an ellipse circumscribing each MCP particle, calculated using the ImageJ software. Data is averaged over at least 100 measurements of MCPs from random frames of view. Error bars represent one standard deviation.

2.3.2 C-terminal truncations of PduA or PduJ destabilize MCP assembly

The shell protein PduA and its close homolog PduJ are implicated in interactions between the MCP shell and its cargo enzymes. Pull down assays indicated that both PduJ and PduA

interact with the PduP cargo enzyme [82]. PduP did not bind to purified PduA if the C-terminus was truncated. Further, the C-terminus of PduA was modeled as a binding site for the PduP signal sequence [82]. An alanine scan of the PduA C-terminus revealed 3 key residue mutations, H81A, V84A, and L88A, that reduced PduP encapsulation (Figure 2.9). We began a study of the PduJ C-terminus to determine whether it would have analogous interactions with the Pdu signal sequences. Instead, we were surprised to find that alterations to this region of PduJ caused aberrant MCP assembly.

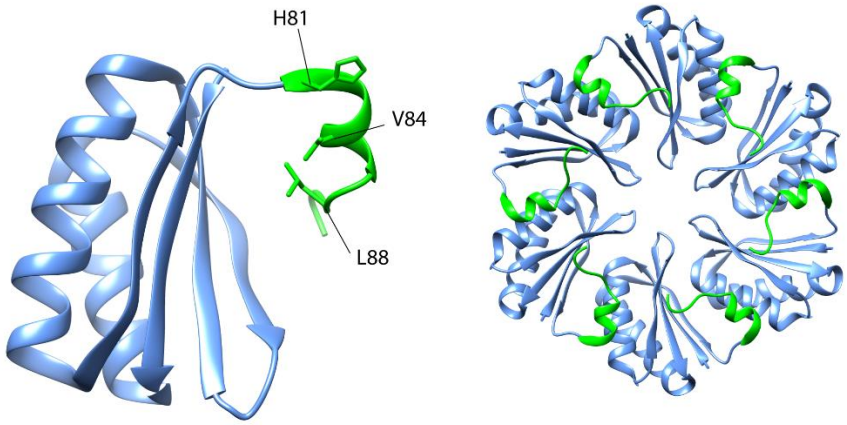


Figure 2.9 Structure of PduA (PDBID 3NGK). The crystal structure of PduA is shown with the c-terminal residues highlighted in green. The image at left is the monomer, and image at reight is the hexamer assembly. The residues labeled H81, V84, and L88 are those implicated in the binding interaction with PduP.

A plasmid expressing all Pdu MCP shell constituents, PduA, -B, -J, -K, -M, -N, -T, and –U, was modified to encode PduJ mutants. These constructs were used to observe the assembly of PduJ mutants with other Pdu MCP components. We truncated the last twelve amino acids from PduJ to create PduJ^{t} by insertion of a stop codon and a restriction site between the sequence encoding PduJ residues 1-78 and the C-terminal PduJ residues 79-90. In this way, the native RBS region upstream of PduK is preserved to avoid disruption of expression levels. The shell-encoding genes on one plasmid were co-expressed with a PduD$^{1-20}$-GFP-ssrA reporter on a second plasmid in *E. coli* DH10B and strains were analyzed by microscopy. The strain with the complete shell plasmid forms puncta similar to the native MCPs produced in *S. enterica*. Conversely, the strain with the shell plasmid encoding PduJ^{t} formed long chains of cells connected by long filaments preventing cell division (Figure 2.10a, b, c). The GFP was localized to most of the filaments, highlighting their structure within the cell. TEM of sectioned cells showed clearly a long ordered structure spanning linked cells and preventing complete membrane separation (Figure 2.11). We were intrigued that although the C-terminus of PduA was amenable to mutation, the truncation of the PduJ C-terminus resulted in such a large effect. The observed filaments may be structures similar to those previously characterized in other mutant shell systems [24,66,68].

To learn whether this phenotype could be replicated in the native expression system, a chromosomal PduJ truncation was made in *S. enterica*. This strain also had elongated and linked cells with GFP localized to both filament structures and large aggregates (Figure 2.10d). Next,

we sought to eliminate the binding site of protein cargos without truncating PduJ. A shell plasmid encoding a mutated PduJ gene called PduJ* was constructed in which three residues (H79, V82, and L86) of the PduJ C-terminal sequence were changed to alanine. Using the SPIDER^2 secondary structure prediction server, the C-terminus of PduJ* is still predicted to form an alpha helix, similarly to the wild-type PduJ structure [83]. Surprisingly, the strain expressing PduJ* with all other Pdu shell proteins also formed filaments (Figure 2.10e). Finally, a chromosomal PduA truncation (PduA^t) was made removing residues 81-84. Removing the C-terminus of PduA also caused a chain-forming structure (Figure 2.10f). We conclude that individual amino acids in the C-terminal helix of PduA could be altered without major MCP structural changes, but eliminating the C-terminal helix of PduA or PduJ or making changes to alter the binding site for PduP and PduD proteins in PduJ results consistently in the observed filament phenotype.

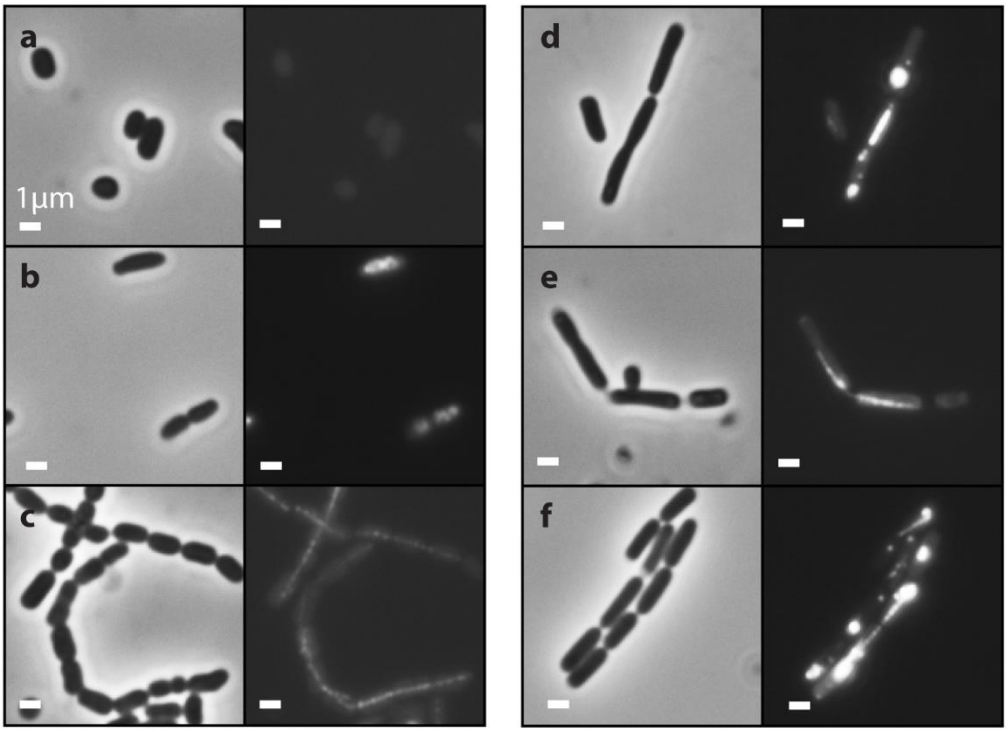


Figure 2.10 Fluorescence microscopy a) *E. coli* containing the Pdu shell plasmid with PduJ^t, uninduced, b) *E. coli* containing the unaltered Pdu shell plasmid, 10 ng/mL aTc, c) *E. coli* containing the Pdu shell plasmid with PduJ^t, 10 ng/mL aTc, d) *S. enterica* with PduJ^t on chromosome, e) *E. coli* with PduJ* mutantion to Pdu shell plasmid, f) *S. enterica* with PduA^t on chromosome. All strains also express the PduD$^{1-20}$-GFP-ssrA reporter plasmid induced with 0.02% arabinose.

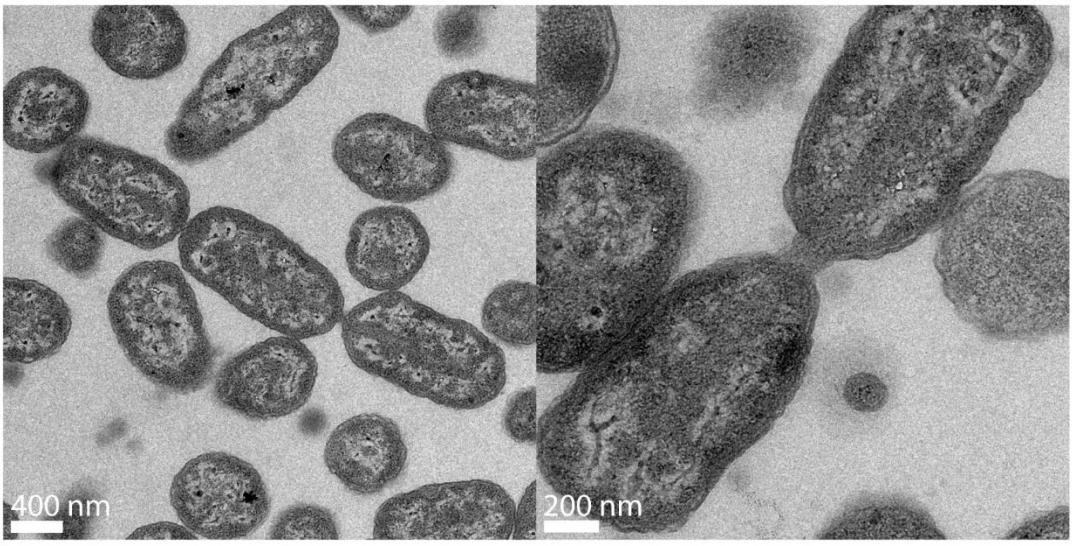


Figure 2.11 TEM of thin cell sections. *S. enterica* expressing PduJ^t mutation on Pdu shell plasmid induced with 10 ng/mL aTc.

We next tested how the truncating PduA affected the behavior of the encapsulated 1,2-PD metabolism. *S. enterica* harboring the PduA^t mutation at PduA locus on the chromosome was grown in minimal medium with 1,2-PD as the sole carbon source and added coenzyme B_{12}. Coenzyme B_{12} is a required cofactor for the first reaction converting 1,2-PD into propionaldehyde by the PduCDE diol dehydratase. This cofactor is supplied in the growth medium because *S. enterica* cannot produce it under our aerobic conditions. Two different levels of coenzyme B_{12} were tested: 150 nM (saturating), and 20 nM (limiting). In the saturating B_{12} condition the *PduA^t* mutant grows poorly compared to the wild-type (Figure 2.12). This could indicate increase intermediate escape or toxicity due to lack of a closed MCP shell. This result is consistent with findings for other non-assembling MCP shell mutants [24,27,31,84].

Growth at low coenzyme B_{12} concentrations has been used previously as a method to sensitively detect whether mutations to MCP proteins result in non-assembly or defective MCP shells [24,27,31,84]. The limiting coenzyme B_{12} condition is advantageous for non-assembling MCP mutants, such as the *pduA^{K26A}* mutant, and allows improved growth. Conversely, the wild type and mutant strains that assemble well grow poorly at 20 nM coenzyme B_{12} compared to mutants that do not assemble MCPs. We exploited this growth condition as an additional assay to detect if the mutations we made to PduA disrupted MCP assembly. We found that the *PduA^t* mutant grew better than the wild type with low B_{12} concentrations, confirming that the observed filament structures do not impose a complete diffusion barrier (Figure 2.12).

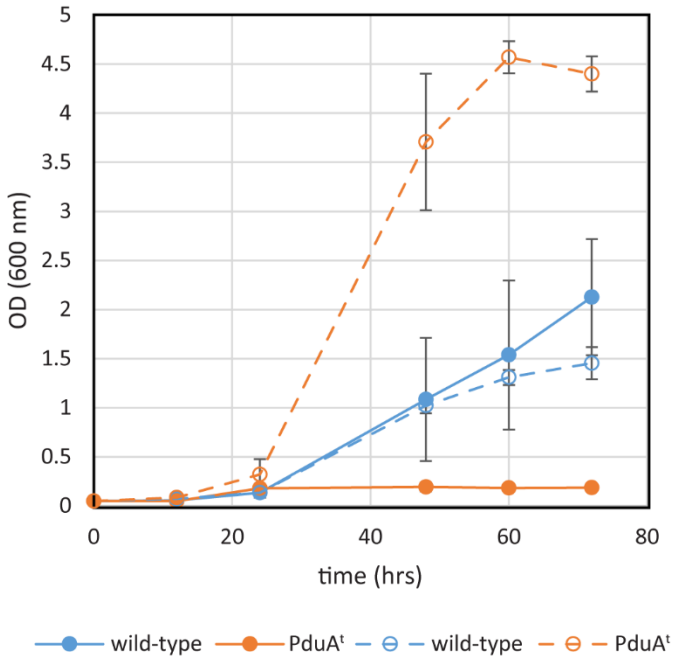


Figure 2.12 Growth of PduA^t strain. Optical density measurements were taken over time for *S. enterica* wild-type and PduA^t strains grown in minimal media supplied with 1,2-PD as the sole carbon source and coenzyme B12. Solid lines with filled circles are measurements for cultures growth with saturating B12 concentrations (150 nM). Dashed lines with open circles are cultures grown with limiting B12 concentration (20 nM)

2.3.3 Comparison of PduA and EutM shell proteins

The following is adapted from Slininger, Jakobson, and Tullman-Ercek; biorXiv, 2017 with permission.

PduA forms one of three canonical shell protein types: a hexameric tile complex containing a central pore approximately 6 Å in diameter [26]. The residues defining the shape and properties of the pore include those contained in a loop at the hexamer symmetry axis with sequence Lys-Ile-Gly-Ser-Gly-Leu. The importance of the K37 and S40 loci of this sequence was previously demonstrated; certain mutations to either of these residues result in diminished growth of *S. enterica* LT2 on 1,2-PD as a sole carbon source [27,31]. EutM, a close PduA homolog from the Eut operon, also forms a hexamer with nearly identical predicted structure (RMSD EutM/PduA = 1.004 Å, PDB ID 3MPY (EutM) and 3NGK (PduA)) and a highly similar amino acid sequence (67% identity)[23,26]. Refer to the protein sequence alignment for PduA and EutM (Figure 2.13a). Key differences between PduA and EutM are found in the pore, particularly at the residues aligning with PduA K37 and S40. The EutM pore loop has the sequence Gln-Ile-Gly-Gly-Gly-Leu (Figure 2.13b). In Figure 2.13b, the pore-lining lysine residue in PduA and the glutamine in EutM are highlighted in green.

a

```
PduA: MQQEALGMVETKGLTAAIEAADAMVKSANVMLVGYEKIGSGLVTVIV
EutM: M--EALGMIETRGLVALIEASDAMVKAARVKLVGVKQIGGGLCTAMV

RGDVGAVKAATDAGAAAARNVGEVKAVHVIPRPHTDVEKILP---KGIS--
 RGDVAACKAATDAGAAAAQRIGELVSVHVIPRPHGDLEEVFPISFKGDSNI
```

b

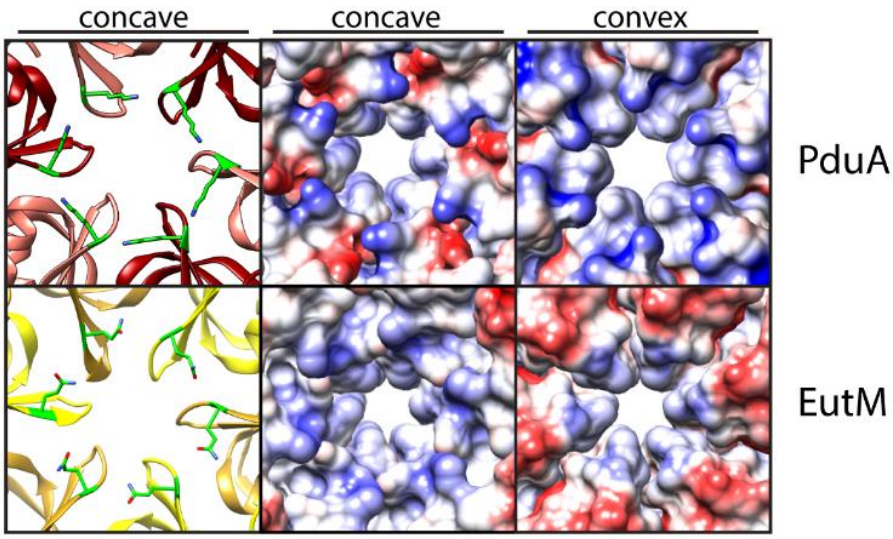
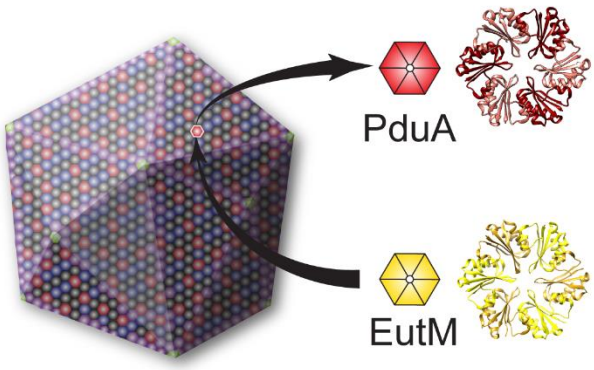


Figure 2.13Comparison of PduA and EutM a) Alignment of PduA and EutM protein sequences. b) Left hand images are ribbon structures of the pores of PduA (red) and EutM (yellow). Residues PduA K37 and the corresponding glutamine of the EutM pore are highlighted in green. To the right are the surface models of the respective pore regions colored by the electrostatic potential. Blue is positive potential, white is neutral, and red is negative potential. (PDB IDs: 3ngk, 3mpy) [23,26]

2.3.4 Pdu shell protein substitution alters 1,2-PD metabolism

The following is adapted from Slininger, Jakobson, and Tullman-Ercek; biorXiv, 2017 with permission. Some figures include data discussed in chapter 3 of this dissertation.

a



b

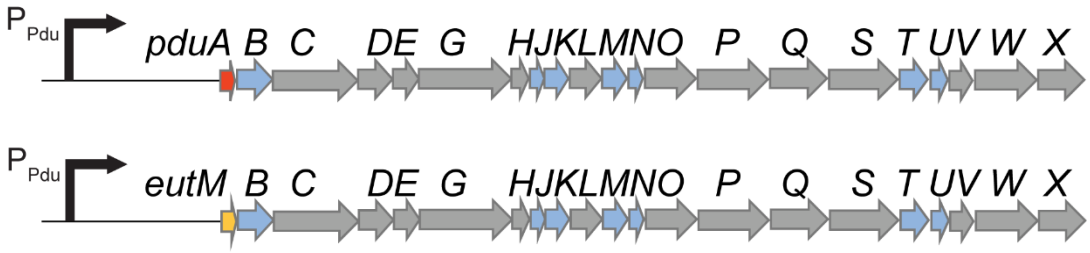


Figure 2.14 Δ*pduA::eutM* substitution a) Schematic demonstrating the hypothesized overall structure of the MCP. Arrows indicate replacement of the PduA shell protein (red) with EutM (yellow). Crystal structures to the right illustrate structural similarity. b) A diagram of the Pdu operon and the operon resulting from recombineering to substitute eutM for *pduA* on the *S. enterica* chromosome. Other Pdu shell proteins are blue arrows, enzymes of the pathway and proteins of unknown function are grey arrows.

We substituted *eutM* for *pduA* to test whether the structural similarity of *eutM* is sufficient to complement the deletion of *pduA* and maintain proper Pdu MCP assembly (Figure 2.14a). We also wanted to learn if small differences between the pore structures would manifest in significant differences in phenotypes attributed to metabolite diffusion. In order to minimize disruption of the native regulation of the Pdu operon, we constructed a genomic integration of *eutM* at the *pduA* locus (Figure 2.14b). To test MCP formation in this Δ*pduA::eutM* strain, the expression of the Pdu operon was induced in minimal NCE media containing 55 mM 1,2-PD as the sole carbon source, and 150 nM coenzyme B_{12}. Growth of the Δ*pduA::eutM* strain in this medium is an indicator of metabolite diffusion through the MCP structure. For instance, a strain harboring a *pduA* deletion without a compensating substitution exhibits decreased cell growth due to increased shell permeability to the toxic intermediate propionaldehyde, despite producing a seemingly closed MCP that may be purified and appears to contain all other necessary shell

proteins and enzymes [24,31]. Furthermore, occlusion or other changes to the architecture of the pores of the MCP shell also lead to growth defects due to restriction of diffusion of the substrate 1,2-PD [27]. Our results show this is not the case for the Δ*pduA::eutM* strain. Not only did the strain harboring the substitution grow without a noticeable defect on 1,2-PD, we observed improved growth over wild type under these growth conditions (Figure 2.15, Table 2.5). To control for changes in cell morphology, we also quantified colony forming units at various growth time points to confirm improved growth of the Δ*pduA::eutM* strain (Figure 2.16). The Δ*pduA::cat/sacb* mutant strain was used as a control for both growth measurements. This strain grew poorly compared to the wild-type (Figure 2.15, Figure 2.16). Moreover, no gross changes to cell morphology or physiology were observed in the Δ*pduA::eutM* strain by phase contrast (Figure 2.17).

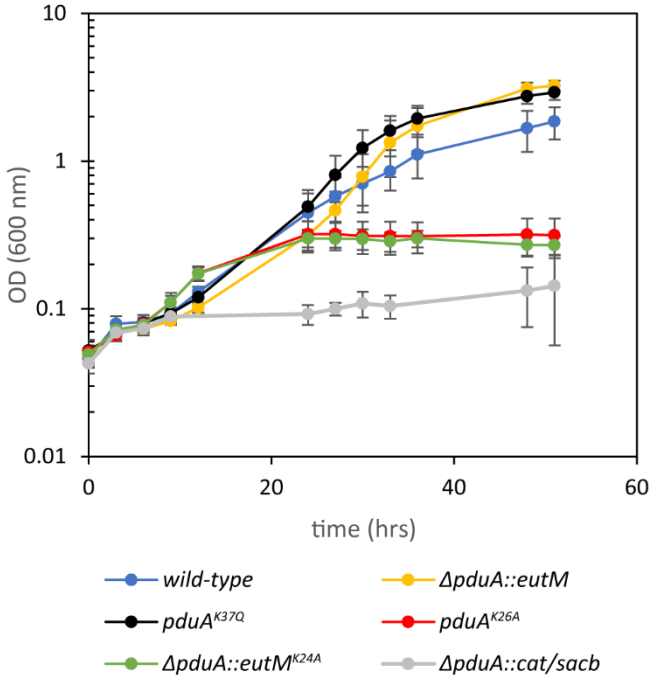


Figure 2.15 Strains harboring shell protein variants grow on 55 mM 1,2-PD and 150 nM coenzyme B_{12}. MCP variant strains were grown on 55 mM 1,2-PD with 150 nM coenzyme B_{12}. Wild-type strain LT2 is compared to LT2 Δ*pduA::eutM* , *pduA*^{K37Q} , *pduA*^{K26A} , Δ*pduA::eutM* ^{K24A}, and Δ*pduA::cat/sacb* strains. Mean optical density at 600 nm is plotted vs time. Error bars represent 95% confidence interval from six measurements started on four different days.

Strain	Doubling time (hr)	Standard error
MFSS044, Wildtype	7.85	±0.89
MFSS252, *ΔpduA::eutM*	5.73	±0.29
MFSS238, *pduA*^{K37Q}	5.75	±0.49

Table 2.5 Doubling times of PduA mutants for growth on 1,2-PD. Both *ΔpduA::eutM* and *pduA*^{K37Q} strains are statistically different than the wildtype strain by two tailed T-test. The *ΔpduA::eutM* and *pduA*^{K37Q} strains are not significantly different from each other by two tailed T-test. n=7 for all cases.

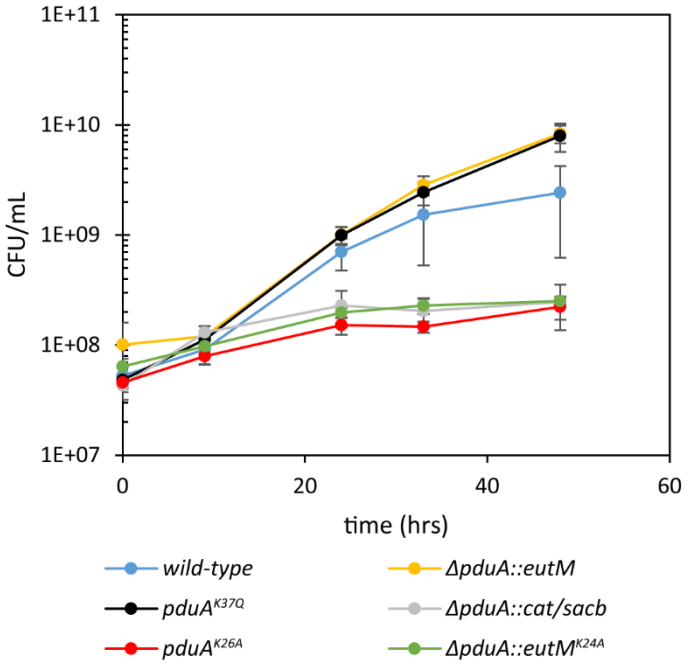


Figure 2.16 Strains harboring shell protein variants grow on 55 mM 1,2-PD and 150 nM coenzyme B_{12} – quantification of cell forming units over time. The same procedure was used as described in Figure 3, except cell forming units (CFU) were quantified by dilution plating and colony counting. Error bars represent 95% confidence interval from ≥ six measurements started on four different days.

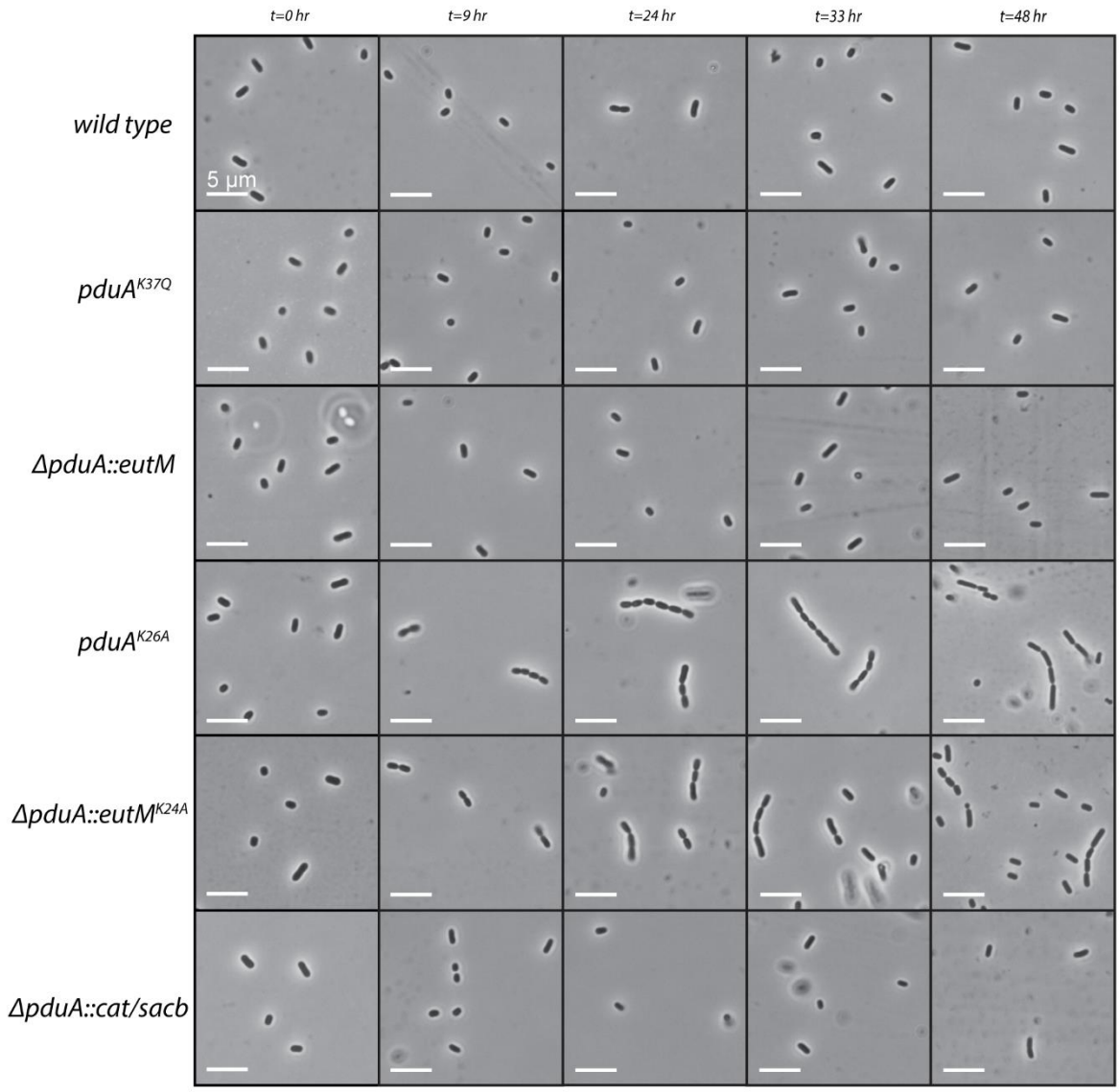


Figure 2.17 Phase contrast microscopy of cell morphology over growth curve with 1,2-PD carbon source. Microscopy images were collected during the growth curves for six strains depicted in Figure 2.15. Representative images are shown for each time point for each strain. All scale bars are 5 µm.

2.3.5 EutM complements the deletion of PduA in S. enterica

The following is adapted from Slininger, Jakobson, and Tullman-Ercek; biorXiv, 2017 with permission. Some figures include data discussed in chapter 3.

We next set out to demonstrate that EutM is incorporated into the Pdu shell, and to determine if the formation of the shell was altered by this incorporation. We found that substituting *eutM* for *pduA* allowed the formation of chimeric MCPs that were morphologically similar to the wildtype, as assessed by a number of assays. First, we targeted GFP for

localization within the MCPs produced in each of the variants and analyzed assembly formation *in vivo* using fluorescence microscopy. GFP was expressed as a fusion to the $PduD^{1-20}$ encapsulation tag and an SsrA degradation tag such that only encapsulated GFP remains protected from proteolysis by the ClpXP machinery [62,82,85]. This results in bright puncta by fluorescence microscopy. Fluorescent puncta were observed in both Δ*pduA::eutM* and wild-type strains, confirming MCP formation and protein encapsulation *in vivo* (Figure 2.18a, e, b, f, and Figure 2.19). Automated image analysis by Cell Profiler yielded a count of the number of puncta per cell and indicated that the number of chimeric MCPs per cell is similar to the number of wild-type MCPs per cell (Figure 2.20)[86]. Variability in puncta brightness is often observed between cells as well as between puncta within a cell. While we have not investigated the cause of this feature, it appears consistent across both the wild-type and Δ*pduA::eutM* strains.

We next purified MCPs by sedimentation to compare the morphology of chimeric MCPs to wild-type *in vitro*. Transmission electron microscopy (TEM) analysis showed that MCPs purified from the Δ*pduA::eutM* strain are morphologically similar to wild-type MCPs (Figure 2.18i, j, and Figure 2.19). An analysis of MCP diameter demonstrated that MCPs purified from the Δ*pduA::eutM* strain have a similar mean diameter to the wild-type but with a somewhat wider distribution of sizes (Figure 2.21).

The protein content of purified MCP samples was analyzed by densitometry measurements of coomassie-stained SDS-PAGE gels. Prominent bands consistently observed in MCP samples were previously identified via mass spectrometry [21]. Those identifications are used here. MCPs purified from the Δ*pduA::eutM* strain exhibited shell protein bands in similar proportions to those found in wild-type MCPs, except that the PduA band was missing and the 13 kDa band expected to be PduU roughly doubled in density (Figures 2.18m, 2.19, 2.22). We hypothesized that this band represented both PduU and EutM superimposed. To verify the localization of EutM, the purification was repeated with a FLAG epitope tag fused to EutM. Western blotting indicated that EutM-FLAG runs at the apparent size of 13 kDa in a 12.5% acrylamide gel, and we confirmed that the 13 kDa band included EutM-FLAG by mass spectrometry (Figures 2.18n, 2.23, Table 2.6).

Next, we constructed a non-polar Δ*pduA::pduA-FLAG* strain. We purified MCPs from this strain along with the Δ*pduA::eutM-FLAG* strain and compared the levels of FLAG-tagged proteins using SDS-PAGE and western blot. Loading of the purified MCP samples was normalized by total protein concentration. We found that all samples contained similar amounts of PduA- or EutM-FLAG (Figure 2.24).

To confirm that EutM is not merely co-purifying with Pdu MCPs, anti-FLAG immuno-gold antibodies were used to label Pdu MCPs purified from the Δ*pduA::eutM-FLAG* strain . Immunostained TEM images show that the anti-FLAG immuno-gold antibodies bind well-formed MCP structures. In contrast, the wild-type MCP sample does not bind any gold particles (Figure 2.18 i, j). Taken together, all evidence supports our hypothesis that EutM properly incorporates into the Pdu MCP shell and causes no morphological defects.

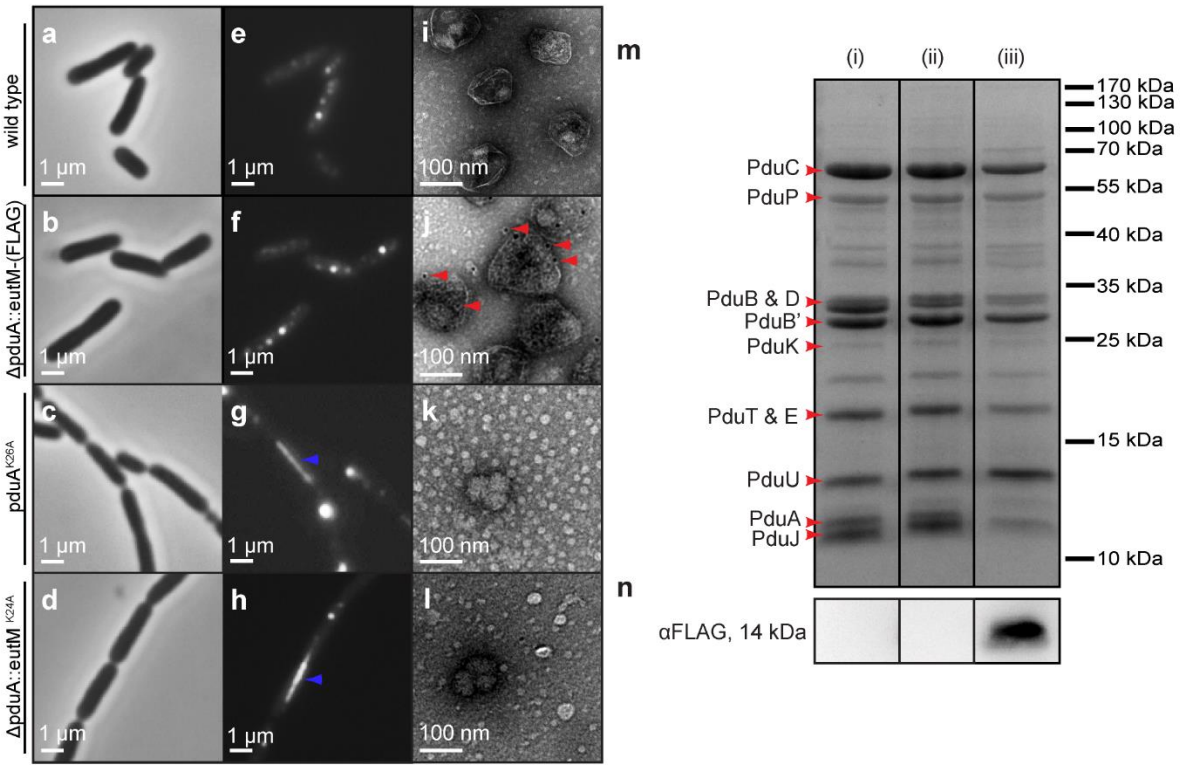


Figure 2.18 EutM incorporates into the Pdu MCP (a, b, c, d) Phase contrast images of *S. enterica* strains LT2 (*wild type*), LT2 Δ*pduA::eutM*, LT2 *pduA*^{K26A}, and LT2 Δ*pduA::eutM* ^{K24A}, respectively. All strains are co-expressing Pdu MCPs and PduD$^{1-20}$-GFP-ssrA. (e, f, g, h) GFP fluorescence microscopy images corresponding to phase contrast images a-d. Large filamentous aggregates are indicated with blue arrows in panels g and h. (i, j) Immunostained transmission electron microscopy of MCPs purified from wild type LT2 and LT2 Δ*pduA::eutM-FLAG*, respectively. Bound IgG-conjugated gold particles are indicated with red arrows. (k, l) Transmission electron microscopy images of MCPs purified from the strains depicted in images c, d. For convenience, larger versions of the images in panels (i-l) are provided in Supplemental Figure 2.25. (m) Coomassie-stained SDS-PAGE of purified MCPs. Lane (i) is a purification from strain LT2 wild type, lane (ii) is from strain LT2 *pduA* ^{K37Q}, lane (iii) is from strain LT2 Δ*pduA::eutM-FLAG*. (n) α-FLAG western blot of purified compartments.

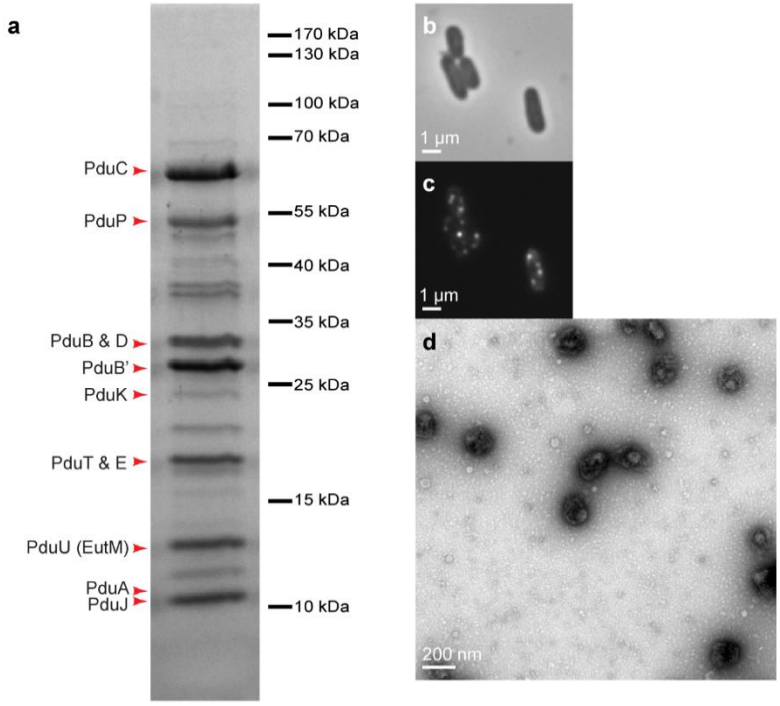


Figure 2.19 Δ*pduA::eutM* strain characterization (a) Coomassie-stained SDS/PAGE of MCPs purified from the *S. enterica LT2* Δ*pduA::eutM* strain. (b) Phase contrast image of Δ*pduA::eutM* strain co-expressing Pdu MCPs and PduD$^{1\text{-}20}$-GFP-ssrA (c) GFP fluorescence microscopy image corresponding to image b. (d) Transmission electron microscopy image of MCPs purified from the Δ*pduA::eutM* strain.

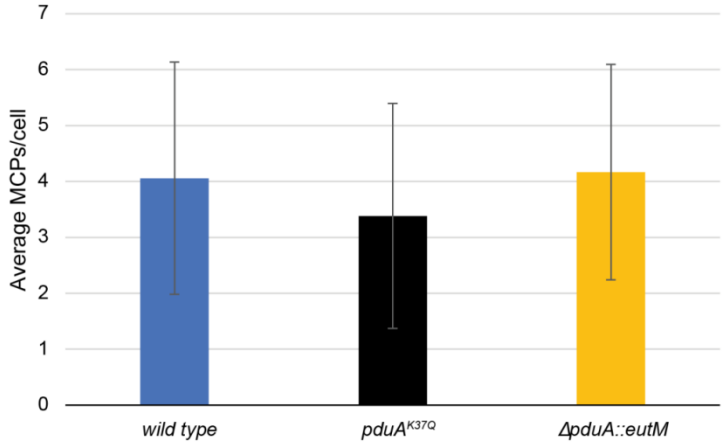


Figure 2.20 Comparison of microcompartment number per cell for three Pdu variants. Cell profiler was used to count the number of fluorescent puncta per cell for LT2 wild type (WT), *pduA*^{K37Q}, and Δ*pduA::eutM* strains coexpressing PduD$^{1\text{-}20}$-GFP-ssrA. Error bars represent one standard deviation for n>5000 cells, generated from six separate cultures, and measured on two different days; representative images are shown in Figure 2.14 and Figure 2.15.

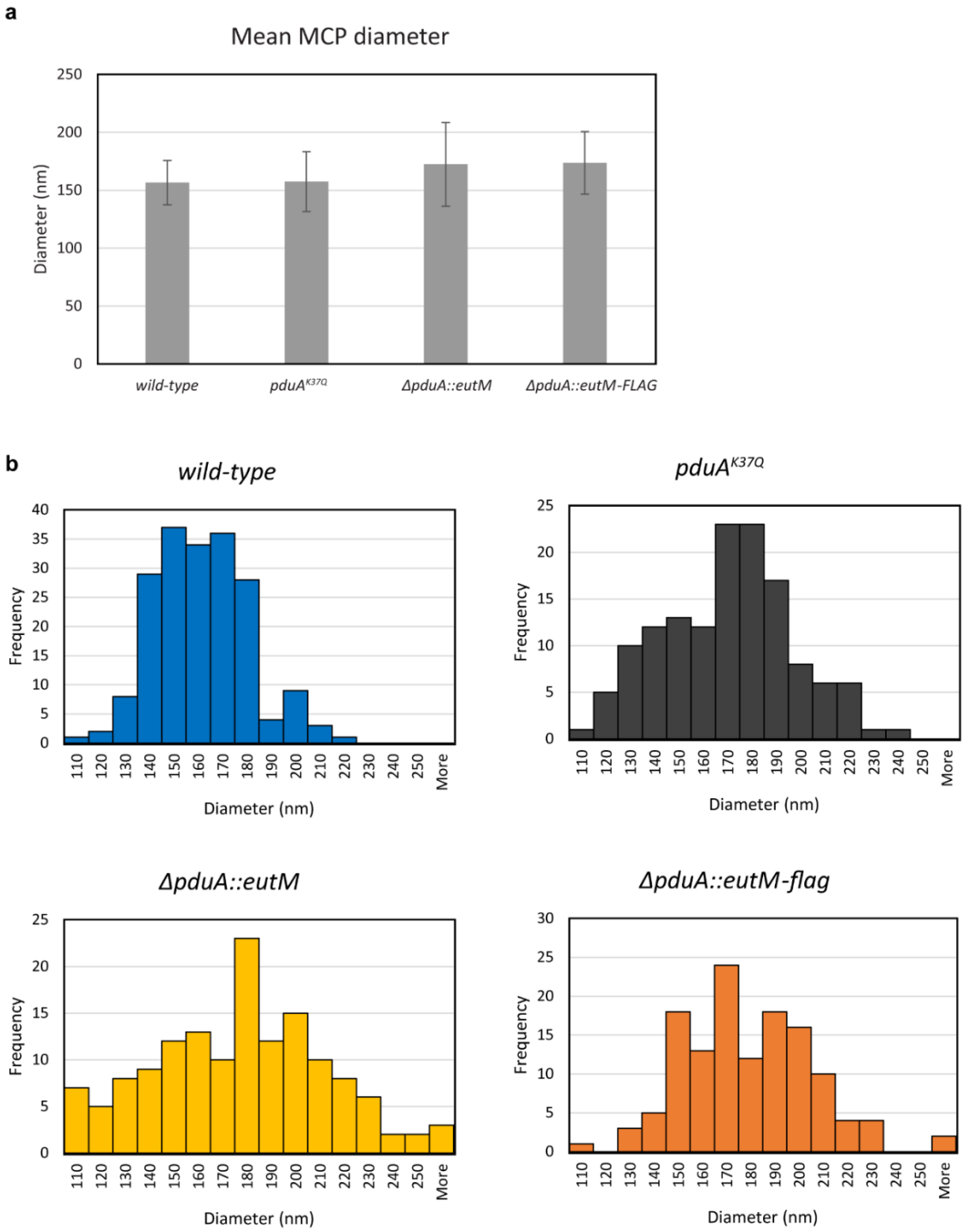


Figure 2.21 Size analysis of MCPs from TEM images. TEM images purified MCPs from each strain were analyzed to determine the mean MCP diameter. In each case, at least 80 individual MCPs were measured from at least 3 separate purifications. (a) the average Feret's diameter as measured using ImageJ software. Error bars represent the standard deviation. (b) histograms for the size distribution in each strain.

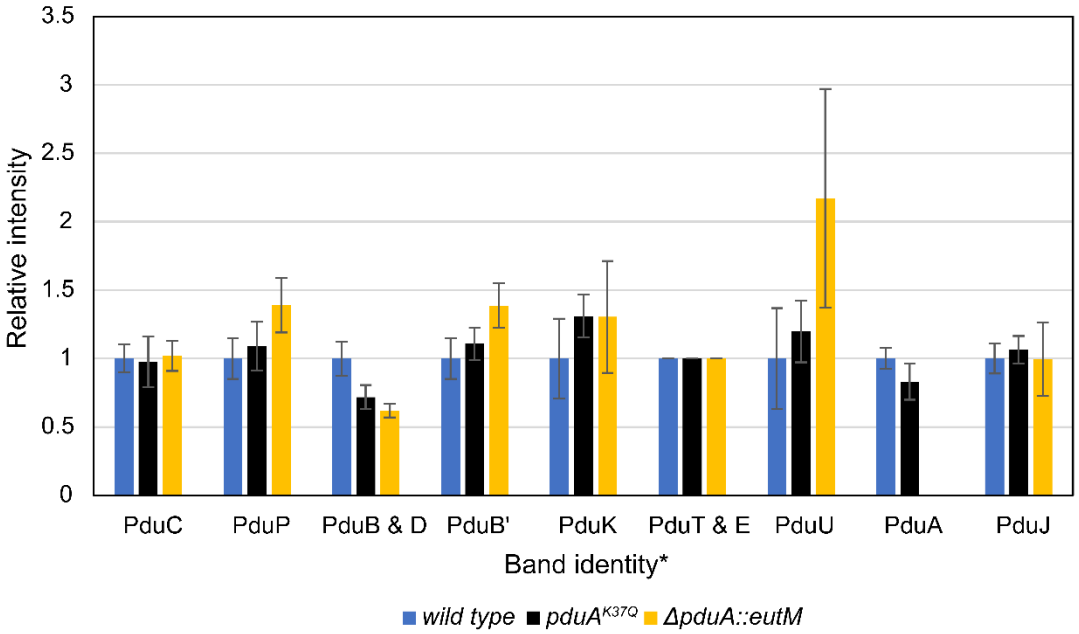


Figure 2.22 Densitometry of purified MCP SDS-PAGE. Purified microcompartments from the strains shown were analyzed by SDS-PAGE. Densitometry was used to estimate the relative ratios of prominent bands; *proteins of the same approximate molecular weight were identified previously by Havemann and Bobik [21]. The raw peak area value for each band was normalized by the value for the band containing PduT and PduE (this band was found to have very consistent density in all lanes) to calculate a band ratio for each band within a lane. A second normalization was applied by dividing all of the band ratios for each strain by the corresponding ratios in the wild type strain, measured from the same gel. In this way, results from different gels may be compared. Error bars represent one standard deviation with $n \geq 3$ purifications, each purification was run on three separate gels to capture error in staining method.

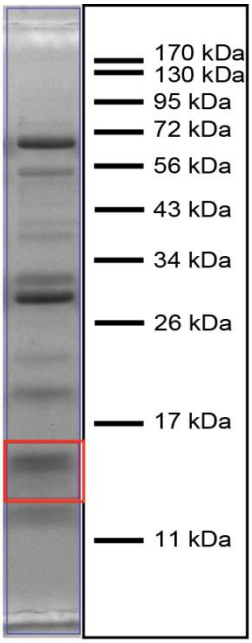


Figure 2.23 SDS-PAGE of purified MCPs for identification of EutM by mass spectrometry. SDS-PAGE of purified MCPs from strain LT2 Δ*pduA::eutM*. The red box outlines the slice of gel used for mass spectrometry identification.

Protein Name		Mass (kDa)	% Seq. Coverage*	Unique Peptide Count	Uniprot Acc.
EutM	Ethanolamine utilization shell protein	10	52	6	P41791

Peptide threshold of 0%

*% coverage of complete EutM sequence

Table 2.6 Mass spectrometry analysis data

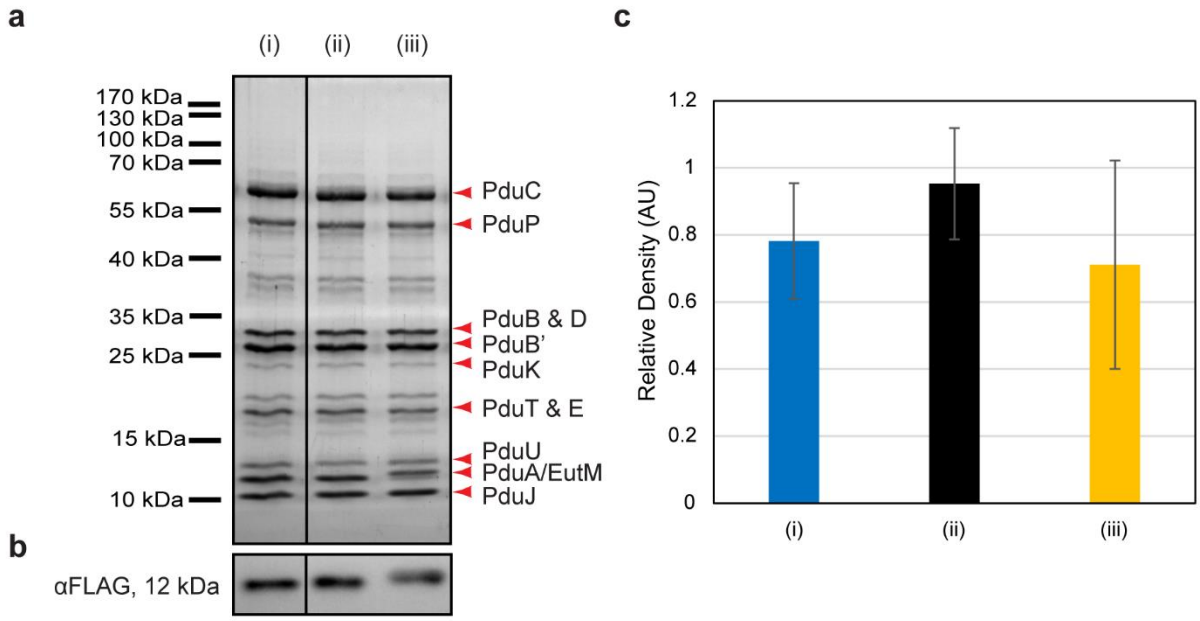


Figure 2.24 Quantification of FLAG-tagged PduA, PduA^{K37Q}, and EutM mutants by western blot (a) Coommassie stained SDS/PAGE of MCP samples purified from strains: (i) Δ*pduA::pduA-FLAG*, (ii) Δ*pduA::pduA*^{K37Q}*-FLAG*, and (iii) Δ*pduA::eutM-FLAG*. The gels used for these measurements were half 12%, half 15% acrylamide to achieve better separation of the low molecular weight species. Under this condition, the EutM-FLAG band separated from the PduU band in contrast with the 12% gel depicted in Figure 2m. Also in this case, loading of the gel was normalized by total protein concentration measured by BCA assay. (b) Representative western blot using anti-FLAG primary antibody to measure the abundance of FLAG-tagged protein in the pure MCP samples. (c) Bar chart generated from densitometry measurements of the western blots depicted in (b). The average relative density represents measurements from at least three MCP purifications. Error bars are the 95% confidence interval.

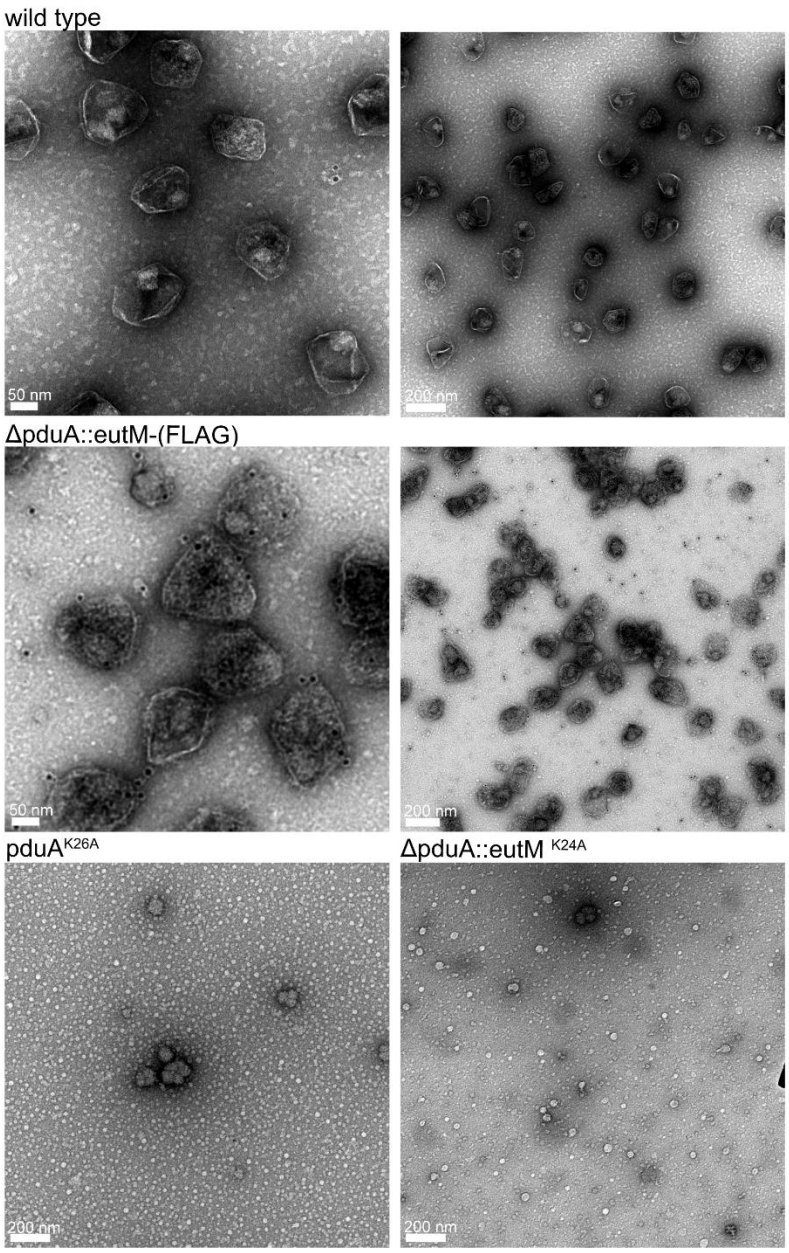


Figure 2.25 TEM of MCPs purified from four strains as labeled. The images from the wild type and the Δ*pduA::eutM-(FLAG)* strains on the left side are simply larger versions of the images featured in Figure 2 of the main text. The right-side wild type and Δ*pduA::eutM-(FLAG)* images are taken at a lower magnification. The bottom panels are uncropped versions of the images in panels k and l.

Sinha *et al*. demonstrated that a PduA^{K26A} mutant does not form MCPs due to disruption of hexamer edge binding contacts, resulting in a growth defect on 1,2-PD even though the enzymes required to metabolize 1,2-PD are present [31]. Also, MCP structures cannot be purified from the *pduA^{K26A}* strain via established methods for purifying the Pdu MCP. We replicated these results (Figures 2.18k, 2.21, 2.14), and in addition used fluorescence microscopy

with our encapsulation reporter PduD$^{1-20}$-GFP-SsrA to observe this mutant *in vivo*. We did not observe an absence of fluorescence as we expected; instead, we saw cells linked by filaments resembling the nanotube or filament structures formed by overexpression of single shell proteins or incomplete MCP operons as described previously (Figure 2.18c, g)[66,68,70,75]. These structures were accompanied by bright puncta varying in appearance and predominantly found at the cell poles. These puncta are presumed to be aggregated or poorly formed MCPs since they cannot be purified in the same manner as non-mutated MCPs. Attempts to purify MCPs from this strain gave only dilute aggregate species as seen by SDS-PAGE analysis as well as by TEM (Figure 2.18k, Figure 2.26).The same phenotype was observed when a similar mutation was made to *eutM* in the Δ*pduA::eutM* strain (Figure 2.18d, h, l, and Figure 2.26). Growth on 1,2-PD is decreased for a Δ*pduA::eutM* ^{K24A} strain, likely due to mis-formed MCPs like those formed in the *pduA*^{K26A} strain (Figure 2.15). Δ*pduA::eutM*^{K24A} and *pduA*^{K26A} mutants were used as non-MCP forming controls in subsequent experiments. Their behaviors coincide in all tests. This is further evidence that EutM is able to complement PduA and fully integrate into the Pdu MCP. Moreover, our results indicate that the same lysine binding partner is important for assembly of both the wild-type Pdu MCP and the EutM Pdu MCP chimera.

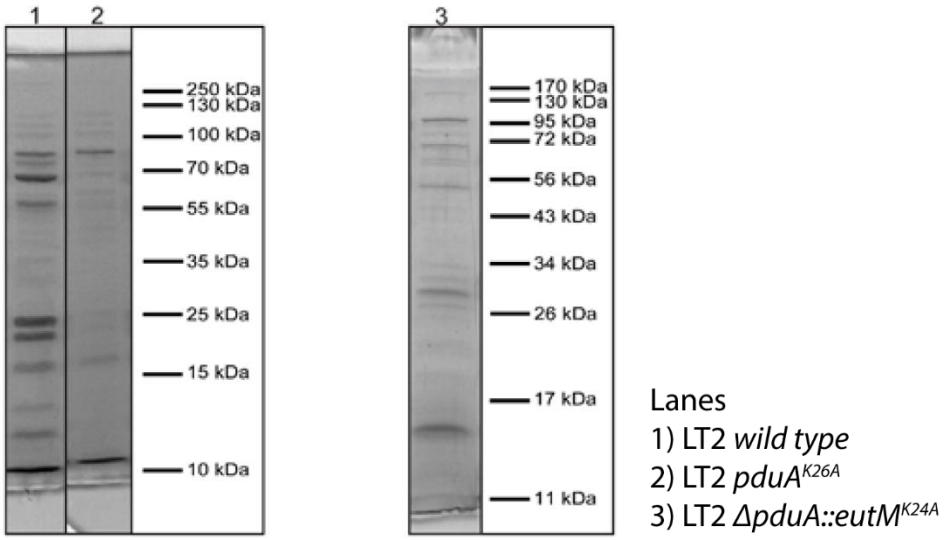


Figure 2.26 SDS-PAGE of MCPs purified from strains expressing non-assembling mutants. The MCP purification protocol was applied to non-assembling PduA and EutM mutants. The product was analyzed by SDS-PAGE to detect MCP banding pattern and discern which strains formed MCPs.

Coenzyme B_{12} is a required cofactor for the first step of 1,2-PD metabolism in the Pdu pathway. This cofactor is supplied in the growth medium because *S. enterica* cannot produce it under our aerobic conditions. Growth at low coenzyme B_{12} concentrations (20 nM) has been used previously as a method to sensitively detect whether mutations to MCP proteins result in non-assembly or defective MCP shells [24,27,31,84]. The limiting coenzyme B_{12} condition is advantageous for non-assembling MCP mutants, such as the *pduA*^{K26A} mutant, and allows improved growth because propionaldehyde production and toxicity is reduced and substrates and cofactors access the enzymes more freely. Conversely, the wild type and mutant strains that assemble well grow poorly at 20 nM coenzyme B_{12} compared to mutants that do not assemble

MCPs. We exploited this growth condition as an additional assay to detect if the mutations we made to PduA disrupted MCP assembly. We found that the Δ*pduA::eutM* mutant grows as poorly as the wild-type strain in this condition (Figure 2.27), indicating that the chimeric MCP shells form a complete diffusion barrier.

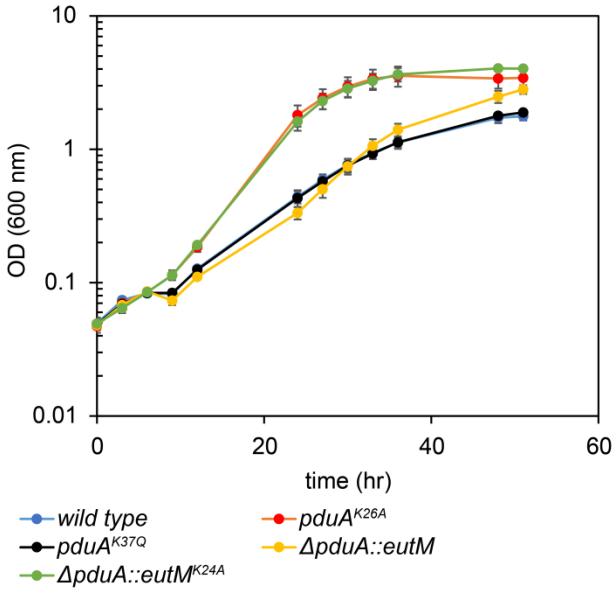


Figure 2.27 Growth curves of strains with limiting concentrations of coenzyme B_{12}. OD (600 nm) was measured over 51 hours for each mutant strain grown in 55 mM 1,2-PD and 20 nM coenzyme B_{12}. Time points represent the average of three biological replicates. Error bars represent standard error.

2.4 Discussion

We can conclude from the study of MCP stability that the Pdu MCP is remarkably stable structure. Purified native Pdu MCPs withstand temperatures up to 60°C. As temperatures increase, MCPs slowly stick together, maintaining much of the shell structure up to 70 or even 80°C before finally denaturing completely. The MCP is sensitive to pH. Stability is maintained from pH 6 to 9. Below pH 5 severe aggregation occurs. However, at pH 5, 10, and 11, MCPs interestingly become smaller. We hypothesize that this could be due to the denaturation of certain shell proteins before others, changing the content of the shell. There are nine known protein constituents of the shell, and some may be more susceptible to denaturation than others. Vernizzi et al. demonstrate how changing the relative abundance of shell components could alter shell morphology with a model of multicomponent elastic membranes [87].

The importance of the C-terminal residues of PduA and PduJ for MCP assembly is an intriguing result. Since empty Pdu shells may be formed without the expression of cargo proteins, this suggests that this portion of the protein may form unforeseen interactions with other structural components of the shell. Thus, engineering shell proteins in order to modify their interaction with the cargo may be a complex enterprise. These observations provide a starting point for designing MCPs with varying shell properties and gaining control over assembly and disassembly. The stability data also indicates that applications for MCPs *in vitro* are possible,

including MCP encapsulated reactions in extracellular environments, and using MCPs as the scaffold for a drug delivery vehicle.

We also demonstrated that the ability to form chimeric MCP structures is an important tool to control the properties of the shell and to customize this organelle for metabolic engineering. We created Pdu MCPs harboring a foreign shell protein substitution and, for the first time, showed that alterations to the Pdu MCP shell can confer enhanced function of the enclosed pathway. PduA shell protein mutants, have improved growth phenotypes with 1,2-PD as a carbon source. We suspect that these phenotypes are a result of modification of small molecule diffusion rates through pores of the MCP. The next chapter of this dissertation describes a detailed analysis of the effect of pore structure on the performance of the encapsulated metabolism.

Chapter 3: USING SINGLE AMINO ACID MUTATIONS TO THE PORE OF THE PDUA SHELL PROTEIN TO ALTER THE PERFORMANCE OF AN MCP-ENCAPSULATED PATHWAY

The following is adapted from Slininger, Jakobson, and Tullman-Ercek; biorXiv, 2017 with permission.

3.1 Background

In the previous chapter, we discovered one strategy for altering MCP function: generation of chimeric MCPs. We hypothesized that the observed change in growth phenotype was a result of altered small molecule diffusion through the pores of the MCP. To support this idea, in this chapter I describe how we incorporated a chromosomal point mutation to encode $PduA^{K37Q}$, to more closely mimic the EutM pore sequence. We then analyzed metabolite levels during growth for various strains to track how mutations affect consumption and release of each small molecule in the pathway. Next we used a saturation mutagenesis of position 37 of PduA to reveal that the pore is highly amenable to mutation, and that the hydrophobicity and charge of the amino acid at this locus are the most important factors affecting growth on 1,2-PD. This study thus demonstrates the potential for success in altering metabolite transport properties with both chimeric MCPs and MCP pore mutants, enabling MCP design and optimization for synthetic biology applications.

.

3.2 Materials and methods

3.2.1 Strains and culture conditions

The bacterial strains used in this study are variants of *Salmonella enterica* serovar Typhimurium LT2 and *Escherichia coli* DH10B (Table 3.1).

For MCP purifications or microscopy experiments, single colonies were grown in 5 mL LB (Lysogeny broth)-Lennox medium at 30°C, 225 rpm, for 24 hours. Cell cultures were diluted from the primary culture 1:1000 into NCE supplemented with 42 mM sodium succinate, 1mM magnesium sulfate, and 50 μM ferric citrate. For cultures containing the plasmid expressing $PduD^{1-20}$-GFP-ssrA, half the usual amount of antibiotic was used (17 mg/mL chloramphenicol). Culture volumes were 400 mL of media in 2L flasks for MCP purification, and 5 mL of media in 24-well blocks (Analytical Sales and Services, Inc., cat. no. 24108) for microscopy. Cells were grown at 37˚C in an orbital shaker at 225 rpm. For cultures with $PduD^{1-20}$-GFP-ssrA expressed from a pBAD33 plasmid, 1.33 mM arabinose was added when subculture reached an OD_{600} of 0.4. After five additional hours of growth, samples were taken for fluorescence microscopy or Pdu MCP purification.

For mutations to *pduA*, the gene was amplified from the *S. enterica* chromosome and inserted into a pBAD33, p15a, cm^R plasmid by golden gate assembly with BsaI restriction endonuclease [79]. This plasmid was transformed into *E. coli*. and was mutated as described above using PCR site-directed mutagenesis. EutM K24A was created in the same manner. Primers for each mutation are shown in Table 3.3.

Chromosomal modifications were made by recombineering as described by Court [80]. First, the native *pduA* gene was replaced with the *cat/sacb* cassette amplified from the TUC01

genome with primers mfsp105 and either mfsp106 or mfsp195 for *eutM* or *pduA* insertions, respectively (primer sequences listed in Table 2.3). This amplicon was used to create Δ*pduA::cat/sacb*. The e*utM* gene was amplified with primers mfsp199 and mfsp200 and used to make Δ*pduA::eutM*. To make mutations in *pduA*, the gene was first subcloned into a plasmid vector and mutated by site directed mutagenesis as described above. The inserts were then amplified from plasmids by PCR with primers mfsp197 and mfsp104 and used to make all chromosomal *pduA* mutants. For FLAG tag fusions, the non-polar Δ*pduA::cat/sacb* strain constructed with mfsp106 was used as a starting point to maintain the same RBS region for the subsequent *pduB* gene. The FLAG sequence was added by PCR to *pduA* and *eutM* sequences using primers mfsp304 and mfsp218, respectively before insertion by recombineering. Each strain was confirmed by Sanger sequencing of PCR amplicons from the PduA region of the *S. enterica* chromosome.

For growth of *S. enterica*, single colonies were picked from freshly streaked plates and were grown in 5 mL of TB (Terrific Broth) medium at 30°C, 225 rpm, for 24 hours. For growth curves with a 1,2-PD carbon source, cultures were diluted to an OD_{600} of 0.05 in 50 mL No-Carbon-E (NCE) minimal medium supplemented with (unless stated otherwise) 55 mM 1,2-PD, 150 nM coenzyme B_{12}, 1 mM magnesium sulfate, and 50 μM ferric citrate.[32] Cultures were grown in 250 mL flasks at 37°C in an orbital shaker at 225 rpm. At every time point, 1.5 mL of cell culture sample was taken for OD_{600} measurement and HPLC analysis. The doubling time of each growth curve was calculated by plotting the log of the OD_{600} vs. time. The linear region of each individual curve was identified (including a minimum of three time points). The slope of the linear regression between time points was determined, and the doubling time was calculated with $t_D = \log(2)\,/slope$.

Strain	Organism	Genotype
MFSS044	*S. enterica* serovar Typhimurium	LT2 wild type
MFSS263	*S. enterica* serovar Typhimurium	LT2 *pduA*^{K26A}
MFSS236	*S. enterica* serovar Typhimurium	LT2 Δ*pduA::cat/sacb*
MFSS252	*S. enterica* serovar Typhimurium	LT2 Δ*pduA::eutM*
MFSS273	*S. enterica* serovar Typhimurium	LT2 Δ*pduA::eutM-FLAG*
MFSS238	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37Q}
MFSS393	*S. enterica* serovar Typhimurium	LT2 Δ*pduA:: pduA* ^{K37Q}*-FLAG*
MFSS392	*S. enterica* serovar Typhimurium	LT2 Δ*pduA:: pduA-FLAG*
MFSS337	*S. enterica* serovar Typhimurium	LT2 Δ*pduA::eutM* ^{K24A}
TUC01 [80]	*E. coli*	DH10B *cat/sacB*
MFSS320	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37H}
MFSS321	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37C}
MFSS322	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37R}
MFSS323	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37G}
MFSS324	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37P}
MFSS325	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37V}
MFSS326	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37F}
MFSS327	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37T}
MFSS328	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37L}

MFSS329	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37D}
MFSS330	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37E}
MFSS331	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37S}
MFSS335	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37N}
MFSS336	*S. enterica* serovar Typhimurium	LT2 *pduA* ^{K37W}

Table 3.1 Bacterial strains used in this study

Plasmid	**Description**
mfspl205	*pduD*$^{1-20}$*-gfp-ssrA*, pBAD, p15A, chloramphenicol^{R}
Psim6 [80]	λ-red, TS (30ºC), carbenicillin^{R}

Table 3.2 Plasmids used in this study

Primer	Description	Sequence
mfsp105	Cat/sacb ΔPduA forward	ttatagtcccaactatcggaacactccatgcgaggtcttttgtgacggaagatcacttcg
mfsp106	cat/sacb ΔPduA reverse (keep PduB RBS)	ttcccttcggtaagatttttttctacatcggtgtgagggcgatcaaagggaaaactgtccatat
mfsp195	complete ΔPduA sacb reverse (no PduB RBS)	caatcacctgcgccatgatctgttccaccagctcattgctgcatcaaagggaaaactgtccatat
mfsp197	PduA knock in reverse	caatcacctgcgccatgatctgttccaccagctcattgctgctcattggctaattcccttcggta
mfsp104	PduA knock in forward	ttatagtcccaactatcggaacactccatgcgaggtctttatgcaacaagaagcactaggaatgg
mfsp199	knock in EutM forward	ttatagtcccaactatcggaacactccatgcgaggtctttatggaagcattaggaatgattgaaacc
mfsp200	knock in EutM reverse	ttcccttcggtaagatttttttctacatcggtgtgagggcg tcaaatgttgctgtcgcctttg
mfsp218	PduA::EutM-FLAG knock in reverse	ttcccttcggtaagatttttttctacatcggtgtgagggcg tcacttgtcatcgtcatctttataatcaatgttgctgtcgcctttg
mfsp304	PduA::PduA-FLAG knockin reverse	ttcccttcggtaagatttttttctacatcggtgtgagggcg tcacttgtcatcgtcatctttataatcttggctaattcccttcggta
mfsp191	PduA K37Q forward	gatgttagtgggctatgaaCagattggctccgggc
mfsp192	PduA K37Q reverse	ccagcccggagccaatctGttcatagcccactaac
mfsp202	PduA K26A forward	gaggccgctgatgcaatggttGCgtcagccaatgtgatg
mfsp203	PduA K26A reverse	ccactaacatcacattggctgacGCaaccattgcatcagc
mfsp271	PduA K37R QC f	gatgttagtgggctatgaaCGTattggctccgggc
mfsp272	PduA K37R QC r	ccagcccggagccaatACGttcatagcccactaac
mfsp273	PduA K37H QC f	gatgttagtgggctatgaaCACattggctccgggc
mfsp274	PduA K37H QC r	ccagcccggagccaatGTGttcatagcccactaac
mfsp275	PduA K37D QC f	gatgttagtgggctatgaaGACattggctccgggc
mfsp276	PduA K37D QC r	ccagcccggagccaatGTCttcatagcccactaac
mfsp277	PduA K37E QC f	gatgttagtgggctatgaaGAGattggctccgggc
mfsp278	PduA K37E QC r	ccagcccggagccaatCTCttcatagcccactaac
mfsp279	PduA K37S QC f	gatgttagtgggctatgaaTCTattggctccgggc

mfsp280	PduA K37S QC r	ccagcccggagccaatAGAttcatagcccactaac
mfsp281	PduA K37T QC f	gatgttagtgggctatgaaACCattggctccgggc
mfsp282	PduA K37T QC r	ccagcccggagccaatGGTttcatagcccactaac
mfsp283	PduA K37N QC f	gatgttagtgggctatgaaCCGattggctccgggc
mfsp284	PduA K37N QC r	ccagcccggagccaatCGGttcatagcccactaac
mfsp285	PduA K37C QC f	gatgttagtgggctatgaaTGCattggctccgggc
mfsp286	PduA K37C QC r	ccagcccggagccaatGCAttcatagcccactaac
mfsp287	PduA K37G QC f	gatgttagtgggctatgaaGGCattggctccgggc
mfsp288	PduA K37G QC r	ccagcccggagccaatGCCttcatagcccactaac
mfsp289	PduA K37P QC f	gatgttagtgggctatgaaCCGattggctccgggc
mfsp290	PduA K37P QC r	ccagcccggagccaatCGGttcatagcccactaac
mfsp291	PduA K37L QC f	gatgttagtgggctatgaaCTGattggctccgggc
mfsp292	PduA K37LQC r	ccagcccggagccaatCAGttcatagcccactaac
mfsp293	PduA K37V QC f	gatgttagtgggctatgaaGTGattggctccgggc
mfsp294	PduA K37V QC r	ccagcccggagccaatCACttcatagcccactaac
mfsp295	PduA K37F QC f	gatgttagtgggctatgaaTTTattggctccgggc
mfsp296	PduA K37F QC r	ccagcccggagccaatAAAttcatagcccactaac
mfsp297	PduA K37W QC f	gatgttagtgggctatgaaTGGattggctccgggc
mfsp298	PduA K37W QC r	ccagcccggagccaatCCAttcatagcccactaac
mfsp301	EutM K24A f	cgatgcgatggtaGCGgccgcgcgcgtgaagctg
mfsp302	EutM K24A r	gcgcgcggcCGCtaccatcgcatcggaggcctcaatcagc

Table 3.3 Primers used in this study

3.2.2 Fluorescence microscopy

Bacteria were viewed using a Nikon Ni-U upright microscope with a 100x 1.45 n.a. plan apochromat objective. Images were captured using an Andor Clara-Lite digital camera and Nikon NIS Elements software. Fluorescence images were collected using a C-FL Endow GFP HYQ bandpass filter. All images were taken with 400 ms exposure and adjusted identically for contrast in Adobe Photoshop software.

3.2.3 MCP purification

5 mL cultures were grown in LB-Miller from a single colony for 24 hours at 30°C, then subcultured 1:1000 into 400 mL of NCE with 55 mM 1,2-PD and grown for 13-15 hours at 37°C. To express heterologous cargo, at OD600~0.4 the appropriate inducer was added at the concentrations indicated and the cultures were grown for a further 5.5 hours. MCPs were purified by sedimentation as previously described [81], with the following modification: in the place of BPER-II bacterial lysis solution (Thermo), a solution of 1% w/v octylthioglucoside (Santa Cruz Biotech) in 20 mM Tris (Fisher) pH 7.5 in water was used for cell lysis. They were stored at 4°C until analysis by SDS-PAGE or TEM. Total protein concentration of pure MCP sample was determined by Pierce BCA protein assay kit (Thermo Fisher).

3.2.4 SDS-PAGE and western blot

Denaturing protein electrophoresis (SDS-PAGE) was performed on purified MCPs using a 12.5% polyacrylamide gel unless otherwise noted. A 130V potential was applied for 70 minutes. Unless otherwise noted, equal volumes (15 µL) of each MCP sample were loaded onto the gel, and densitometry quantifications were normalized by sample protein concentration. Gels

were either stained with Coomassie dye or further processed for western blot. Western blotting was done with a PVDF membrane according to standard protocols. Blotting was done with a mouse anti-FLAG primary antibody (Sigma) 1:2000 dilution in 50 mM Tris 150 mM NaCl pH 7.6 with 0.05% Tween-20 (TBST) with 1% w/v dry milk. The secondary antibody was HRP-conjugated goat anti-mouse antibody (Thermo) diluted 1:1000 in TBST. Labeling was visualized with west-pico chemiluminescent substrate (Thermo) using a Bio-Rad ChemiDoc XRS+.

Densitometry analysis of Coomassie stained gels was performed using at least three different MCP purifications for each strain. Each purified sample was run on three different gels. Images were analyzed using the Image Lab software (Bio-Rad). The absolute signal measurement for each band was normalized by the signal for the band containing proteins PduT and PduE (band 7 in Figure 2.17). This band had the least variation in signal over all replicates.

3.2.5 Transmission electron microscopy

Samples were placed on 400 mesh formvar coated copper grids with a carbon film after the grids were treated by glow discharge. 10 μL of purified MCPs at a concentration of about 0.1 mg/mL was placed on each grid for 2 min. The grids were washed three times with deionized water before fixation. Then, 10 μL of 2% glutaraldehyde in water was placed on each grid for 1 min. Grids were then washed an additional three times with deionized water. Finally, samples were stained in 1.6% aqueous uranyl acetate for 1 min. All solutions were centrifuged directly before use to avoid aggregate or particulate contact with grids. Samples were imaged with an REI Tecnai T12 transmission electron microscope and a Gatan Ultrascan 1000 camera (Gatan, Pleasanton, CA).

3.2.6 Size distribution analysis

TEM images were imported into ImageJ and the ellipse tool was used to manually draw an ellipse circumscribing each MCP. The measure function was used to determine the Feret's diameter for at least 100 MCPs per condition tested, obtained from random fields of view. The diameters were averaged and reported as a bar graph, with the error bars representing one standard deviation from the mean Feret's diameter for MCPs at that condition.

3.2.7 High Performance Liquid Chromatography (HPLC) media analysis

At each time point, a 750 µL sample of whole cell culture was taken. This was centrifuged for 5 min at 13,000 g to remove cells. The supernatant was decanted and frozen until time of analysis. Before analysis samples were thawed and filtered. The Shimadzu HPLC detection system was equipped with a Bio-Rad Aminex HPX-87H (300 by 7.8 mm) ion exclusion column, LC-20AD solvent delivery system, SIL-20AC auto sampler, RID-10A refractive index detector, and SPD-M20A diode array detector (Shimadzu, Kyoto, Japan). Mobile phase is 5 mM H_2SO_4 with 0.4 mL min^{-1} flow rate [11]. The column temperature is 35°C.

3.2.8 Multiple Sequence Alignment construction and evolutionary rate analysis

Evolutionary rate of each amino acid locus of PduA was calculated using the Rate4Site program version 2.01 [88]. The input for this tool was a multiple sequence alignment (MSA) of 193 non-redundant PduA homologs. The homologs were retrieved and the MSA assembled by the jackhmmer web server (https://www.ebi.ac.uk/Tools/hmmer/search/jackhmmer). Rate4Site uses this data to construct a phylogenetic tree. With both MSA and tree information, a rate-

scaling factor is calculated that indicates how rapidly each residue evolves relative to the mean protein rate. Default settings (including Bayesian framework) were used here. A larger MSA of about 1000 sequences, also retrieved and aligned with jackhammer, was used to calculate amino acid frequency.

3.3 Results

3.3.1 Single amino acid mutation to PduA pore alters pathway behavior

We hypothesized that the difference in growth caused by the EutM substitution was the result of a change in diffusion through the central pore of the shell protein. To support this idea, we incorporated a chromosomal point mutation to encode PduA^{K37Q}, to more closely mimic the EutM pore sequence. The analyses performed for the Δ*pduA::eutM* substitution mutant were repeated with *pduA*^{K37Q}. This strain was grown in minimal media with 1,2-PD as the sole carbon source to test the performance of the encapsulated 1,2-PD utilization metabolism. Strikingly, the *pduA*^{K37Q} mutation confers improved growth (Figure 2.14, 2.15) over wild type while still maintaining normal MCP morphology and protein content (Figures 2.17m, 2.19, 2.20, 2.21, 2.23, 3.1, 3.2).

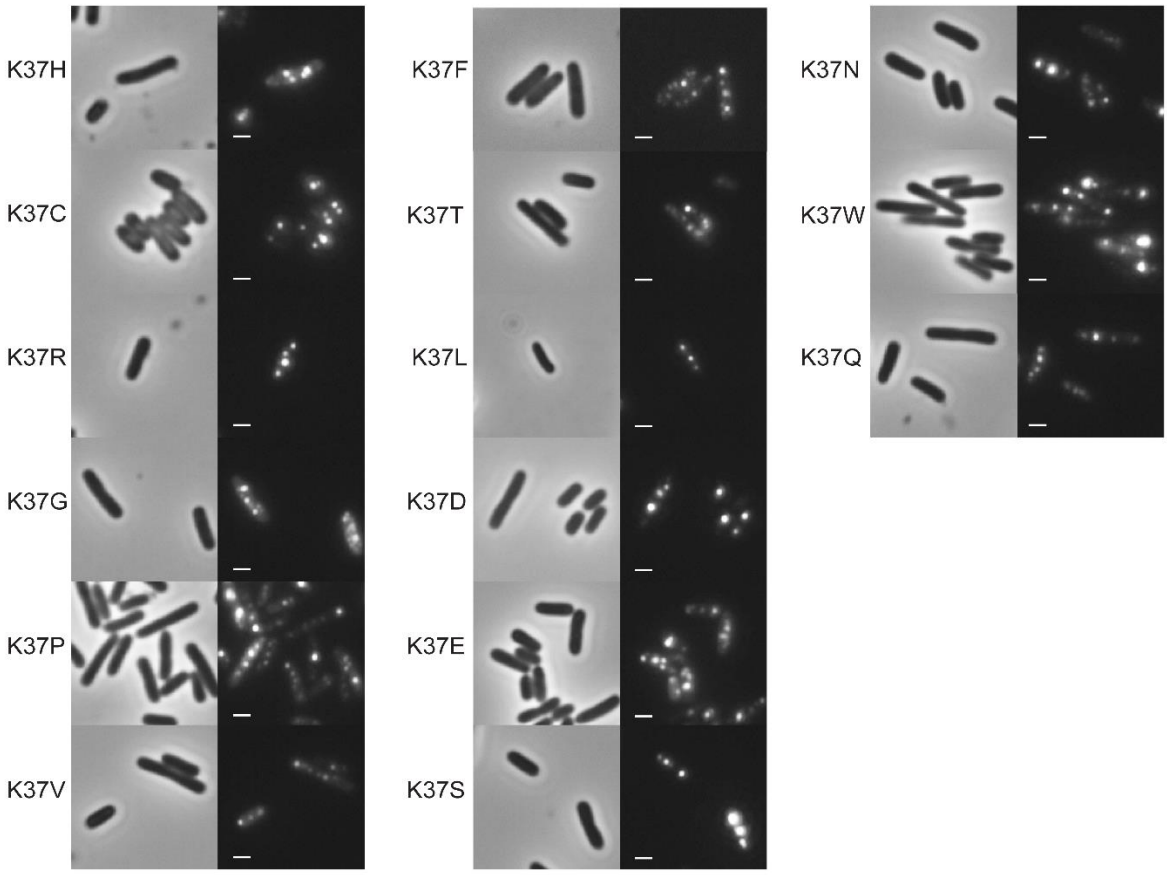


Figure 3.1 Fluorescent microscopy of PduA variant strains. Representative phase contrast and fluorescent microscopy images of strains harboring *pduA* variations co-expressing PduD$^{1-20}$-GFP-ssrA. Images taken at 17 hours after subculturing into media containing 1,2-PD, and 5

hours after induction of pBAD PduD$^{1-20}$-GFP-ssrA with .02% arabinose. Labels indicate the mutation incorporated into PduA at residue 37. Scale bar is 1 μm.

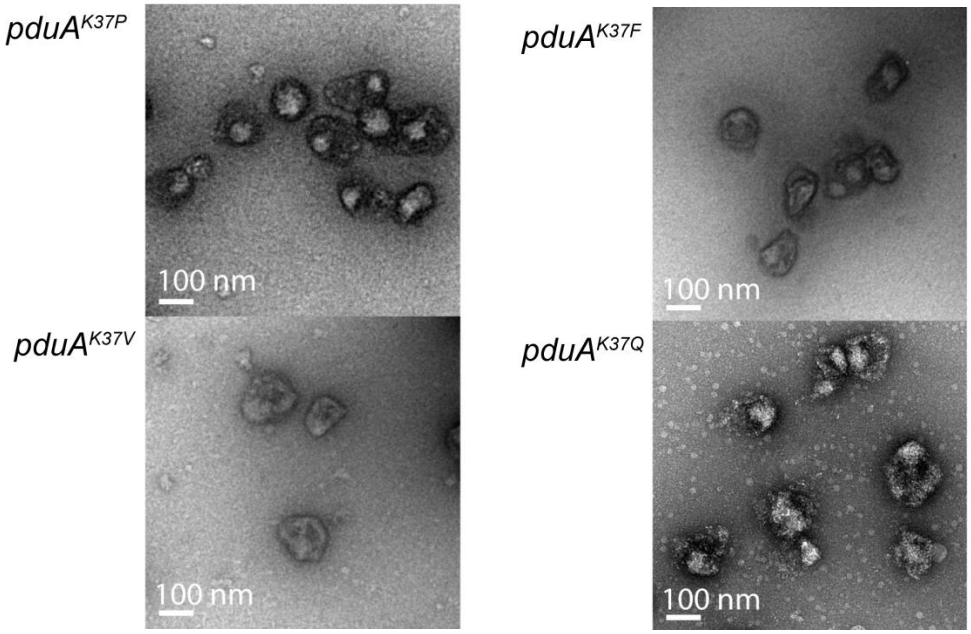


Figure 3.2 TEM of MCPs purified from strains expressing PduA variants. MCPs were purified from select strains from the PduA K37 library, including: LT2 *pduA*^{K37P}, *pduA*^{K37V}, *pduA*^{K37F}, *pdu*^{K37Q}. Strains were grown in minimal medium with 55 mM 1,2-PD to induce MCP formation. MCPs were purified by centrifugation and imaged by TEM.

HPLC analysis was performed to measure metabolite levels in the media over the course of growth on 1,2-PD (Figure 3.3). Fast growth on 1,2-PD is associated with decreased aldehyde levels in the media.[11] Therefore, we expected that the Δ*pduA::eutM* and *pduA*^{K37Q} strains would accumulate less aldehyde in the media. Indeed, we found that the accumulation of downstream products propionaldehyde, 1-propanol, and propionate are decreased in both the Δ*pduA::eutM* and *pduA*^{K37Q} mutants. We also found that the consumption of 1,2-PD occurred at roughly the same rate in wild-type, Δ*pduA::eutM*, and *pduA*^{K37Q} strains (the differences in calculated rate are not statistically significant).

Non-assembling mutants (*pduA*^{K26A} and Δ*pduA::eutM* ^{K24A}) show the highest levels of propionaldehyde, at 8.6 and 10 mM, respectively, compared to 2.1 mM for the wild-type MCP (Figure 3.4). Propionaldehyde is toxic to *S. enterica* when exogenously added at concentrations of about 8 mM.[14] Also, since propionaldehyde is volatile, accumulation of this molecule in the media results in some unmeasured amount of propionaldehyde evaporation to the head space. The differences in propionaldehyde concentrations among wild type, Δ*pduA::eutM*, and *pduA*^{K37Q} are more subtle. Wild-type *S. enterica* maintains about 2 mM propionaldehyde for much of the exponential phase. We observed no detectable propionaldehyde in Δ*pduA::eutM* cultures apart from a single spike usually occurring between 27 and 30 hours reaching a level averaging 2 mM. The *pduA*^{K37Q} strain behaved similarly to the Δ*pduA::eutM* strain with respect to propionaldehyde accumulation. Therefore, while these strains reach similar maximum concentrations of propionaldehyde in the media, the Δ*pduA::eutM*, and *pduA*^{K37Q} *mutants* sustain this concentration for a shorter period.

In wild-type cultures the 1-propanol concentration level stabilizes at around 25 mM from 36 hours onward. In contrast, Δ*pduA::eutM* and *pduA*^{K37Q} strains reach a maximum of about 20 mM 1-propanol, and the concentration decreases after 36 hours. This indicates that the accumulation rate of 1-propanol is slower in the Δ*pduA::eutM* and *pduA*^{K37Q} mutants. Non-assembling mutants accumulate even less 1-propanol, likely due to the greatly increased escape of the aldehyde before conversion. After the supply of 1,2-PD is exhausted, 1-propanol in the media is likely converted back to propionaldehyde and then consumed.[23] This process in addition to some evaporation from the headspace results in the decline in 1-propanol concentration at the end of each growth period. A surprising trend is seen in the concentration of propionate observed in the media over time. We observed increased propionate accumulation in strains that grew poorly. Non-assembling mutants ranked highest in propionate accumulation, followed by the wild type, then *pduA*^{K37Q} and Δ*pduA::eutM*.

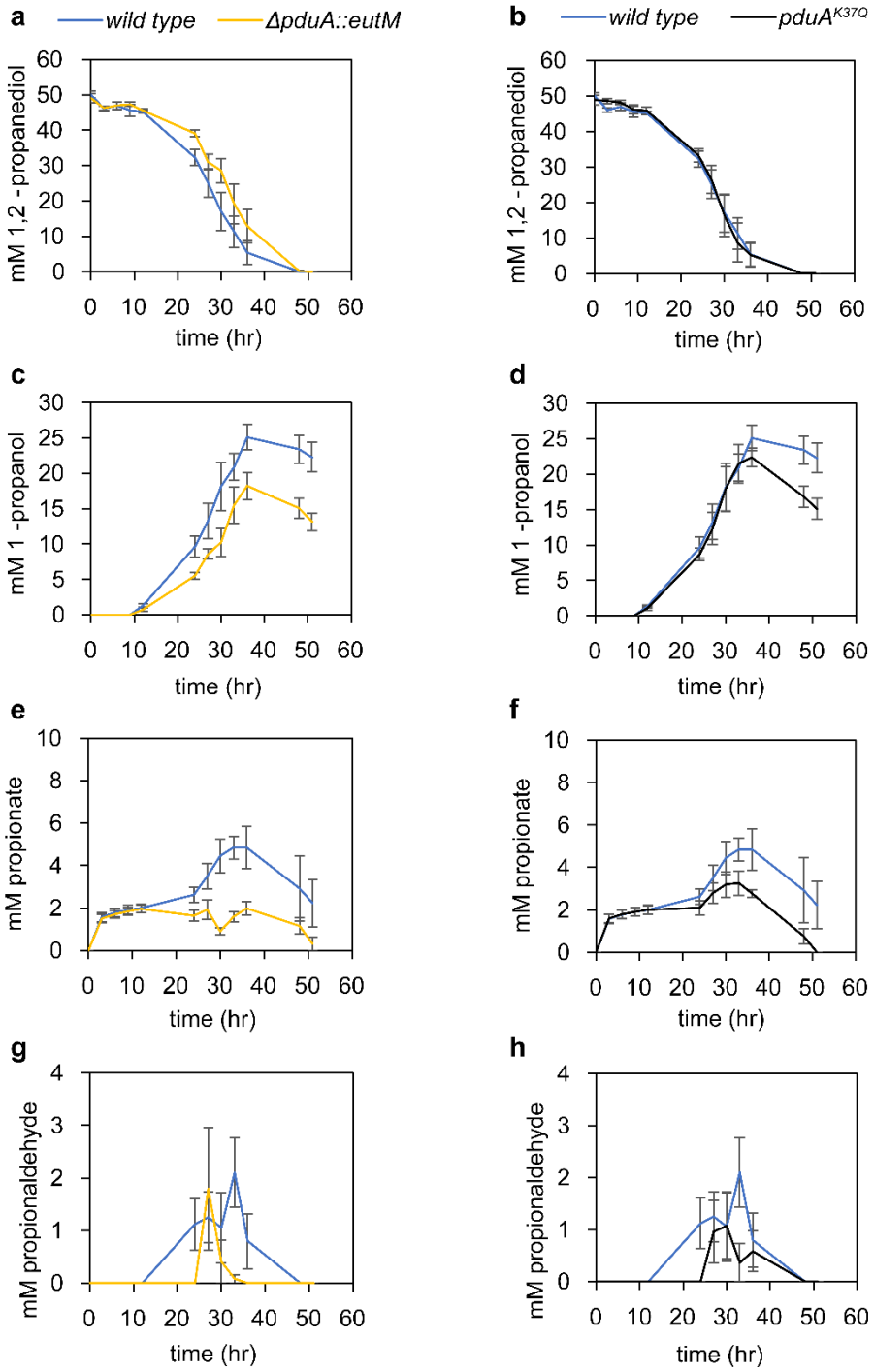


Figure 3.3 HPLC analysis of MCP-forming strains during growth on 1,2-PD. The concentrations of 1,2-PD and products of the Pdu pathway were monitored during growth on 55 mM 1,2-PD and 150 nM coenzyme B_{12}. Mean values of metabolites are plotted vs time. Metabolites measured are 1,2-PD, 1-propanol, propionate, and propionaldehyde, as labeled. In the left hand column (a, c, e, g) the wild-type *S. enterica* LT2 is compared to LT2 Δ*pduA::eutM*. In the right hand column (b, d, f, h) the wild-type strain is compared to LT2 *pduA* ^{K37Q}. Error bars represent standard error from six measurements started on four different days.

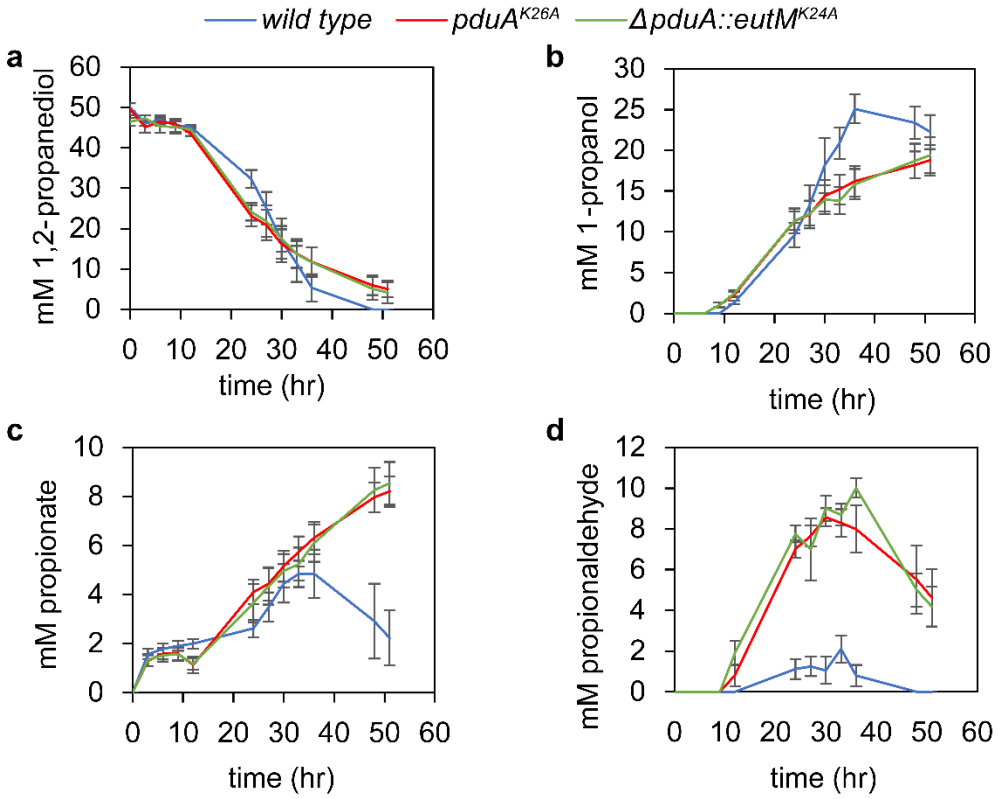


Figure 3.4 HPLC analysis of non-assembling shell protein mutant strains. The concentrations of 1,2-PD and products of the Pdu pathway were monitored during growth on 55 mM 1,2-PD and 150 nM coenzyme B_{12}. Mean values of metabolites are plotted vs time. Metabolites measured are 1,2-PD, 1-propanol, propionate, and propionaldehyde, as labeled. The wild-type LT2 strain is compared to LT2 *pduA* ^{K26A}and LT2 Δ*pduA::eutM* ^{K24A} strains. Error bars represent standard error from six measurements started on four different days.

3.3.2 Saturation mutagenesis of K37 of PduA produces a range of pathway behaviors

We further explored the relationship between the pore and 1,2-PD metabolism by creating a genomically-encoded library of PduA variants that differ at residue 37. We checked each mutant for MCP formation by fluorescence microscopy using our encapsulation reporter PduD$^{1-20}$-GFP-ssrA (Figure 3.1). All mutants exhibited fluorescent puncta, indicating the formation of MCPs. Growth on 1,2-PD with a limiting concentration of coenzyme B12 further demonstrated that all mutants did not form "leaky" MCP structures (Figure 3.5). Finally, we selected a few of the library members for purification. Purified MCPs from these strains are all similar to wild-type MCPs when examined by SDS-PAGE banding pattern and by TEM. (Figure 3.6 and 3.2).

The library was then screened for growth in NCE with 1,2-PD and 150 nM coenzyme B12 (Figure 3.7). The end-point cell density for each mutant is summarized in Figure 3.8a. To analyze these data we compared average end-point cell density to the characteristics of the amino acid substituted in each mutant. The variables considered were the predicted amino acid pK_a,

hydropathy index, and residue molecular weight. We found a moderate negative correlation between hydropathy index and final cell density (correlation coefficient of -0.47), indicating more hydrophobic side chains generally caused a decrease in growth (Figure 3.8c). Final cell density was also plotted versus pK_a for several charged amino acid side chains. There was a negative correlation (correlation coefficient of -0.6) suggesting that a more positively charged pore is detrimental to growth on 1,2-PD (Figure 3.8b). There was no clear trend when growth was compared to size of the amino acid side chain alone (Figure 3.9). This indicates that, of the three characteristics considered, the hydrophobicity and charge of the amino acids in the pore were most important for retention of small molecules like propionaldehyde within the MCP.

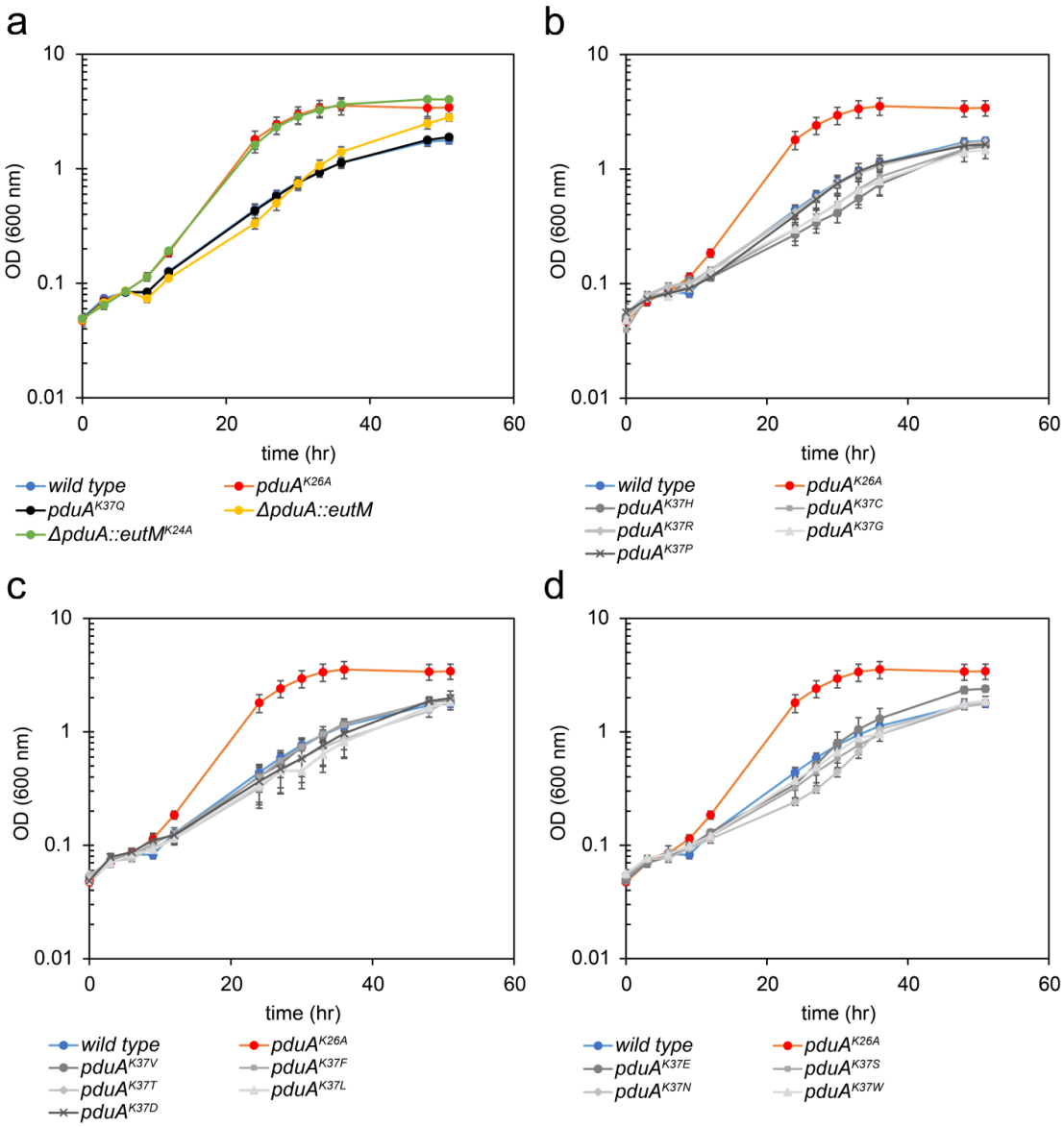


Figure 3.5 Growth curves of strains expressing the PduA K37 mutagenesis library in media with 1,2-PD and limiting concentrations of coenzyme B_{12}. OD (600 nm) was measured over 51 hours for each mutant strain grown in 55 mM 1,2-PD and 20 nM coenzyme B_{12}. Time points represent the average of three biological replicates. Error bars represent standard error.

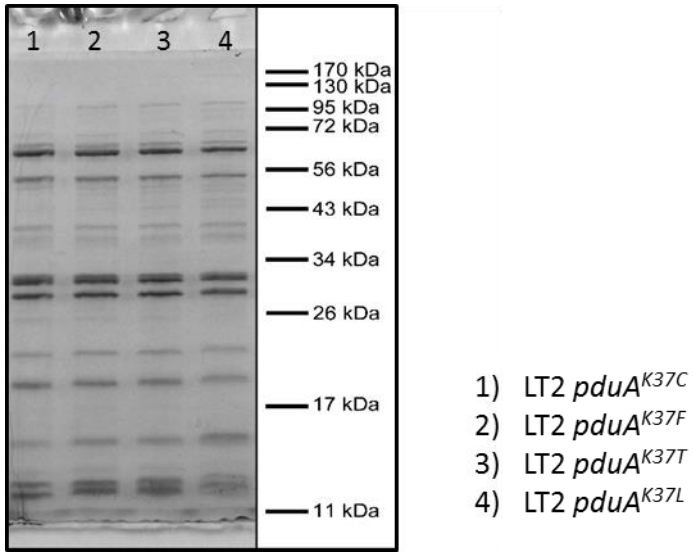


Figure 3.6 SDS-PAGE of MCPs purified from strains expressing PduA variants. The MCP purification protocol was applied to a selection of strains. The product was analyzed by SDS-PAGE to detect MCP banding pattern and discern which strains formed MCPs.

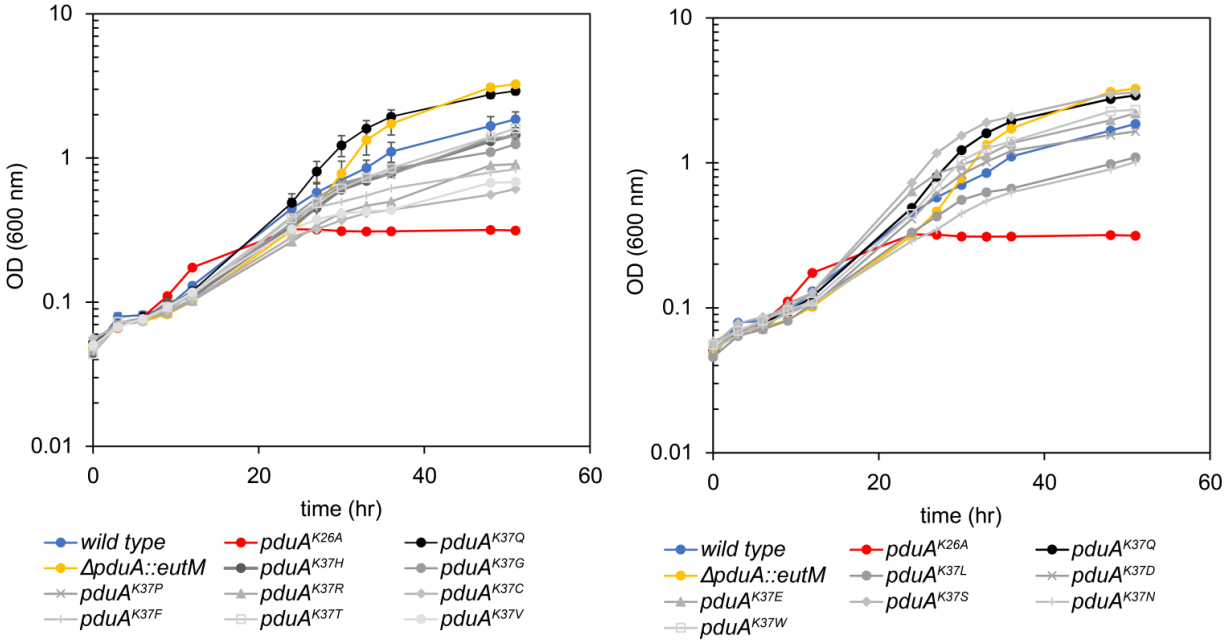


Figure 3.7. Growth curves of strains expressing the PduA K37 mutagenesis library in media with 55 mM 1,2-PD and 150 nM coenzyme B_{12}. OD (600 nm) was measured over 51 hours for each mutant strain grown in 55 mM 1,2-PD and 150 nM coenzyme B_{12}. Each time point is an average for $n \geq 4$.

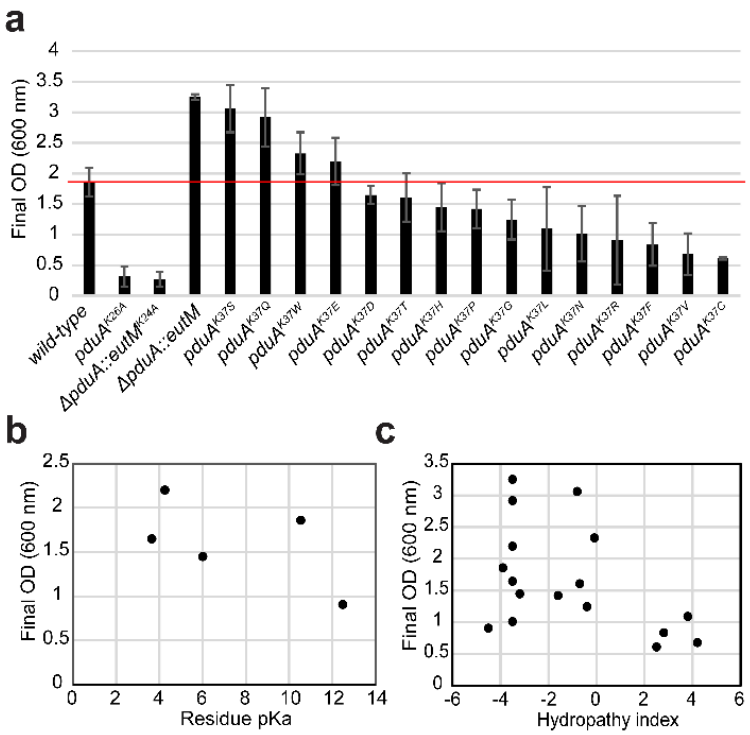


Figure 3.8 A library of amino acids at residue 37 of PduA confers a variety of growth effects (a) Mean values of final OD (600nm) are plotted for various PduA mutants from growth on 1,2-PD with 150 nm coenzyme B_{12}. Error bars represent standard error from four different measurements started on separate days. The red line highlights wild-type growth for comparison. (b) Plot of final OD vs pK_a of charged amino acids substituted at position 37 of PduA. (c) Plot of final OD vs hydropathy index of amino acids substituted at position 37 of PduA.

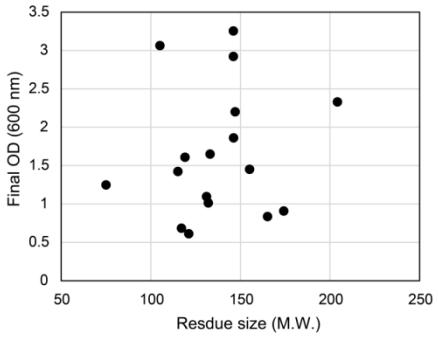


Figure 3.9 Comparison of amino acid size versus growth on 1,2-PD for site-saturated library at residue 37 of PduA. Strains harboring mutations at amino acid position 37 were grown in 55 mM 1,2-PD and 150 nM coenzyme B_{12}. Plot of final OD (600 nm) after growth for 51 hours vs molecular weight (M.W.) of amino acids substituted at position 37 of PduA.

3.3.3 Evolutionary rate analysis of PduA

Upon analysis of the PduA K37 mutagenesis library, we were surprised to observe that the pore loop of PduA is very tolerant to mutation; indeed, all mutants formed MCPs similar to wild type. It is remarkable that this locus of PduA can accommodate any amino acid without a compensating mutation. For future engineering of the MCP shell, it is desirable to know if other residues may be mutated to change MCP function or to add chemical handles for the conjugation of useful moieties. We therefore used the Rate4Site program (version 2.01) to calculate an evolutionary rate scaling factor for each residue of PduA relative to the overall evolution rate of the protein [88]. The input of this program is a multiple sequence alignment of 193 non-redundant PduA homologs from a variety of MCP systems. The resulting output is mapped over the structure of PduA (Figure 3.10a). Residues of a darker red color have high evolutionary rates, while white residues evolve slowly relative to the average rate for the entire protein. The evolutionary rate parameter represents the number of non-synonymous replacements expected per unit time for a specific residue. This parameter is related to conservation score but the rate calculation also uses the architecture and branch lengths of the phylogenetic tree in order to account for evolutionary time. If we detect residues that change often over time, but greatly affect the protein function, this could indicate a structural feature important for adapting function to environmental change.

The results indicate that pore positions 37 and 40 have a higher evolutionary rate than the average evolution rate for the entire protein, while the isoleucine and glycine at positions 38 and 39 evolve more slowly than the protein average. This suggests that a hydrophobic motif at the narrowest part of the pore is essential for MCP function in general, whichever pathway is encapsulated within. On the other hand, while the residues forming rings on either side of the hydrophobic motif evolve more quickly, they evidently also affect encapsulated pathway function, most likely by inhibiting or facilitating diffusion of one or more metabolites. The mutability of certain pore residues may allow the flexibility to encapsulate a number of pathways involving a diversity of small molecules.

Interestingly, the residues K90-I92 near the C-terminus of PduA, as well as several other residues on both concave and convex surfaces, are also very rapidly evolving (labeled on Figure 3.10a). These areas of the protein structure are less constrained and may serve as good targets for protein modifications, particularly for control of cargo loading or for labeling the outer surface for engineering applications.

Among the strains in the K37 mutation library, the strains with Gln, Glu, Ser, and Trp at position 37 grew to a greater final OD_{600} than the wild-type strain under laboratory conditions. Curiously, this indicates that the *S. enterica* serovar lab strain used in this study is not optimized for flux through the Pdu pathway and other factors may affect the fitness of LT2 in the host gut. We searched among PduA homologs to find out if any of the amino acid variants that resulted in highest growth were among those commonly found in nature. A distribution of the frequencies of each amino acid aligned with position 37 is given in Figure 3.10b. Together, strains with Gln, Glu, Ser, or Trp aligned with position 37 only represent 9.38% of over 1000 related sequences. The population of PduA homologs in nature is not enriched for these high performers.

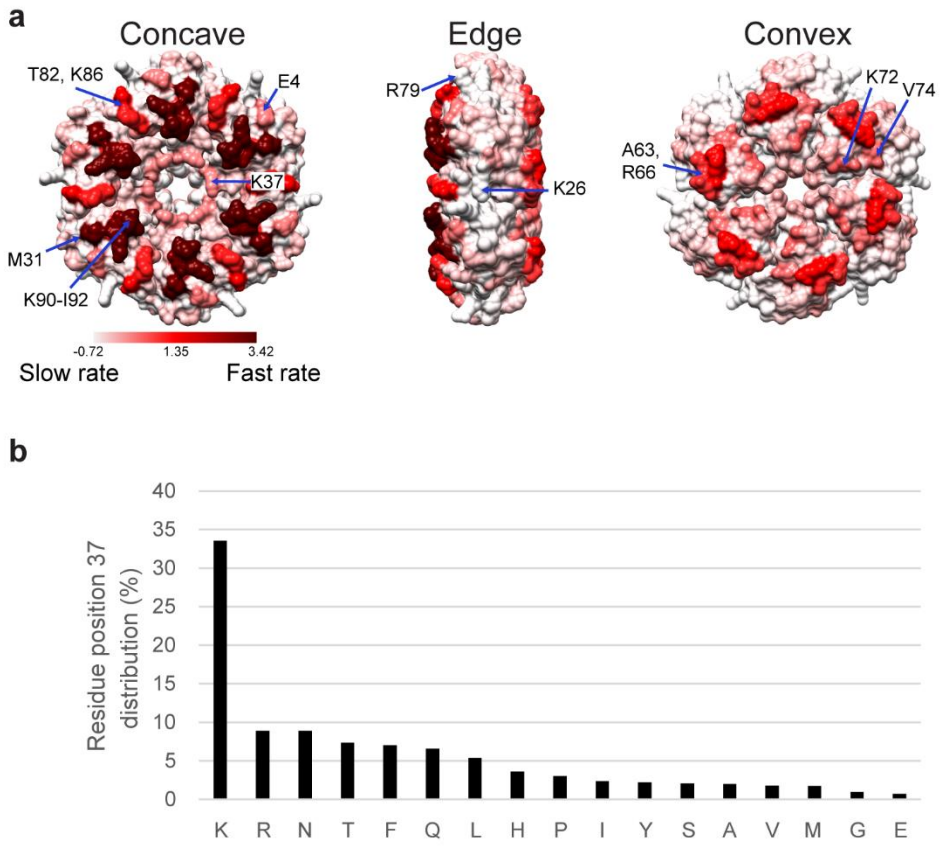


Figure 3.10 Phylogenetic analysis of PduA homologs. (a) Rate4Site evolutionary rate scaling factor mapped over the crystal structure of PduA. White represents most conserved residues. Dark red represents residues with high evolution rates. Some residues with highest and lowest evolutionary rates are labeled for reference. (b) Histogram showing the % frequency of each amino acid aligning with PduA K37 in an MSA of n=1000 homolog sequences.

3.4 Discussion

A single point mutation to the pore causes the same effect as a complete shell protein substitution. This is strong evidence indicating that diffusion of molecules through the pore of the MCP shell is altered, causing the observed changes in growth and metabolite accumulation. Specifically, the modification of lysine to glutamine changes the charge environment in the pore, and may also change the polarity or hydrogen bonding configuration. Both of these factors likely alter the diffusion barrier small molecules experience [27,28]. It is convenient for MCP engineering purposes that such small mutations to this region can confer improved function of the whole organelle structure.

The HPLC analysis of metabolites in the media demonstrated that both the *ΔpduA::eutM* and $pduA^{K37Q}$ mutations resulted in altered levels of metabolite accumulation over the growth period. To offer an explanation for these behaviors, we propose that the *ΔpduA::eutM* and $pduA^{K37Q}$ mutants restrict the diffusion of intermediate propionaldehyde, the negatively charged propionyl-phosphate product, or both. In the following discussion, we consider how this hypothesis is supported by the observed metabolite concentrations and growth phenotype.

The restriction of propionaldehyde diffusion by the *ΔpduA::eutM* and *pduA*^{K37Q} mutants would directly result in the observed decreases in propionaldehyde accumulation compared to the wild-type. Since propionaldehyde is volatile, increased accumulation in the media leads not only to toxicity, but also to a loss of carbon to the gas phase. Preventing aldehyde escape may also result in a higher local propionaldehyde concentration and a kinetic enhancement of pathway throughput. Both of these effects contribute to the differences in final cell density observed between strains, leading to fastest growth in the *ΔpduA::eutM* and *pduA*^{K37Q} mutants.

Restriction of propionyl-phosphate diffusion could result in the decreased accumulation of both propionate and 1-propanol by the *ΔpduA::eutM* and *pduA*^{K37Q} mutants. To illustrate, non-assembling mutants allow free diffusion of propionyl-phosphate, facilitating its conversion to propionate by the PduW kinase, which is not localized to the MCP. In contrast, propionyl-phosphate molecules face a moderate diffusion barrier in the wild-type pore, resulting in decreased propionate release compared to non-assembling mutants. Thus, the *ΔpduA::eutM* and *pduA*^{K37Q} mutants may form a greater diffusion barrier since even less propionate is released than in the wild-type strain. The mitigation of 1-propanol accumulation in *ΔpduA::eutM* and *pduA*^{K37Q} mutants might also be due to restricted propionyl-phosphate diffusion. Decreased product escape could cause a pathway back-up reducing the rate of NADH production, leading to slower turnover in the 1-propanol branch of the pathway involved in NADH recycling. Considering pore structure, it is intuitive that a positively charged lysine in the wild-type PduA pore would facilitate the interaction of a negative product such as propionyl-phosphate more than the neutral glutamine, and allow a greater rate of egress [28,89]. Propionate escape in the wild-type could lead directly to reduced growth rate, or slower propionate uptake into central metabolism could be a result of aldehyde toxicity.

We recognize that there is still much to learn about the fate of propionate and the relationship between the 1,2-PD utilization pathway and the rest of central metabolism. According to our current understanding, the effect of decreased propionaldehyde leakage should result in flux enhancement. However, though strains that appear to restrict aldehyde leakage grow better, they also release less of the downstream products, propionate and 1-propanol. Further study is necessary resolve these two effects.

In this work, we demonstrate the flexibility to use protein engineering to alter the pore structures of PduA without abrogating MCP structure. For the first time, mutations to the Pdu MCP shell have resulted in enhanced function of the enclosed pathway. PduA shell protein mutants have improved growth phenotypes with 1,2-PD as a carbon source. We uncovered significant evidence that these phenotypes are due to modification of small molecule diffusion rates through the pores of the MCP, leading to escape of the intermediate aldehyde and possibly preventing the escape of the product propionyl-phosphate. These trends are observable through HPLC analysis of metabolites and are supported by our analysis of protein structure and MCP assembly. An investigation of the relationship between amino acid chemistry in the pore and the resulting growth performance indicates that charge and hydrophobicity of the side chain play key roles. Shells with more negatively charged pores tended to grow poorly on 1,2-PD, while an increase in the hydropathy of the residue at position 37 also decreased the fitness of the strain. The PduA K37 mutagenesis library produced a range of growth phenotypes with strains that grew on 1,2-PD faster or slower than the wild type. Taken together, data in this and the previous chapter demonstrate that generation of chimeras and/or modification of shell protein pore residues are both promising techniques to optimize MCP shell characteristics for an encapsulated pathway of interest and to achieve useful phenotypes not found in nature.

Chapter 4: TOWARDS THE ENCAPSULATION OF MULTI-ENZYMATIC CARGOS

4.1 Background

Several single heterologous reporter proteins or enzymes have been successfully encapsulated in well-studied MCP systems such as the 1,2-propanediol utilization (Pdu) and ethanolamine utilization (Eut) MCPs of numerous enteric bacteria, as well as the carboxysome of cyanobacteria. These engineered cargos include various fluorescent proteins [66,85,69,37,42,29,62], β-galactosidase [37,60], and polyphosphate kinase [90](Figure 1b). Also, efforts to target a two-enzyme system to the MCP resulted in some success. Lawrence et al loaded a pyruvate decarboxylase and an alcohol dehydrogenase within the shell of the 1,2-propanediol utilization (Pdu) MCP of *Citrobacter freundii* recombinantly expressed in *Escherichia coli* to form ethanol. Through encapsulation, the production of ethanol was increased by roughly 50%, with the greatest advantage happening when high levels of shell proteins were expressed [34]. However, researchers have not yet demonstrated the encapsulation of pathways with three or more enzymes, nor the encapsulation of classes of enzymes that are functionally dissimilar from those associated natively with MCP systems.

Over time and through trial and error, our lab and others in the field have generated insights that will help guide future efforts to use the MCP as a metabolic engineering tool. This chapter highlights the progress made and discusses the factors influencing the encapsulation of multi-enzyme cascades within MCP systems, with the aim of facilitating the expansion of this technology.

The first step in sequestering a heterologous pathway within bacterial microcompartment is ensuring that each enzyme of the pathway can be encapsulated. Therefore, a major aim of MCP research is to develop robust methods to encapsulate a wide range of non-native enzymes for the construction of a custom nano-bioreactor. Towards this goal we have used enzymes from the violacein biosynthesis pathway as a model multi-enzyme cargo to refine enzyme targeting methods.

Violacein is a purple pigment produced by the bacterium *Chromobacterium violaceum* which has antibiotic properties as well as cytotoxicity towards some tumor cells [91,92]. The violacein production pathway comprises the enzymes VioA, -B, -C, -D, and -E. However, if just VioA, VioB, and VioE are expressed, a green pigment intermediate, prodeoxyviolacein (PDV), is formed (Figure 4.1). VioA, -B, and –E were chosen as a model multi-enzyme cargo since the product is easily assayed by absorbance, lending this construct to future high-throughput experiments.

Frequently, GFP is used as a reporter to screen functionality of signaling peptides *in vivo*. $PduP^{1-18}$, $PduD^{1-20}$, and the other signal sequences described in Chapter 1 are necessary and sufficient to target other proteins to lumen of the Pdu MCP if translationally fused to the N-terminus. While attempting to attach this tag to violacein enzyme targets, among others, we determined that many enzymes are not amenable to N-terminal fusions, and a more flexible tagging system would be beneficial. In this chapter we report the development of C-terminal encapsulation peptides. We also describe how various encapsulation peptides and configurations impact the expression level and resultant encapsulation of reporter proteins. We find that choice of encapsulation tag has multi-layered effects for the optimization of encapsulated enzyme

activity. Finally, we describe the difficulties encountered for encapsulating the violacein pathway and summarize the parameters influencing this process.

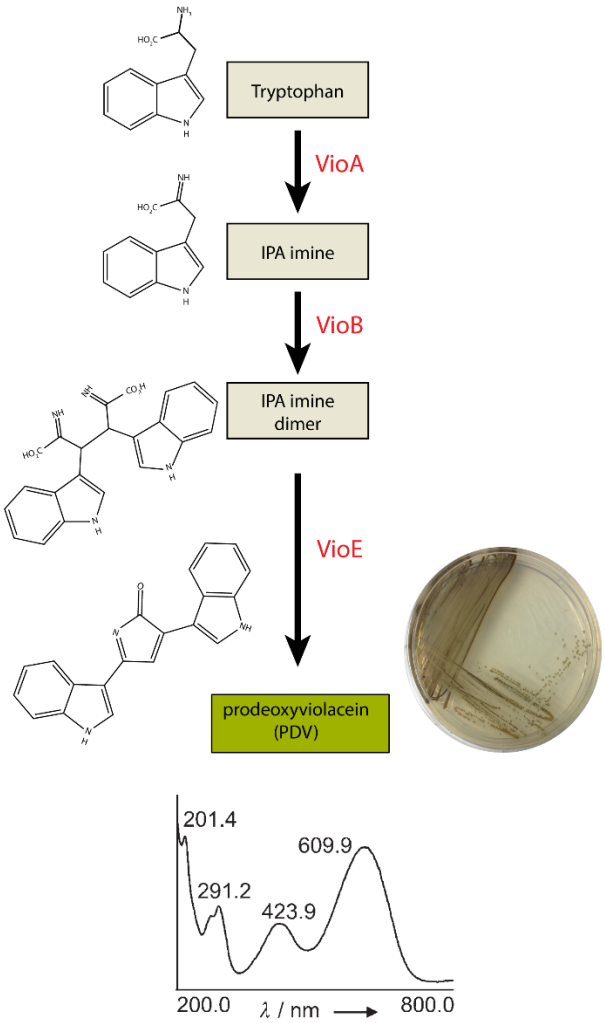


Figure 4.1 Diagram of violacein pathway. Structures of the substrate, intermediates, and product are shown at left. At the bottom the absorbance spectrum for prodeoxyviolacein is shown. At right is an image of *S. enterica* expressing VioA, VioB, and VioE and producing visible green pigment [93].

4.2 Materials and methods

4.2.1 Strains and culture conditions

The bacterial strain used in this study is *Salmonella enterica* serovar Typhimurium LT2. For MCP purifications or microscopy experiments, single colonies were grown in 5 mL LB (Lysogeny broth)-Lennox medium at 30°C, 225 rpm, for 24 hours. Cell cultures were diluted from the primary culture 1:1000 into NCE supplemented with 42 mM sodium succinate, 1mM magnesium sulfate, and 50 μM ferric citrate. For cultures containing plasmid expressing GFP reporter, half the usual amount of antibiotic was used (17 mg/mL chloramphenicol). Culture volumes were 400 mL of media in 2L flasks for MCP purification, and 5 mL of media in 24-well blocks (Analytical Sales and Services, Inc., cat. no. 24108) for microscopy. Cells were grown at

37˚C in an orbital shaker at 225 rpm. For cultures with GFP reporter expressed from a pBAD33 plasmid, 1.33 mM arabinose was added when subculture reached an OD_{600} of 0.4. After five additional hours of growth, samples were taken for fluorescence microscopy or Pdu MCP purification.

The plasmids encoding GFP reporters with N-terminal signal sequence fusions were constructed by Chris Jakobson as described in published work [39]. The plasmid encoding the GFP-L-PduP$^{1-18}$ reporter was constructed by amplifying GFP with primers appending the linker and the PduP signal sequence. This amplicon was inserted into a pBAD33 vector backbone with the EcoRI restriction site. The other GFP reporters with C-terminal signal peptide fusions were similarly amplified and inserted into the same pBAD33 vector, except scarless golden gate assembly was used. The primers used are listed in Table 4.2.

Plasmid	**Description**
eyk056	*PduP$^{1-18}$-GFP*, pBAD, p15A, chloramphenicol^{R}
cmj012	*PduD$^{1-20}$-GFP*, pBAD, p15A, chloramphenicol^{R}
mfspl099	*EutC$^{1-20}$-GFP*, pBAD, p15A, chloramphenicol^{R}
mfspl140	*EutE$^{1-20}$-GFP*, pBAD, p15A, chloramphenicol^{R}
mfspl080	*GFP-L-PduP$^{1-18}$*, pBAD, p15A, chloramphenicol^{R}
mfspl129	*GFP-L-PduD$^{1-20}$*, pBAD, p15A, chloramphenicol^{R}
mfspl118	*GFP-L-EutC$^{1-20}$*, pBAD, p15A, chloramphenicol^{R}
mfspl119	*GFP-L-EutE$^{1-20}$*, pBAD, p15A, chloramphenicol^{R}
mfspl042	*VioA, VioB, VioE,* , pBAD, p15A, chloramphenicol^{R}
mfspl032	*PduP$^{1-18}$-VioA, PduP$^{1-18}$-VioB, PduP$^{1-18}$-VioE,* , pBAD, p15A, chloramphenicol^{R}
mfspl037	*PduP$^{1-18}$-VioA, VioB, VioE,* , pBAD, p15A, chloramphenicol^{R}
mfspl041	*VioA, PduP$^{1-18}$-VioB, VioE,* , pBAD, p15A, chloramphenicol^{R}
mfspl043	*VioA, VioB, PduP$^{1-18}$-VioE,* , pBAD, p15A, chloramphenicol^{R}
mfspl089	*PduP$^{1-18}$-VioA, VioB-L-PduP$^{1-18}$, PduP$^{1-18}$-L-VioE*, pBAD, p15A, chloramphenicol^{R}

Table 4.1 Plasmids used in this study

Primer	Description	Sequence
eykp567	PduP$^{1-18}$-VioA f2	cctgattcgcaccattcttagcgagcaattatctagaaaacattcttccgatatctgcattg
eykp568	PduP$^{1-18}$-VioA f1	tattgaattcggtctcacatgaatacttctgaactcgaaaccctgattcgcaccattcttag c
eykp569	VioA f	tattgaattcggtctcacatgaaacattcttccgatatctgcattg
eykp570	VioA r	tattctcgagggtctcagtaccttattacgcggcgatacgc
eykp572	PduP$^{1-18}$-VioB f2	cctgattcgcaccattcttagcgagcaatta tctaga agcattctggatttcccgc
eykp573	PduP$^{1-18}$-VioB f1	tattgaattcggtctcagtaccttaaagaggagaaaggtcatgaatacttctgaactc gaaaccctgattcgcaccattcttagc
eykp574	VioB f	tattgaattcggtctcagtaccttaaagaggagaaaggtcatgagcattctggatttc ccgc
eykp575	VioB r	tattctcgagggtctcactagtttattaggcctcgcggctc
mfsp036	VioB-L- PduP$^{1-18}$ r1	gcgaatcagggtttcgagttcagaagtattagaaccagaaccagaaccggcctcg cggctc

mfsp037	VioB-L- PduP$^{1-18}$ r2	tattctcgagggtctcactagtttattataattgctcgctaagaatggtgcgaatcaggg gtttcgagttc
eykp577	PduP$^{1-18}$-VioE f2	cctgattcgcaccattcttagcgagcaattatctagagagaaccgtgagccaccac
eykp578	PduP$^{1-18}$-VioE f1	tattgaattcggtctcactagtttaaagaggagaaaggtcatgaatacttctgaactcgaa accctgattcgcaccattcttagc
eykp579	VioE f	tattgaattcggtctcactagtttaaagaggagaaaggtcatggagaaccgtgagccacc ac
eykp580	VioE r	tattctcgagggtctcatttattagcgcttggccgcg
mfsp011	PduP$^{1-18}$-L-VioE f2	cctgattcgcaccattcttagcgagcaattaggttctggtagtggttctgagaaccgtgag ccaccac
mfsp023	GFP-L-PduP$^{1-18}$ pbad33 r1	gcgaatcagggtttcgagttcagaagtattagaaccagaaccagaacctttgtatagttc atccatgccatg
mfsp024	GFP-L-PduP$^{1-18}$ pbad33 r2	tattaagcttttataattgctcgctaagaatggtgcgaatcagggtttcgagttc
mfsp025.2	GFP f	tattgaattcatttattaaagaggagaaaggtcatgagtaaaggagaagaacttttcact gg
mfsp060	GFP-Linker f	tatt ggtctc a catg agtaaaggagaagaacttttcactgg
mfsp087	GFP-Linker r	tattggtctc a agaa ccagaaccagaacctttgtatag
mfsp085	EutC$^{1-20}$ f	tatt ggtctc a ttct atggatcaaaaacagattgaagaaattgtacg
mfsp086	EutC$^{1-20}$ r	tattggtctc a ttta tta gtcctgtcccattgacgc
mfsp088	EutE$^{1-20}$ f	tatt ggtctc a ttct atgaatcaacaggatattgaacaggtg
mfsp089	EutE$^{1-20}$ r	tattggtctc a ttta tta gctgtctttcattttcagcagtac
mfsp107	PduD$^{1-20}$ f	tatt ggtctc a ttct atggaaattaatgaaaaattgctgcgcc
mfsp108	PduD$^{1-20}$ r	tatt ggtctc a ttta tta cttcatatcgcggagtacgtcttc

Table 4.2 Primers used in this study

4.2.2 Violacein pathway expression and PDV measurement

Plasmids expressing VioA, VioB, and VioE were assembled by golden gate assembly. Constructs with pathway activity formed green colonies. To further test the formation of PDV, single colonies were used to inoculate 5 mL of LB medium and grown at 37°C for 17 hours. These cultures were diluted 1:1000 into 5 mL of fresh LB medium. After about 2.5 hours of growth, the cultures reached an OD (600 nm) of .4 and were induced for Vio enzyme production with .02% arabinose. After 5 hours cells were pelleted at 17000 xg for 1 minute. PDV was extracted by resuspending the cell pellet in 0.5 mL methanol and incubating at 80°C for 15 min with periodic vortexing. The insoluble material is pelleted by centrifugation at 17,000 xg for 1 minute and the supernatant is filtered prior to analysis by HPLC and UV/vis absorbance measurements. PDV has absorbance peaks at 424 and 610 nm (Figure 4.1).

4.2.3 Fluorescence microscopy

Bacteria were viewed using a Nikon Ni-U upright microscope with a 100x 1.45 n.a. plan apochromat objective. Images were captured using an Andor Clara-Lite digital camera and Nikon NIS Elements software. Fluorescence images were collected using a C-FL Endow GFP HYQ bandpass filter. All images were taken with 400 ms exposure and adjusted identically for contrast in Adobe Photoshop software.

4.2.4 MCP purification

5 mL cultures were grown in LB-Miller from a single colony for 24 hours at 30°C, then subcultured 1:1000 into 400 mL of NCE with 55 mM 1,2-PD and grown for 13-15 hours at 37°C. To express heterologous cargo, at OD600~0.4 the appropriate inducer was added at the concentrations indicated and the cultures were grown for a further 5.5 hours. MCPs were purified by sedimentation as previously described [81], with the following modification: in the place of BPER-II bacterial lysis solution (Thermo), a solution of 1% w/v octylthioglucoside (Santa Cruz Biotech) in 20 mM Tris (Fisher) pH 7.5 in water was used for cell lysis. They were stored at 4°C until analysis by SDS-PAGE or TEM. Total protein concentration of pure MCP sample was determined by Pierce BCA protein assay kit (Thermo Fisher).

4.2.5 SDS-PAGE and western blot

Denaturing protein electrophorese (SDS-PAGE) was performed on purified MCPs using a 12.5% polyacrylamide gel unless otherwise noted. A 130V potential was applied for 70 minutes. Sample loading was normalized by total protein concentration measured by BCA assay. Gels were either stained with Coomassie dye or further processed for western blot. Western blotting was done with a PVDF membrane according to standard protocols. Blotting was done with a mouse anti-GFP primary antibody (Sigma) 1:2000 dilution in 50 mM Tris 150 mM NaCl pH 7.6 with 0.05% Tween-20 (TBST) with 1% w/v dry milk. The secondary antibody was HRP-conjugated goat anti-mouse antibody (Thermo) diluted 1:1000 in TBST. Labeling was visualized with west-pico chemiluminescent substrate (Thermo) using a Bio-Rad ChemiDoc XRS+.

4.2.6 Flow cytometry

Cultures were grown as described above for fluorescence microscopy. Samples were diluted 1:40 in phosphate-buffered saline (PBS) supplemented with 2 g/L kanamycin (to halt translation) in 96-well plates. 10,000 events were collected for each sample on a Millipore Guava easyCyte 5HT instrument. Cells were distinguished from debris by gating on the forward and side scatter channels using the FlowJo software. Reported fluorescence values are the arithmetic mean of the geometric mean green fluorescence of three independent samples acquired on three different days. Error bars represent one standard deviation.

4.3 Results

4.3.1 Development of C-terminal signal peptide

A plasmid expressing VioA, VioB, and VioE of the violacein biosynthesis pathway was constructed to test the function of these enzymes in the cytosol of *S. enterica*. The production of all three enzymes was under the control of a pBAD reporter. The functional activity of the pathway expressed from this plasmid was confirmed through the observed production of green prodeoxyviolacein (PDV) pigment upon induction. To test the functionality of these enzymes with N-terminal $PduP^{1-18}$ signal peptide fusions, new plasmids were constructed with $PduP^{1-18}$ fused to VioA, VioB, VioE, or all three enzymes. By observing the presence or absence of green pigment, it was determined that VioB and VioE did not function when tagged at the N-terminus with $PduP^{1-18}$ (Figure 4.2). The activity of VioE was rescued by the addition of a six amino acid, alternating glycine-serine linker. However, VioB would not tolerate any tag on the N-terminus. Therefore, a C-terminal encapsulation tag was needed.

A plasmid expressing PduP$^{1-18}$ fused to the C-terminus of GFP was constructed to test encapsulation. The glycine-serine linker was added between the signal peptide and GFP (GFP-L-PduP$^{1-18}$). The flexible hydrophilic linker allows increased accessibility of the signal peptide to its binding sites. When coexpressed with Pdu MCPs in *S. enterica*, the c-terminally tagged GFP was localized to the MCPs and appeared as fluorescent puncta by fluorescence microscopy (Figure 4.2). This experiment was repeated with three other previously characterized signal sequence peptides (GFP-L-PduD$^{1-20}$, GFP-L-EutC$^{1-20}$, and GFP-L-EutE$^{1-20}$). All were found to localize GFP to Pdu MCPs (Figure 4.2). MCPs were purified from these strains and protein content was assayed by SDS-PAGE and α-GFP western blot. Additionally, MCPs were purified from a strain expressing GFP without a signal peptide fusion as a negative control. All lanes aside from the negative control have a GFP band by western blot, demonstrating that GFP is co-purified with Pdu MCPs (Figure 4.3a). The MCPs harvested from all strains have the normal profile of MCP proteins (Figure 4.3b)

The C-terminal PduP$^{1-18}$ encapsulation peptide was applied to VioB, and this configuration demonstrated pathway activity. Then, a plasmid expressing PduP$^{1-18}$-VioA, VioB-L- PduP$^{1-18}$, and PduP$^{1-18}$-L-VioE was constructed, and also produced PDV pigment. Hence the C-terminal tag provides flexibility to accommodate sensitive enzyme structures.

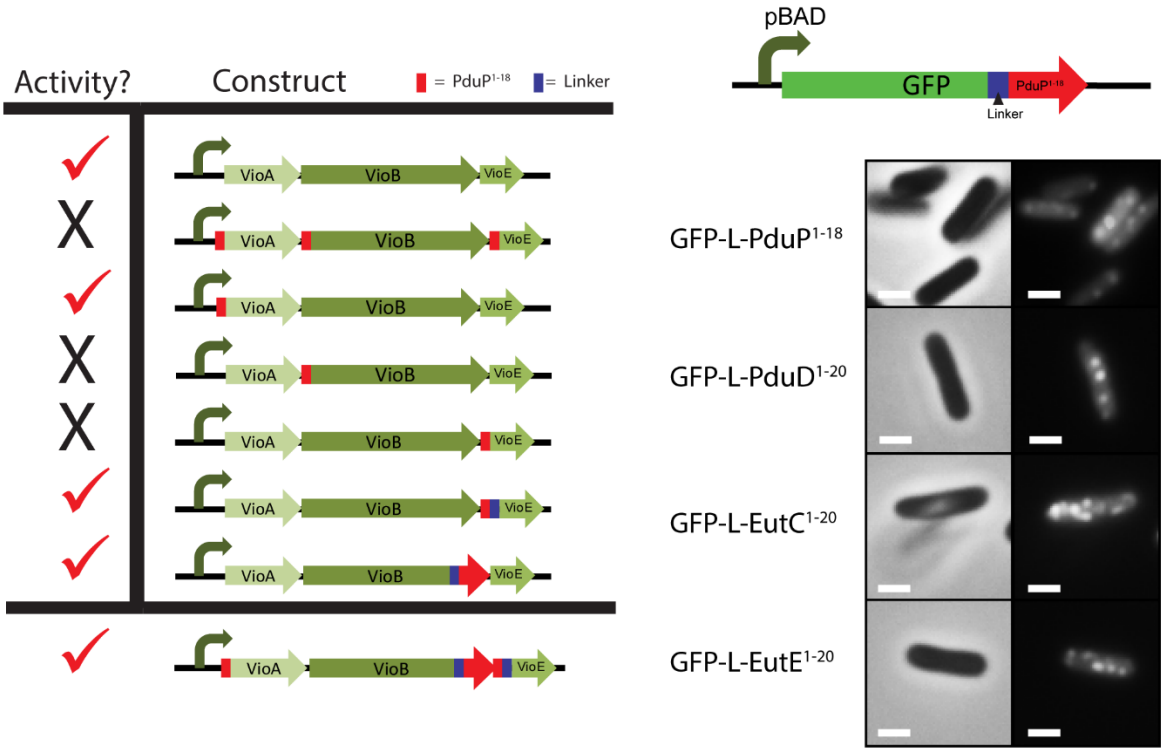


Figure 4.2 C-terminal signal sequence fusions. At left: diagram indicating several configurations of the plasmid expressing VioA, VioB, and VioE. Check marks indicate constructs that produce PDV pigment when expressed. Upper right: a cartoon depiction of the construct expressing GFP with a C-terminal linker and signal sequence fusion.

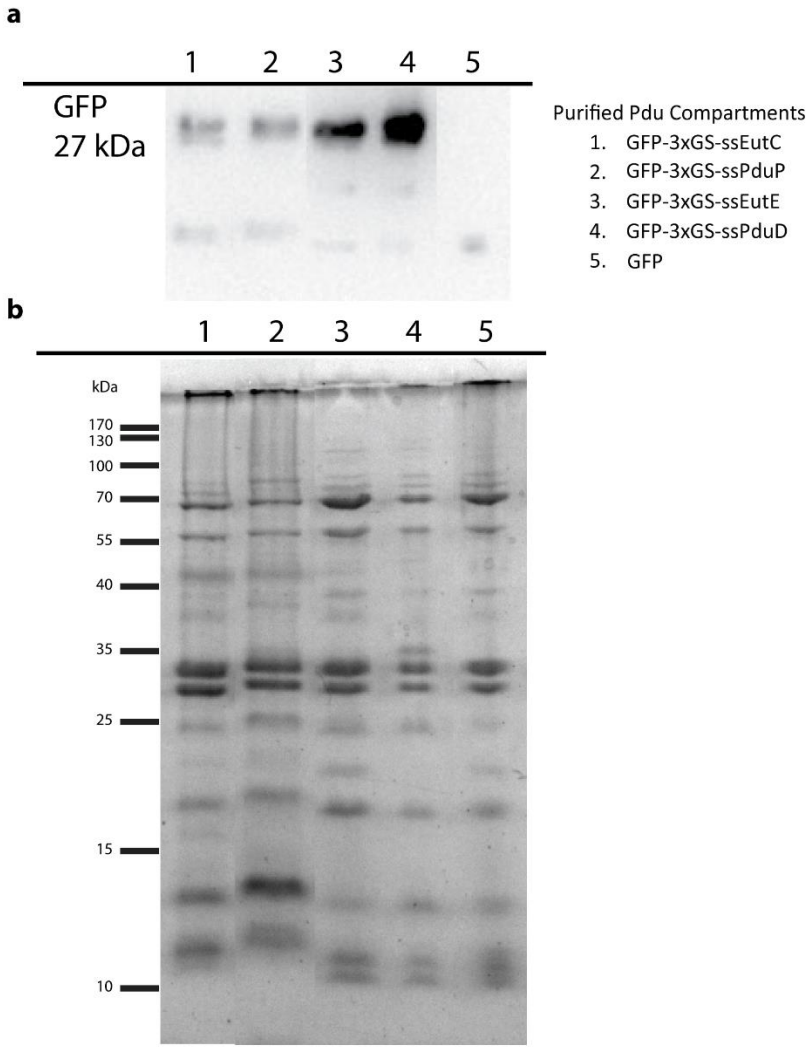


Figure 4.3 SDS-PAGE and western of purified MCPs with GFP reporters. a) western of purified Pdu MCP samples developed with mouse anti-GFP primary antibody. b) SDS-PAGE gel of the same samples depicted in (a). The normal band pattern formed by purified MCPs is visible in all lanes.

4.3.2 Effect of signal peptides on protein expression and encapsulation

Flow cytometry was used to measure the expression level of GFP in *S. enterica* when signal sequences are fused to the N- or C-terminus. Signal sequences from PduP, PduD, EutC, and EutE were all tested for expression. The results demonstrate that the use of different signal sequences and fusion configurations dramatically impact expression level (Figure 4.4a). For instance, relative to PduP$^{1-18}$-GFP, GFP-L- PduP$^{1-18}$ has 60 fold greater fluorescence, and EutC$^{1-30}$-GFP has 140 fold greater fluorescence. Further, when these reporter fusions are co-expressed with Pdu MCPs, the differences in expression level translate to different levels of loading within MCPs. This was measured by densitometry of αGFP western blots of purified MCPs (Figure 4b,c). This effect may be harnessed to control the ratios of custom cargos within synthetic MCP systems.

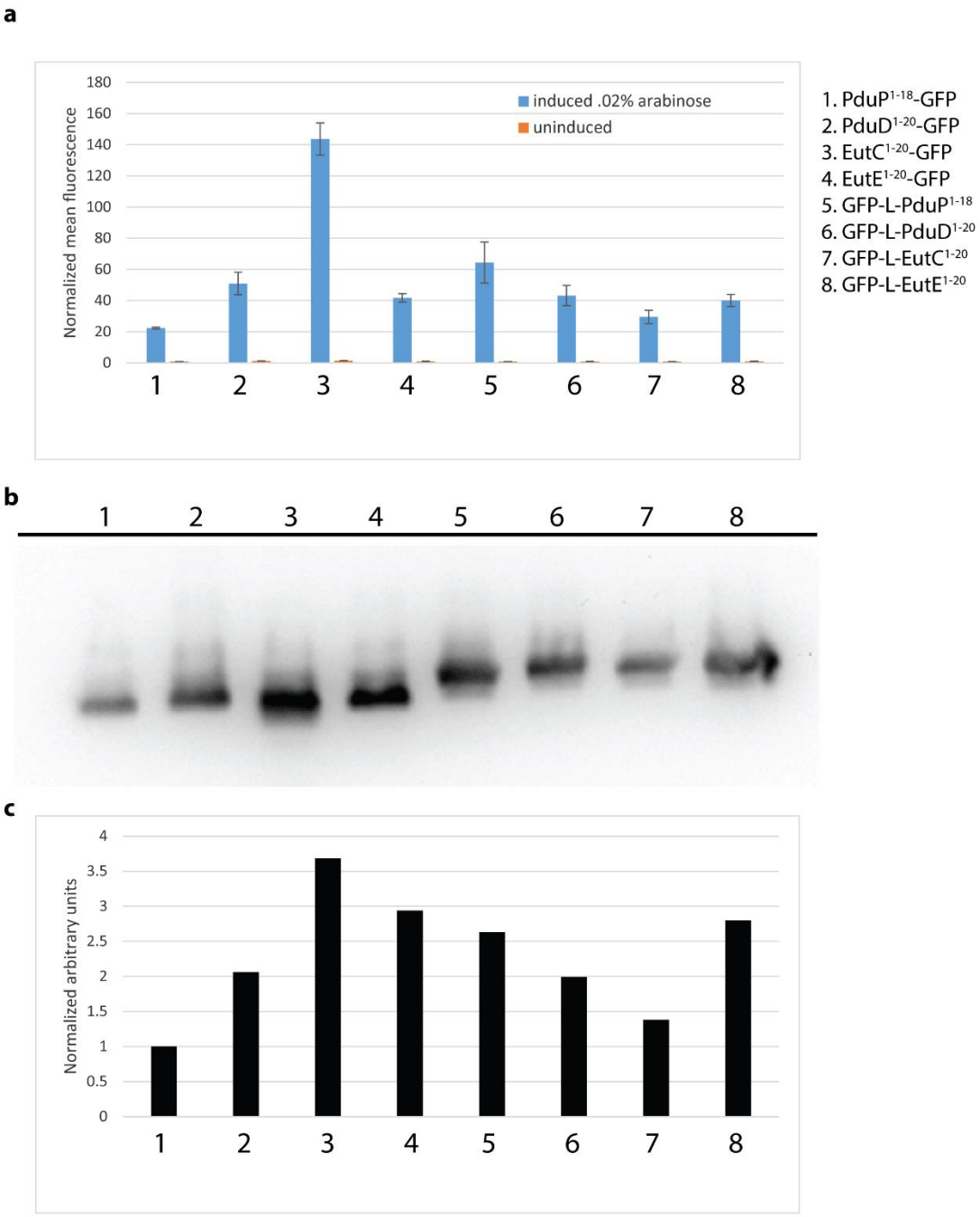


Figure 4.4 Flow cytometry and western densitometry of GFP reporters. a) Flow cytometry measuring GFP fluorescence in eight strains expressing GFP with different signal sequence tags and fusion configurations. The y-axis is the geometric mean fluorescence of a population of single cell measurements. The mean is normalized to the un-induced control measured for each strain. b) αGFP western blot c) densitometry measurements from the blot pictured in (b).

4.3.3 Identifying barriers and methods to improve encapsulation of multi-enzyme systems

By observing the native architecture of MCPs, we learn that there are two potential approaches for constructing an MCP with user-defined cargos. First, each individual enzyme may be fused with an encapsulation peptide (Figure 4.5a). Alternatively, a core of enzymes may be designed where one protein interacts with the shell and other proteins "tag along." In theory, an enzyme core could be constructed by non-covalent protein-protein interactions, using previously designed scaffolds (Figure 4.5b), or by fusing enzymes together with linkers (Figure 4.5c). So far, only the first approach with all proteins fused to encapsulation peptides has been employed to target heterologous cargo to the MCP. More exploration of the second approach may simplify the challenges of encapsulating many cargos at once. In any case, it is evident that the MCP structure is capable of sequestering many different cargos simultaneously. Thus, the prospects for encapsulating heterologous pathways of interest are promising.

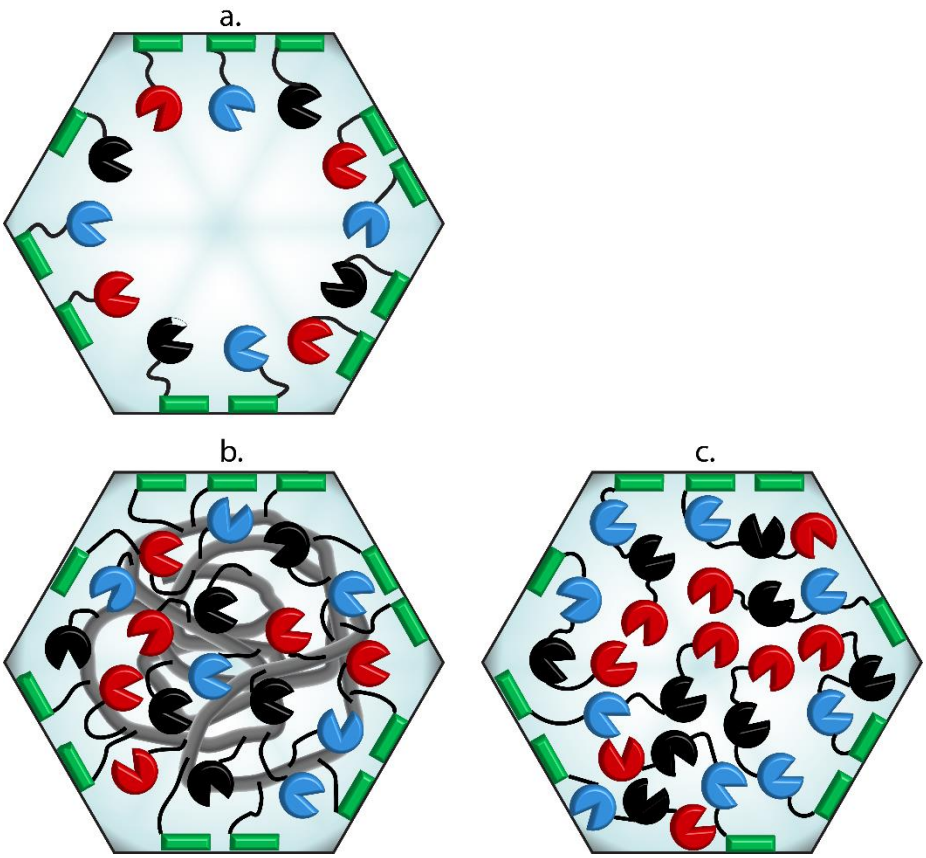


Figure 4.5: Cartoon model of encapsulation organization strategies. (a) all three enzymes (red, blue, and black) are fused to encapsulation peptides (green) that interact with shell proteins. (b) All enzymes bind to a core scaffold (grey) with controllable stoichiometry. Additional encapsulation peptides (green) bind to both the scaffold and shell proteins to induce shell formation around scaffold particle. (c) The enzymes (red, blue, and black) all bind to each other, one enzyme is also fused to the encapsulation peptide (green).

Before attempting to sequester a pathway within an MCP, it is important to consider whether encapsulation will be valuable given the particular properties of the pathway of interest.

MCPs function by imposing a diffusion barrier to restrict the escape or toxicity of intermediates, allow for the saturation of "slow" enzymes, and create a private pool of cofactors. This list of potential benefits helps to evaluate whether a given pathway will be significantly improved by sequestration within an MCP. To elaborate, a researcher should especially consider encapsulation if the pathway of interest involves: an enzyme with a large Km, an enzyme with promiscuous activity, a toxic intermediate, a volatile or reactive intermediate, or a requisite cofactor concentration different than that available in the cytosol. Quantitative modeling of native MCP systems has shown that different encapsulated pathways may be enhanced through different mechanisms depending on the kinetic properties of the enzymes involved [10,12]. Reaction diffusion models have also shown that the geometry of the protein shell and the density of enzyme organization can impact pathway performance, depending on again on the system kinetics [94]. Modeling is an invaluable tool for vetting target pathways and can also help to identify the best MCP design.

After discerning whether encapsulation could be worthwhile, one must assess the properties of each enzyme and determine whether they are amenable to encapsulation strategies. Current methods require a 20 amino acid encapsulation tag to be fused the N- or C- terminus of the cargo protein. In addition, the existence of a native mid-domain signal sequence was very recently discovered, so this may also be a viable way to include a signal sequence on a protein of interest [95]. As observed for VioB and VioE of the model violacein pathway in section 4.3.1, these fusions often reduce the activity of an enzyme if not carefully optimized. Fortunately, there are several ways to minimize any negative effects. Flexible linker additions between enzyme and signal sequence peptide enable control over signal sequence orientation. Additionally, the sequence of the signal peptide itself is flexible, with thousands of possible sequences able to mediate encapsulation [41]. Changing the peptide sequence can improve both protein expression and peptide-enzyme compatibility [39]. A high throughput search for the best signal peptide and linker combination from a random library may be necessary to guarantee the performance of a sensitive enzyme.

A pathway chosen for encapsulation must involve small molecules that are able to diffuse through the MCP shell. The pores found at the symmetric center of some shell protein hexamers are identified as the most likely conduits through which small molecules diffuse in and out of the organelle [26,27,96]. These pores are about 0.5-1.5 nm in diameter [26,97]. The size and charge of small molecules involved in the pathway of interest should be considered, however the limits on these parameters are not fully mapped for molecules passing through the shell. The only molecule shown to be unable to pass through the shell so far is polyphosphate, a large polymer [90]. As a result, the behavior of a particular small molecule with respect to the MCP shell cannot be predicted at this time. Fortunately, techniques for engineering the pores of the MCP shell have been developed and result in altered diffusion properties of the shell [27,77,96]. It is likely that mutating pore residues of shell proteins can optimize the performance of the encapsulated pathway.

Competition between multiple cargo proteins influences the success of sequestering an enzyme cascade within MCPs. It would be advantageous to control loading of each enzyme in the target pathway independently. However, this capability is not currently achievable with available techniques. All of the signal peptides in the Pdu and Eut systems that are characterized thus far appear to compete for the same binding sites within Pdu MCPs [39]. For instance, when co-expressing $PduP^{1-20}$-GFP and $PduD^{1-20}$-mCherry, increased loading of GFP resulted in a decrease in mCherry loading, and vice versa [39].

This effect could occur because all of the signal sequences have the same binding partners, or it could be a result of competition for limited volume within MCPs. Cargo competition introduces another parameter to tune for multi-enzyme encapsulation. The relative expression levels of cargo enzymes greatly impact the ratios of enzymes encapsulated. In addition, different signal peptides may also have different binding affinities, although this has not been evaluated. Therefore, it is important to tune the expression levels of cargo proteins with the knowledge that signal peptides greatly affect expression levels and encapsulation efficiency.

4.4 Discussion

We present a new functional configuration of signal sequence peptides for the encapsulation of heterologous cargo within MCPs. We find that attaching signal sequences with a linker to the C-terminus of GFP mediated robust encapsulation. As an example, we also demonstrated that the C-terminal encapsulation peptide allowed the VioB enzyme to remain active when N-terminal modifications were not tolerated.

Next it was determined that the choice of signal sequence and the configuration of the signal sequence both affect the expression level of reporter protein GFP. We found that high expression of cargo GFP caused increased loading within MCPs. The idea that cargo loading is influenced by cargo expression level was further investigated by another member of the Tullman-Ercek lab, Christopher Jakobson, leading to the finding that the relative timing of expression of cargo and shell is shown to control amounts of total cargo loaded [60]. Thus, optimizing the cargo loading is achieved by manipulating both the time between cargo and shell protein induction, and the overall expression level of the cargo. The best combination of these factors will vary depending on the dynamics of the expression system used. It is also limited by the optimal protein production window of a given strain. Induction outside of the log phase of growth usually results in poor protein production. Further simplifying the control of MCP shell and cargo expression is an important topic for future research.

Many advances in MCP engineering have brought the encapsulation of multi-enzyme cargos within reach. In particular, gains have been made in knowledge of MCP encapsulation peptides, the assembly pathway of carboxysomes, the effects of MCP pore chemistries on encapsulated pathway performance, and the mechanisms by which sequestration can improve pathway function. Many of the basic tools for encapsulation of multi-enzyme pathways are established, and in future work it will be important to attempt encapsulation for more diverse test cases. Tuning cargo loading and small molecule diffusion through the shell for the optimization of a pathway un-related to natively encapsulated metabolism is an important proof of concept. Still, a better understanding of the requirements for shell assembly and an expansion of cargo loading methods will simplify engineering efforts.

Chapter 5: CONCLUSIONS

Over the past decade, many advances were made in the understanding of native bacterial microcompartment structures and functions. Before this work began, some fundamental tools to control MCP formation were established. The above chapters describe new contributions to the

next generation of synthetic biology tools necessary to apply bacterial organelles to sequester heterologous pathways.

To start, we studied MCP stability build a baseline understanding of the structural limits of MCPs *in vitro*. We found that the C-termini of PduA and PduJ shell proteins are necessary for normal MCP shell formation, and a remarkable filament structure is formed when the C-terminal motif is disrupted. Then, we demonstrated that the Eut shell protein EutM can complement the deletion of PduA to form complete chimeric MCP structures. Further, we found that substituting EutM for PduA enhanced growth of the strain on 1,2-PD as the sole carbon source, indicating that chimeric MCPs can improve the function of an encapsulated pathway.

Next, we explored the role of shell protein pores in the function of the Pdu pathway in *S. enterica*. We found that a single point mutation, $PduA^{K37Q}$, causes the same improved growth phenotype as the complete *ΔpduA::eutM* substitution. We measured differences in the accumulation of metabolites in the growth medium when culturing these strains, and provide evidence that these changes are a result of altered metabolite diffusion across the shell of the MCP. The pore of PduA is amenable to mutations, enabling the generation of many possible diffusion phenotypes. This may be used to tailor synthetic organelle systems to benefit a heterologous pathway of interest.

Finally, we successfully developed a C-terminal tag system to target enzymes for encapsulation, better accommodating diverse enzyme structures. We demonstrated that the signal sequence and the configuration of the signal sequence fusion greatly affect the expression level of proteins. Further, we found that expression level is positively correlated with encapsulation level, revealing an important parameter for tuning MCP assembly. We summarize all of the factors impacting the encapsulation of multiple enzymes within the MCP to facilitate the application of this technology. With these techniques, it is possible to make use of MCPs for metabolic engineering.

REFERENCES

1. Bobik TA, Havemann GD, Busch RJ, Williams DS, Aldrich HC: The Propanediol Utilization (pdu) Operon ofSalmonella enterica Serovar Typhimurium LT2 Includes Genes Necessary for Formation of Polyhedral Organelles Involved in Coenzyme B12-Dependent 1, 2-Propanediol Degradation. *J. Bacteriol.* 1999, 181:5967–5975.

2. Kofoid E, Rappleye C, Stojiljkovic I, Roth J: The 17-gene ethanolamine (eut) operon of Salmonella typhimurium encodes five homologues of carboxysome shell proteins. *J. Bacteriol.* 1999, 181:5317–5329.

3. Petit E, LaTouf WG, Coppi MV, Warnick T a, Currie D, Romashko I, Deshpande S, Haas K, Alvelo-Maurosa JG, Wardman C, et al.: Involvement of a bacterial microcompartment in the metabolism of fucose and rhamnose by Clostridium phytofermentans. *PloS One* 2013, 8:e54337.

4. Erbilgin O, McDonald KL, Kerfeld CA: Characterization of a Planctomycetal Organelle: a Novel Bacterial Microcompartment for the Aerobic Degradation of Plant Saccharides. *Appl. Environ. Microbiol.* 2014, 80:2193–2205.

5. Axen SD, Erbilgin O, Kerfeld CA: A Taxonomy of Bacterial Microcompartment Loci Constructed by a Novel Scoring Method. *PLoS Comput. Biol.* 2014, 10:e1003898.

6. Shively JM, Ball F, Brown DH, Saunders RE: Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of Thiobacillus neapolitanus. *Science* 1973, 182:584–586.

7. Codd GA, Stewart WDP: Polyhedral bodies and ribulose 1,5-diphosphate carboxylase of the blue-green alga Anabaena cylindrica. *Planta* 1976, 130:323–326.

8. So AK-C, Espie GS, Williams EB, Shively JM, Heinhorst S, Cannon GC: A novel evolutionary lineage of carbonic anhydrase (epsilon class) is a component of the carboxysome shell. *J. Bacteriol.* 2004, 186:623–630.

9. Penrod JT, Roth JR: Conserving a Volatile Metabolite : a Role for Carboxysome-Like Organelles in Salmonella enterica. *J. Bacteriol.* 2006, 188:2865–2874.

10. Jakobson CM, Slininger MF, Tullman-Ercek D, Mangan NM: A Systems-Level Model Reveals That 1,2-Propanediol Utilization Microcompartments Enhance Pathway Flux Through Intermediate Sequestration. *bioRxiv* 2016, doi:10.1101/069542.

11. Sampson EM, Bobik T a: Microcompartments for B12-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate. *J. Bacteriol.* 2008, 190:2966–71.

12. Mangan NM, Brenner MP: Systems analysis of the CO_2 concentrating mechanism in cyanobacteria [Internet]. *eLife* 2014, 3.

13. Cheng S, Fan C, Sinha S, Bobik T a: The PduQ enzyme is an alcohol dehydrogenase used to recycle NAD+ internally within the Pdu microcompartment of Salmonella enterica. *PloS One* 2012, 7:e47144.

14. Huseby DL, Roth JR: Evidence that a Metabolic Microcompartment Contains and Recycles Private Cofactor Pools. *J. Bacteriol.* 2013, 195:2864–2879.

15. Cheng S, Bobik TA: Characterization of the PduS Cobalamin Reductase of Salmonella enterica and Its Role in the Pdu Microcompartment. *J. Bacteriol.* 2010, 192:5071–5080.

16. Kerfeld CA, Heinhorst S, Cannon GC: Bacterial Microcompartments. *Annu. Rev. Microbiol.* 2010, 64:391–408.

17. Thiennimitr P, Winter SE, Winter MG, Xavier MN, Tolstikov V, Huseby DL, Sterzenbach T, Tsolis RM, Roth JR, Bäumler AJ: Intestinal inflammation allows Salmonella to use ethanolamine to compete with the microbiota. *Proc. Natl. Acad. Sci.* 2011, 108:17480–17485.

18. Rivera-Chávez F, Bäumler A: The pyromaniac inside you: Salmonella metabolism in the host gut. *Annu. Rev. Microbiol.* 2015, 69:31–48.

19. Jakobson CM, Tullman-Ercek D: Dumpster Diving in the Gut: Bacterial Microcompartments as Part of a Host-Associated Lifestyle. *PLOS Pathog.* 2016, 12:e1005558.

20. Liu Y, Jorda J, Yeates TO, Bobik TA: The PduL Phosphotransacylase Is Used To Recycle Coenzyme A within the Pdu Microcompartment. *J. Bacteriol.* 2015, 197:2392–2399.

21. Havemann GD, Bobik TA: Protein Content of Polyhedral Organelles Involved in Coenzyme B12 - Dependent Degradation of 1,2-Propanediol in Salmonella enterica Serovar Typhimurium. *J. Bacteriol.* 2003, 185:5086–5095.

22. Yeates TO, Crowley CS, Tanaka S: Bacterial microcompartment organelles: protein shell structure and evolution. *Annu. Rev. Biophys.* 2010, 39:185–205.

23. Takenoya M, Nikolakakis K, Sagermann M: Crystallographic insights into the pore structures and mechanisms of the EutL and EutM shell proteins of the ethanolamine-utilizing microcompartment of Escherichia coli. *J. Bacteriol.* 2010, 192:6056–63.

24. Cheng S, Sinha S, Fan C, Liu Y, Bobik T a: Genetic analysis of the protein shell of the microcompartments involved in coenzyme B12-dependent 1,2-propanediol degradation by Salmonella. *J. Bacteriol.* 2011, 193:1385–92.

25. Pang A, Liang M, Prentice MB, Pickersgill RW: Substrate channels revealed in the trimeric Lactobacillus reuteri bacterial microcompartment shell protein PduB. *Acta Crystallogr. D Biol. Crystallogr.* 2012, 68:1642–52.

26. Crowley CS, Cascio D, Sawaya MR, Kopstein JS, Bobik T a, Yeates TO: Structural insight into the mechanisms of transport across the Salmonella enterica Pdu microcompartment shell. *J. Biol. Chem.* 2010, 285:37838–46.

27. Chowdhury C, Chun S, Pang A, Sawaya MR, Sinha S, Yeates TO, Bobik TA: Selective molecular transport through the protein shell of a bacterial microcompartment organelle. *Proc. Natl. Acad. Sci.* 2015, 112:2990–2995.

28. Glasgow JE, Asensio MA, Jakobson CM, Francis MB, Tullman-Ercek D: Influence of Electrostatics on Small Molecule Flux through a Protein Nanoreactor. *ACS Synth. Biol.* 2015, 4:1011–1019.

29. Lassila JK, Bernstein SL, Kinney JN, Axen SD, Kerfeld CA: Assembly of Robust Bacterial Microcompartment Shells Using Building Blocks from an Organelle of Unknown Function. *J. Mol. Biol.* 2014, 426:2217–2228.

30. Tanaka S, Kerfeld C a, Sawaya MR, Cai F, Heinhorst S, Cannon GC, Yeates TO: Atomic-level models of the bacterial carboxysome shell. *Science* 2008, 319:1083–6.

31. Sinha S, Cheng S, Sung YW, McNamara DE, Sawaya MR, Yeates TO, Bobik T a: Alanine scanning mutagenesis identifies an asparagine-arginine-lysine triad essential to assembly of the shell of the pdu microcompartment. *J. Mol. Biol.* 2014, 426:2328–45.

32. Fan C, Cheng S, Liu Y, Escobar CM, Crowley CS, Jefferson RE, Yeates TO, Bobik T a: Short N-terminal sequences package proteins into bacterial microcompartments. *Proc. Natl. Acad. Sci. U. S. A.* 2010, 107:7509–14.

33. Fan C, Bobik T a: The N-terminal region of the medium subunit (PduD) packages adenosylcobalamin-dependent diol dehydratase (PduCDE) into the Pdu microcompartment. *J. Bacteriol.* 2011, 193:5623–8.

34. Lawrence AD, Frank S, Newnham S, Lee MJ, Brown IR, Xue W-F, Rowe ML, Mulvihill DP, Prentice MB, Howard MJ, et al.: Solution Structure of a Bacterial Microcompartment Targeting Peptide and Its Application in the Construction of an Ethanol Bioreactor. *ACS Synth. Biol.* 2014, 3:454–465.

35. Lehman BP, Chowdhury C, Bobik TA: The N-terminus of the PduB protein binds the protein shell of the Pdu microcompartment to its enzymatic core. *J. Bacteriol.* 2017, doi:10.1128/JB.00785-16.

36. Parsons JB, Lawrence AD, McLean KJ, Munro AW, Rigby SEJ, Warren MJ: Characterisation of PduS, the pdu Metabolosome Corrin Reductase, and Evidence of Substructural Organisation within the Bacterial Microcompartment. *PLoS ONE* 2010, 5:e14009.

37. Choudhary S, Quin MB, Sanders M a, Johnson ET, Schmidt-Dannert C: Engineered protein nano-compartments for targeted enzyme localization. *PloS One* 2012, 7:e33342.

38. Jorda J, Lopez D, Wheatley NM, Yeates TO: Using comparative genomics to uncover new kinds of protein-based metabolic organelles in bacteria. *Protein Sci.* 2013, 22:179–195.

39. Jakobson CM, Kim EY, Slininger MF, Chien A, Tullman-Ercek D: Localization of Proteins to the 1,2-Propanediol Utilization Microcompartment by Non-native Signal Sequences Is Mediated by a Common Hydrophobic Motif. *J. Biol. Chem.* 2015, 290:24519–24533.

40. Aussignargues C, Paasch BC, Gonzalez-Esquer R, Erbilgin O, Kerfeld CA: Bacterial microcompartment assembly: The key role of encapsulation peptides. *Commun. Integr. Biol.* 2015, 8:e1039755.

41. Jakobson CM, Slininger Lee MF, Tullman-Ercek D: *De Novo* Design of Signal Sequences to Localize Cargo to the 1,2-Propanediol Utilization Microcompartment: Design of signal sequences [Internet]. *Protein Sci.* 2017, doi:10.1002/pro.3144.

42. Cameron JC, Wilson SC, Bernstein SL, Kerfeld CA: Biogenesis of a Bacterial Organelle: The Carboxysome Assembly Pathway. *Cell* 2013, 155:1131–1140.

43. Kinney JN, Salmeen A, Cai F, Kerfeld C a: Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly. *J. Biol. Chem.* 2012, 287:17729–36.

44. Long BM, Badger MR, Whitney SM, Price GD: Analysis of Carboxysomes from Synechococcus PCC7942 Reveals Multiple Rubisco Complexes with Carboxysomal Proteins CcmM and CcaA. *J. Biol. Chem.* 2007, 282:29323–29335.

45. Cot SS-W, So AK-C, Espie GS: A Multiprotein Bicarbonate Dehydration Complex Essential to Carboxysome Function in Cyanobacteria. *J. Bacteriol.* 2008, 190:936–945.

46. Long BM, Tucker L, Badger MR, Price GD: Functional Cyanobacterial -Carboxysomes Have an Absolute Requirement for Both Long and Short Forms of the CcmM Protein. *PLANT Physiol.* 2010, 153:285–293.

47. Kaneko Y, Danev R, Nagayama K, Nakamoto H: Intact Carboxysomes in a Cyanobacterial Cell Visualized by Hilbert Differential Contrast Transmission Electron Microscopy. *J. Bacteriol.* 2006, 188:805–808.

48. Iancu CV, Ding HJ, Morris DM, Dias DP, Gonzales AD, Martino A, Jensen GJ: The Structure of Isolated Synechococcus Strain WH8102 Carboxysomes as Revealed by Electron Cryotomography. *J. Mol. Biol.* 2007, 372:764–773.

49. Chen AH, Robinson-Mosher A, Savage DF, Silver P a, Polka JK: The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo. *PloS One* 2013, 8:e76127.

50. Cai F, Bernstein SL, Wilson SC, Kerfeld CA: Production and Characterization of Synthetic Carboxysome Shells with incorporated Luminal Proteins. *Plant Physiol.* 2016, doi:10.1104/pp.15.01822.

51. Niederhuber MJ, Lambert TJ, Yapp C, Silver PA, Polka JK: Super-resolution microscopy of the β-carboxysome reveals a homogenous matrix. *bioRxiv* 2017, doi:https://doi.org/10.1101/086090.

52. Chen P, Andersson DI, Roth JR: The control region of the pdu/cob regulon in Salmonella typhimurium. *J. Bacteriol.* 1994, 176:5474–82.

53. Bobik TA, Ailion M, Roth JR: A single regulatory gene integrates control of vitamin B12 synthesis and propanediol degradation. *J. Bacteriol.* 1992, 174:2253–2266.

54. Rondon MR, Escalante-Semerena JC: The poc locus is required for 1,2-propanediol-dependent transcription of the cobalamin biosynthetic (cob) and propanediol utilization (pdu) genes of Salmonella typhimurium. *J. Bacteriol.* 1992, 174:2267–2272.

55. Chen P, Ailion M, Bobik T, Stormo G, Roth J: Five promoters integrate control of the cob/pdu regulon in Salmonella typhimurium. *J. Bacteriol.* 1995, 177:5401–5410.

56. Rondon MR, Escalante-Semerena JC: In vitro analysis of the interactions between the PocR regulatory protein and the promoter region of the cobalamin biosynthetic (cob) operon of Salmonella typhimurium LT2. *J. Bacteriol.* 1996, 178:2196–2203.

57. Ailion M, Bobik TA, Roth JR: Two global regulatory systems (Crp and Arc) control the cobalamin/propanediol regulon of Salmonella typhimurium. *J. Bacteriol.* 1993, 175:7200–7208.

58. Roof DM, Roth JR: Autogenous regulation of ethanolamine utilization by a transcriptional activator of the eut operon in Salmonella typhimurium. *J. Bacteriol.* 1992, 174:6634–6643.

59. Sheppard DE, Roth JR: A rationale for autoinduction of a transcriptional activator: ethanolamine ammonia-lyase (EutBC) and the operon activator (EutR) compete for adenosyl-cobalamin in Salmonella typhimurium. *J. Bacteriol.* 1994, 176:1287–1296.

60. Jakobson CM, Chen Y, Slininger MF, Valdivia E, Kim EY, Tullman-Ercek D: Tuning the Catalytic Activity of Subcellular Nanoreactors. *J. Mol. Biol.* 2016, 428:2989–2996.

61. Kim EY, Jakobson CM, Tullman-Ercek D: Engineering Transcriptional Regulation to Control Pdu Microcompartment Formation. *PLoS ONE* 2014, 9:e113814.

62. Kim EY, Tullman-Ercek D: A rapid flow cytometry assay for the relative quantification of protein encapsulation into bacterial microcompartments. *Biotechnol. J.* 2014, 9:348–354.

63. Farrell CM, Grossman AD, Sauer RT: Cytoplasmic degradation of ssrA-tagged proteins: Cellular degradation of GFP-ssrA. *Mol. Microbiol.* 2005, 57:1750–1761.

64. Gottesman S, Roche E, Zhou Y, Sauer RT: The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev.* 1998, 12:1338–1347.

65. Karzai AW, Roche ED, Sauer RT: The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue. *Nat. Struct. Biol.* 2000, 7:449–455.

66. Parsons JB, Frank S, Bhella D, Liang M, Prentice MB, Mulvihill DP, Warren MJ: Synthesis of Empty Bacterial Microcompartments, Directed Organelle Protein Incorporation, and Evidence of Filament-Associated Organelle Movement. *Mol. Cell* 2010, 38:305–315.

67. Parsons JB, Dinesh SD, Deery E, Leech HK, Brindley A a, Heldt D, Frank S, Smales CM, Lünsdorf H, Rambach A, et al.: Biochemical and structural insights into bacterial organelle form and biogenesis. *J. Biol. Chem.* 2008, 283:14366–75.

68. Noël CR, Cai F, Kerfeld CA: Purification and Characterization of Protein Nanotubes Assembled from a Single Bacterial Microcompartment Shell Subunit. *Adv. Mater. Interfaces* 2016, 3:1500295.

69. Bonacci W, Teng PK, Afonso B, Niederholtmeyer H, Grob P, Silver PA, Savage DF: Modularity of a carbon-fixing protein organelle. *Proc. Natl. Acad. Sci.* 2012, 109:478–483.

70. Pang A, Frank S, Brown I, Warren MJ, Pickersgill RW: Structural Insights into Higher Order Assembly and Function of the Bacterial Microcompartment Protein PduA. *J. Biol. Chem.* 2014, 289:22377–22384.

71. Crowley CS, Sawaya MR, Bobik T a, Yeates TO: Structure of the PduU shell protein from the Pdu microcompartment of Salmonella. *Struct. Lond. Engl. 1993* 2008, 16:1324–32.

72. Thompson MC, Crowley CS, Kopstein J, Bobik TA, Yeates TO: Structure of a bacterial microcompartment shell protein bound to a cobalamin cofactor. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* 2014, 70:1584–1590.

73. Thompson MC, Cascio D, Leibly DJ, Yeates TO: An allosteric model for control of pore opening by substrate binding in the EutL microcompartment shell protein: Pore Regulation by Substrate Binding. *Protein Sci.* 2015, 24:956–975.

74. Wheatley NM, Gidaniyan SD, Liu Y, Cascio D, Yeates TO: Bacterial microcompartment shells of diverse functional types possess pentameric vertex proteins. *Protein Sci.* 2013, 22:660–665.

75. Heldt D, Frank S, Seyedarabi A, Ladikis D, Parsons JB, Warren MJ, Pickersgill RW: Structure of a trimeric bacterial microcompartment shell protein, EtuB, associated with ethanol utilization in Clostridium kluyveri. *Biochem. J.* 2009, 423:199–207.

76. Dou Z, Heinhorst S, Williams EB, Murin CD, Shively JM, Cannon GC: CO2 Fixation Kinetics of Halothiobacillus neapolitanus Mutant Carboxysomes Lacking Carbonic Anhydrase Suggest the Shell Acts as a Diffusional Barrier for CO2. *J. Biol. Chem.* 2008, 283:10377–10384.

77. Cai F, Sutter M, Bernstein SL, Kinney JN, Kerfeld C a: Engineering Bacterial Microcompartment Shells: Chimeric Shell Proteins and Chimeric Carboxysome Shells. *ACS Synth. Biol.* 2014, 4:444–453.

78. Sturms R, Streauslin NA, Cheng S, Bobik TA: In Salmonella enterica, Ethanolamine Utilization Is Repressed by 1,2-Propanediol To Prevent Detrimental Mixing of Components of Two Different Bacterial Microcompartments. *J. Bacteriol.* 2015, 197:2412–2421.

79. Engler C, Gruetzner R, Kandzia R, Marillonnet S: Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. *PLoS ONE* 2009, 4:e5553.

80. Datta S, Costantino N, Court DL: A set of recombineering plasmids for gram-negative bacteria. *Gene* 2006, 379:109–115.

81. Sinha S, Cheng S, Fan C, Bobik T a: The PduM protein is a structural component of the microcompartments involved in coenzyme B(12)-dependent 1,2-propanediol degradation by Salmonella enterica. *J. Bacteriol.* 2012, 194:1912–8.

82. Fan C, Cheng S, Sinha S, Bobik TA: Interactions between the termini of lumen enzymes and shell proteins mediate enzyme encapsulation into bacterial microcompartments. *Proc. Natl. Acad. Sci.* 2012, 109:14995–15000.

83. Heffernan R, Dehzangi A, Lyons J, Paliwal K, Sharma A, Wang J, Sattar A, Zhou Y, Yang Y: Highly accurate sequence-based prediction of half-sphere exposures of amino acid residues in proteins. *Bioinformatics* 2016, 32:843–849.

84. Havemann GD, Sampson EM, Bobik TA: PduA Is a Shell Protein of Polyhedral Organelles Involved in Coenzyme B12-Dependent Degradation of 1,2-Propanediol in Salmonella enterica Serovar Typhimurium LT2. *J. Bacteriol.* 2002, 184:1253–1261.

85. Savage DF, Afonso B, Chen AH, Silver PA: Spatially Ordered Dynamics of the Bacterial Carbon Fixation Machinery. *Science* 2010, 327:1258–1261.

86. Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, et al.: CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* 2006, 7:R100.

87. Vernizzi G, Sknepnek R, Olvera de la Cruz M: Platonic and Archimedean geometries in multicomponent elastic membranes. *Proc. Natl. Acad. Sci. U. S. A.* 2011, 108:4292–6.

88. Mayrose I: Comparison of Site-Specific Rate-Inference Methods for Protein Sequences: Empirical Bayesian Methods Are Superior. *Mol. Biol. Evol.* 2004, 21:1781–1791.

89. Welte W, Nestel U, Wacker T, Diederichs K: Structure and function of the porin channel. *Kidney Int.* 1995, 48:930–940.

90. Liang M, Frank S, Lünsdorf H, Warren MJ, Prentice MB: Bacterial microcompartment-directed polyphosphate kinase promotes stable polyphosphate accumulation in E. coli. *Biotechnol. J.* 2017, doi:10.1002/biot.201600415.

91. Balibar CJ, Walsh CT: In vitro biosynthesis of violacein from L-tryptophan by the enzymes VioA-E from Chromobacterium violaceum. *Biochemistry (Mosc.)* 2006, 45:15444–57.

92. Rodrigues AL, Trachtmann N, Becker J, Lohanatha AF, Blotenberg J, Bolten CJ, Korneli C, de Souza Lima AO, Porto LM, Sprenger G a, et al.: Systems metabolic engineering of Escherichia coli for production of the antitumor drugs violacein and deoxyviolacein. *Metab. Eng.* 2013, 20:29–41.

93. Sánchez C, Braña AF, Méndez C, Salas J a: Reevaluation of the violacein biosynthetic pathway and its relationship to indolocarbazole biosynthesis. *Chembiochem Eur. J. Chem. Biol.* 2006, 7:1231–40.

94. Hinzpeter F, Gerland U, Tostevin F: Optimal Compartmentalization Strategies for Metabolic Microcompartments. *Biophys. J.* 2017, 112:767–779.

95. Zarzycki J, Sutter M, Cortina NS, Erb TJ, Kerfeld CA: In Vitro Characterization and Concerted Function of Three Core Enzymes of a Glycyl Radical Enzyme - Associated Bacterial Microcompartment. *Sci. Rep.* 2017, 7:42757.

96. Slininger Lee MF, Jakobson CM, Tullman-Ercek D: Engineered shell proteins confer improved encapsulated pathway behavior in a bacterial microcompartment. *bioRxiv* 2017, doi:10.1101/106716.

97. Klein MG, Zwart P, Bagby SC, Cai F, Chisholm SW, Heinhorst S, Cannon GC, Kerfeld C a: Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport. *J. Mol. Biol.* 2009, 392:319–33.