

RESEARCH ARTICLE

# The Structural Basis of Coenzyme A Recycling in a Bacterial Organelle

Onur Erbilgin<sup>1</sup>✉, Markus Sutter<sup>2,3</sup>✉, Cheryl A. Kerfeld<sup>1,2,3\*</sup>

**1** Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, California, United States of America, **2** MSU-DOE Plant Research Laboratory and Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan, United States of America, **3** Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, California, United States of America

✉ These authors contributed equally to this work.

\* [ckerfeld@lbl.gov](mailto:ckerfeld@lbl.gov)



**OPEN ACCESS**

**Citation:** Erbilgin O, Sutter M, Kerfeld CA (2016) The Structural Basis of Coenzyme A Recycling in a Bacterial Organelle. *PLoS Biol* 14(3): e1002399. doi:10.1371/journal.pbio.1002399

**Academic Editor:** Gregory A. Petsko, Brandeis University, UNITED STATES

**Received:** August 25, 2015

**Accepted:** February 3, 2016

**Published:** March 9, 2016

**Copyright:** © 2016 Erbilgin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** PDB files are available from the Protein Data Bank under accession codes 5CUO and 5CUP.

**Funding:** This work was funded by the National Institutes of Health, National Institute of Allergy and Infectious Diseases (NIAID) (<http://www.niaid.nih.gov/>) grant 1R01AI114975-01. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

**Abbreviations:** Ack, acetate kinase; BMC, Bacterial Microcompartment; EP, encapsulation peptide; EUT,

## Abstract

Bacterial Microcompartments (BMCs) are proteinaceous organelles that encapsulate critical segments of autotrophic and heterotrophic metabolic pathways; they are functionally diverse and are found across 23 different phyla. The majority of catabolic BMCs (metabolosomes) compartmentalize a common core of enzymes to metabolize compounds via a toxic and/or volatile aldehyde intermediate. The core enzyme phosphotransacylase (PTAC) recycles Coenzyme A and generates an acyl phosphate that can serve as an energy source. The PTAC predominantly associated with metabolosomes (PduL) has no sequence homology to the PTAC ubiquitous among fermentative bacteria (Pta). Here, we report two high-resolution PduL crystal structures with bound substrates. The PduL fold is unrelated to that of Pta; it contains a dimetal active site involved in a catalytic mechanism distinct from that of the housekeeping PTAC. Accordingly, PduL and Pta exemplify functional, but not structural, convergent evolution. The PduL structure, in the context of the catalytic core, completes our understanding of the structural basis of cofactor recycling in the metabolosome lumen.

## Author Summary

In metabolism, molecules with “high-energy” bonds (e.g., ATP and Acetyl~CoA) are critical for both catabolic and anabolic processes. Accordingly, the retention of these bonds during biochemical transformations is incredibly important. The phosphotransacylase (Pta) enzyme catalyzes the conversion between acyl-CoA and acyl-phosphate. This reaction directly links an acyl-CoA with ATP generation via substrate-level phosphorylation, producing short-chain fatty acids (e.g., acetate), and also provides a path for short-chain fatty acids to enter central metabolism. Due to this key function, Pta is conserved across the bacterial kingdom. Recently, a new type of phosphotransacylase was described that shares no evolutionary relation to Pta. This enzyme, PduL, is exclusively associated with organelles called bacterial microcompartments, which are used to catabolize various compounds. Not only does PduL facilitate substrate level phosphorylation, but it also is critical

ethanolamine-utilizing;  $M_n$ , number average;  $M_w$ , weighted average; PDU, propanediol-utilizing; Pta, phosphotransacylase; PTAC, phosphotransacylase; RuBisCO, Ribulose Bisphosphate Carboxylase/Oxygenase; SEC-MALS, multiangle light scattering.

for cofactor recycling within, and product efflux from, the organelle. We solved the structure of this convergent phosphotransacylase and show that it is completely structurally different from Pta, including its active site architecture. We also discuss features of the protein important to its packaging in the organelle.

## Introduction

Bacterial Microcompartments (BMCs) are organelles that encapsulate enzymes for sequential biochemical reactions within a protein shell [1–4]. The shell is typically composed of three types of protein subunits, which form either hexagonal (BMC-H and BMC-T) or pentagonal (BMC-P) tiles that assemble into a polyhedral shell. The facets of the shell are composed primarily of hexamers that are typically perforated by pores lined with highly conserved, polar residues [1] that presumably function as the conduits for metabolites into and out of the shell [5,6].

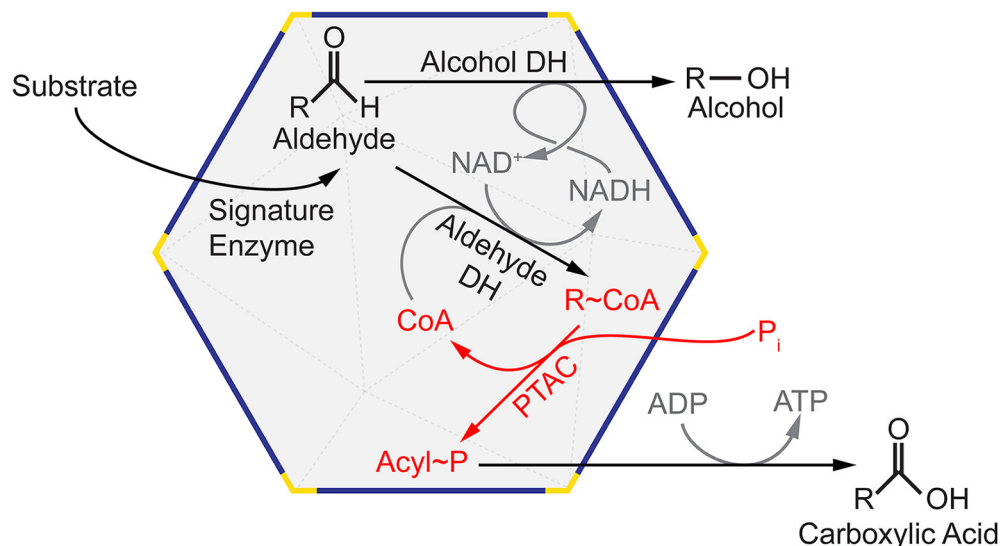
The vitamin B<sub>12</sub>-dependent propanediol-utilizing (PDU) BMC was one of the first functionally characterized catabolic BMCs [7]; subsequently, other types have been implicated in the degradation of ethanolamine, choline, fucose, rhamnose, and ethanol, all of which produce different aldehyde intermediates (Table 1). More recently, bioinformatic studies have demonstrated the widespread distribution of BMCs among diverse bacterial phyla [2,8,9] and grouped them into 23 different functional types [2]. The reactions carried out in the majority of catabolic BMCs (also known as metabolosomes) fit a generalized biochemical paradigm for the oxidation of aldehydes (Fig 1) [2]. This involves a BMC-encapsulated signature enzyme that generates a toxic and/or volatile aldehyde that the BMC shell sequesters from the cytosol [1]. The aldehyde is subsequently converted into an acyl-CoA by aldehyde dehydrogenase, which uses NAD<sup>+</sup> and CoA as cofactors [10,11]. These two cofactors are relatively large, and their diffusion across the protein shell is thought to be restricted, necessitating their regeneration within the BMC lumen [3,12,13]. NAD<sup>+</sup> is recycled via alcohol dehydrogenase [13], and CoA is recycled via phosphotransacylase (PTAC) [3,12] (Fig 1). The final product of the BMC, an acyl-phosphate, can then be used to generate ATP via acyl kinase, or revert back to acyl-CoA

**Table 1. Characterized and predicted catabolic BMC (metabolosome) types that represent the aldehyde-degrading paradigm (for definition of types see Kerfeld and Erbilgin [1]).**

Name	PTAC Type	Sequestered Aldehyde
PDU*	PduL	propionaldehyde
EUT1	PTA_PTB	acetaldehyde
EUT2	PduL	acetaldehyde
ETU	None	acetaldehyde
GRM1/CUT	PduL	acetaldehyde
GRM2	PduL	acetaldehyde
GRM3*,4	PduL	propionaldehyde
GRM5/GRP	PduL	propionaldehyde
PVM*	PduL	lactaldehyde
RMM1,2	None	unknown
SPU	PduL	unknown

\* PduL from these functional types of metabolosomes were purified in this study.

doi:10.1371/journal.pbio.1002399.t001



**Fig 1. General biochemical model of aldehyde-degrading BMCs (metabolosomes) illustrating the common metabolosome core enzymes and reactions.** Substrates and cofactors involving the PTAC reaction are shown in red; other substrates and enzymes are shown in black, and other cofactors are shown in gray.

doi:10.1371/journal.pbio.1002399.g001

by Pta [14] for biosynthesis. Collectively, the aldehyde and alcohol dehydrogenases, as well as the PTAC, constitute the common metabolosome core.

The activities of core enzymes are not confined to BMC-associated functions: aldehyde and alcohol dehydrogenases are utilized in diverse metabolic reactions, and PTAC catalyzes a key biochemical reaction in the process of obtaining energy during fermentation [14]. The concerted functioning of a PTAC and an acetate kinase (Ack) is crucial for ATP generation in the fermentation of pyruvate to acetate (see **Reactions 1** and **2**). Both enzymes are, however, not restricted to fermentative organisms. They can also work in the reverse direction to activate acetate to the CoA-thioester. This occurs, for example, during acetoclastic methanogenesis in the archaeal *Methanosarcina* species [15,16].

**Reaction 1:** acetyl-S-CoA + P<sub>i</sub> ↔ acetyl phosphate + CoA-SH (PTAC)

**Reaction 2:** acetyl phosphate + ADP ↔ acetate + ATP (Ack)

The canonical PTAC, Pta, is an ancient enzyme found in some eukaryotes [17] and archaea [16], and widespread among the bacteria; 90% of the bacterial genomes in the Integrated Microbial Genomes database [18] contain a gene encoding the PTA\_PTB phosphotransacetylase (Pfam domain PF01515 [19,20]). Pta has been extensively characterized due to its key role in fermentation [14,21]. More recently, a second type of PTAC without any sequence homology to Pta was identified [4]. This protein, PduL (Pfam domain PF06130), was shown to catalyze the conversion of propionyl-CoA to propionyl-phosphate and is associated with a BMC involved in propanediol utilization, the PDU BMC [4].

Both *pduL* and *pta* genes can be found in genetic loci of functionally distinct BMCs, although the PduL type is much more prevalent, being found in all but one type of metabolosome locus: EUT1 (Table 1) [2]. Furthermore, in the Integrated Microbial Genomes Database [18], 91% of genomes that encode PF06130 also encode genes for shell proteins. As a member of the core biochemical machinery of functionally diverse aldehyde-oxidizing metabolosomes, PduL must have a certain level of substrate plasticity (see Table 1) that is not required of Pta, which has generally been observed to prefer acetyl-CoA [22,23]. PduL from the PDU BMC of

*Salmonella enterica* favors propionyl-CoA over acetyl-CoA [4], and it is likely that PduL orthologs in functionally diverse BMCs would have substrate preferences for other CoA derivatives. Another distinctive feature of BMC-associated PduL homologs is an N-terminal encapsulation peptide (EP) that is thought to “target” proteins for encapsulation by the BMC shell [3,24]. EPs are frequently found on BMC-associated proteins and have been shown to interact with shell proteins [25,26]. EPs have also been observed to cause proteins to aggregate [27,28], and this has recently been suggested to be functionally relevant as an initial step in metabolosome assembly, in which a multifunctional protein core is formed, around which the shell assembles [24].

Of the three common metabolosome core enzymes, crystal structures are available for both the alcohol and aldehyde dehydrogenases. In contrast, the structure of PduL, the PTAC found in the vast majority of catabolic BMCs, has not been determined. This is a major gap in our understanding of metabolosome-encapsulated biochemistry and cofactor recycling. Structural information will be essential to working out how the core enzymes and their cofactors assemble and organize within the organelle lumen to enhance catalysis. Moreover, it will be useful for guiding efforts to engineer novel BMC cores for biotechnological applications [1,29,30].

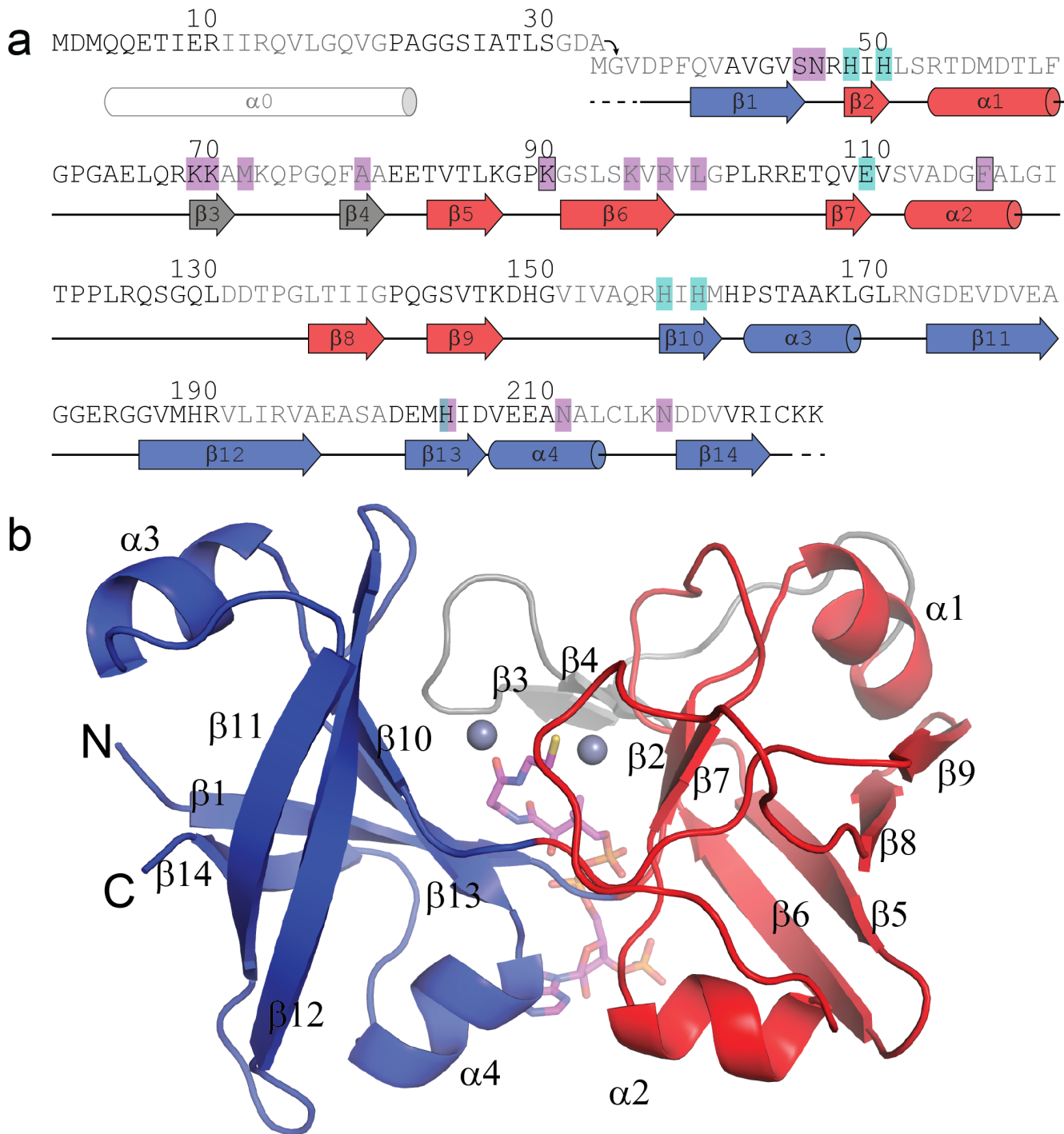
The primary structure of PduL homologs is subdivided into two PF06130 domains, each roughly 80 residues in length. No available protein structures contain the PF06130 domain, and homology searches using the primary structure of PduL do not return any significant results that would allow prediction of the structure. Moreover, the evident novelty of PduL makes its structure interesting in the context of convergent evolution of PTAC function; to-date, only the Pta active site and catalytic mechanism is known [31]. Here we report high-resolution crystal structures of a PduL-type PTAC in both CoA- and phosphate-bound forms, completing our understanding of the structural basis of catalysis by the metabolosome common core enzymes. We propose a catalytic mechanism analogous but yet distinct from the ubiquitous Pta enzyme, highlighting the functional convergence of two enzymes with completely different structures and metal requirements. We also investigate the quaternary structures of three different PduL homologs and situate our findings in the context of organelle biogenesis in functionally diverse BMCs.

## Results

### Structure Determination of PduL

We cloned, expressed, and purified three different PduL homologs from functionally distinct BMCs (Table 1): from the well-studied *pdu* locus in *S. enterica* Typhimurium LT2 (sPduL) [3,4], from the recently characterized *pvm* locus in *Planctomyces limnophilus* (pPduL) [32], and from the *grm3* locus in *Rhodospseudomonas palustris* BisB18 (rPduL) [2]. While purifying full-length sPduL, we observed a tendency to aggregation as described previously [4], with a large fraction of the expressed protein found in the insoluble fraction in a white, cake-like pellet. Remarkably, after removing the N-terminal putative EP (27 amino acids), most of the sPduL $\Delta$ EP protein was in the soluble fraction upon cell lysis. Similar differences in solubility were observed for pPduL and rPduL when comparing EP-truncated forms to the full-length protein, but none were quite as dramatic as for sPduL. We confirmed that all homologs were active (S1a and S1b Fig). Among these, we were only able to obtain diffraction-quality crystals of rPduL after removing the N-terminal putative EP (33 amino acids, also see Fig 2a) (rPduL $\Delta$ EP). Truncated rPduL $\Delta$ EP had comparable enzymatic activity to the full-length enzyme (S1a Fig).

We collected a native dataset from rPduL $\Delta$ EP crystals diffracting to a resolution of 1.54 Å (Table 2). Using a mercury-derivative crystal form diffracting to 1.99 Å (Table 2), we obtained



**Fig 2. Structural overview of *R. palustris* PduL from the *gpm3* locus.** (a) Primary and secondary structure of rPduL (tubes represent  $\alpha$ -helices, arrows  $\beta$ -sheets and dashed line residues disordered in the structure. Blocks of ten residues are shaded alternatively black/dark gray. The first 33 amino acids are present only in the wildtype construct and contains the predicted EP alpha helix,  $\alpha_0$ ; the truncated rPduL $\Delta$ EP that was crystallized begins with M-G-V. Coloring is according to structural domains (domain 1 D36-N46/Q155-C224, blue; loop insertion G61-E81, grey; domain 2 R47-F60/E82-A154, red). Metal coordination residues are highlighted in light blue and CoA contacting residues in magenta, residues contacting the CoA of the other chain are also outlined. (b) Cartoon representation of the structure colored by domains and including secondary structure numbering. The N- and C-termini are in close proximity. Coenzyme A is shown in magenta sticks and Zinc (grey) as spheres.

doi:10.1371/journal.pbio.1002399.g002

**Table 2. Data collection and refinement statistics**

	PduL native	PduL mercury derivative	PduL phosphate soaked
<b>Data collection</b>			
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	57.7, 56.4, 150.4	55.6, 57.7, 150.2	57.1, 58.8, 136.7
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	31.4 – 1.54 (1.60 – 1.54)*	35.3 – 1.99 (2.07 – 1.99)	39.2 – 2.10 (2.21 – 2.10)
<i>R</i> <sub>merge</sub>	0.169 (1.223)	0.084 (0.299)	0.122 (0.856)
<i>I</i> / $\sigma$ ( <i>I</i> )	12.9 (1.7)	22.1 (7.1)	12.6 (2.0)
Completeness (%)	99.4 (94.4)	99.3 (93.3)	100 (99.9)
Redundancy	13.9 (12.1)	7.2 (7.0)	6.5 (6.1)
<b>Refinement</b>			
Resolution (Å)	31.4 – 1.54 (1.60 – 1.54)*		39.2 – 2.10 (2.18 – 2.1)
No. reflections	72,698		27,554
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub> (%)	18.9 (30.7) / 22.1 (34.7)		17.5 (24.2) / 22.6 (30.0)
No. atoms	3,453		3,127
Protein	2,841		2,838
Ligand/ion	100		24
Water	512		265
B-factors	22.8		34.7
Protein	21.5		24.3
Ligand/ion	21.9		40.6
Water	30.3		37.9
R.m.s deviations			
Bond lengths (Å)	0.006		0.013
Bond angles (°)	1.26		1.30
Ramachandran Plot			
favoured (%)	99		99
allowed (%)	1		1
disallowed (%)	0		0

\*Highest resolution shell is shown in parentheses.

doi:10.1371/journal.pbio.1002399.t002

high quality electron density for model building and used the initial model to refine against the native data to *R*<sub>work</sub>/*R*<sub>free</sub> values of 18.9/22.1%. There are two PduL molecules in the asymmetric unit of the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> unit cell. We were able to fit all of the primary structure of PduLΔEP into the electron density with the exception of three amino acids at the N-terminus and two amino acids at the C-terminus (Fig 2a); the model is of excellent quality (Table 2). A CoA cofactor as well as two metal ions are clearly resolved in the density (for omit maps of CoA see S2 Fig).

Structurally, PduL consists of two domains (Fig 2, blue/red), each a beta-barrel that is capped on both ends by short  $\alpha$ -helices.  $\beta$ -Barrel 1 consists of the N-terminal  $\beta$  strand and  $\beta$  strands from the C-terminal half of the polypeptide chain ( $\beta$ 1,  $\beta$ 10– $\beta$ 14; residues 37–46 and 155–224).  $\beta$ -Barrel 2 consists mainly of the central segment of primary structure ( $\beta$ 2,  $\beta$ 5– $\beta$ 9; residues 47–60 and 82–154) (Fig 2, red), but is interrupted by a short two-strand beta sheet ( $\beta$ 3– $\beta$ 4, residues 61–81). This  $\beta$ -sheet is involved in contacts between the two domains and forms a lid over the active site. Residues in this region (Gln42, Pro43, Gly44), covering the

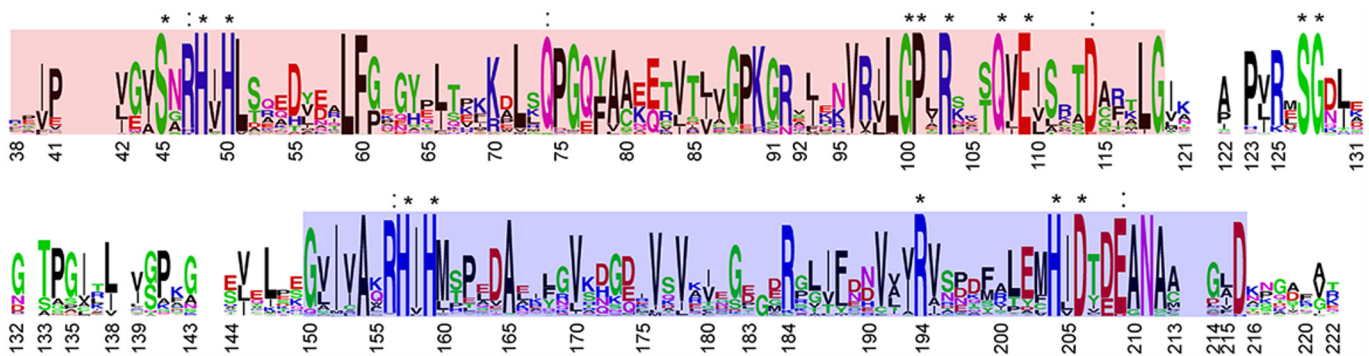
active site, are strongly conserved (Fig 3). This structural arrangement is completely different from the functionally related Pta, which is composed of two domains, each consisting of a central flat beta sheet with alpha-helices on the top and bottom [31].

There are two PduL molecules in the asymmetric unit forming a butterfly-shaped dimer (Fig 4c). Consistent with this, results from size exclusion chromatography of rPduLΔEP suggest that it is a dimer in solution (Fig 5e). The interface between the two chains buries 882 Å<sup>2</sup> per monomer and is mainly formed by α-helices 2 and 4 and parts of β-sheets 12 and 14, as well as a π–π stacking of the adenine moiety of CoA with Phe116 of the adjacent chain (Fig 4c). The folds of the two chains in the asymmetric unit are very similar, superimposing with a rmsd of 0.16 Å over 2,306 aligned atom pairs. The peripheral helices and the short antiparallel β3–4 sheet mediate most of the crystal contacts.

### Active Site Properties

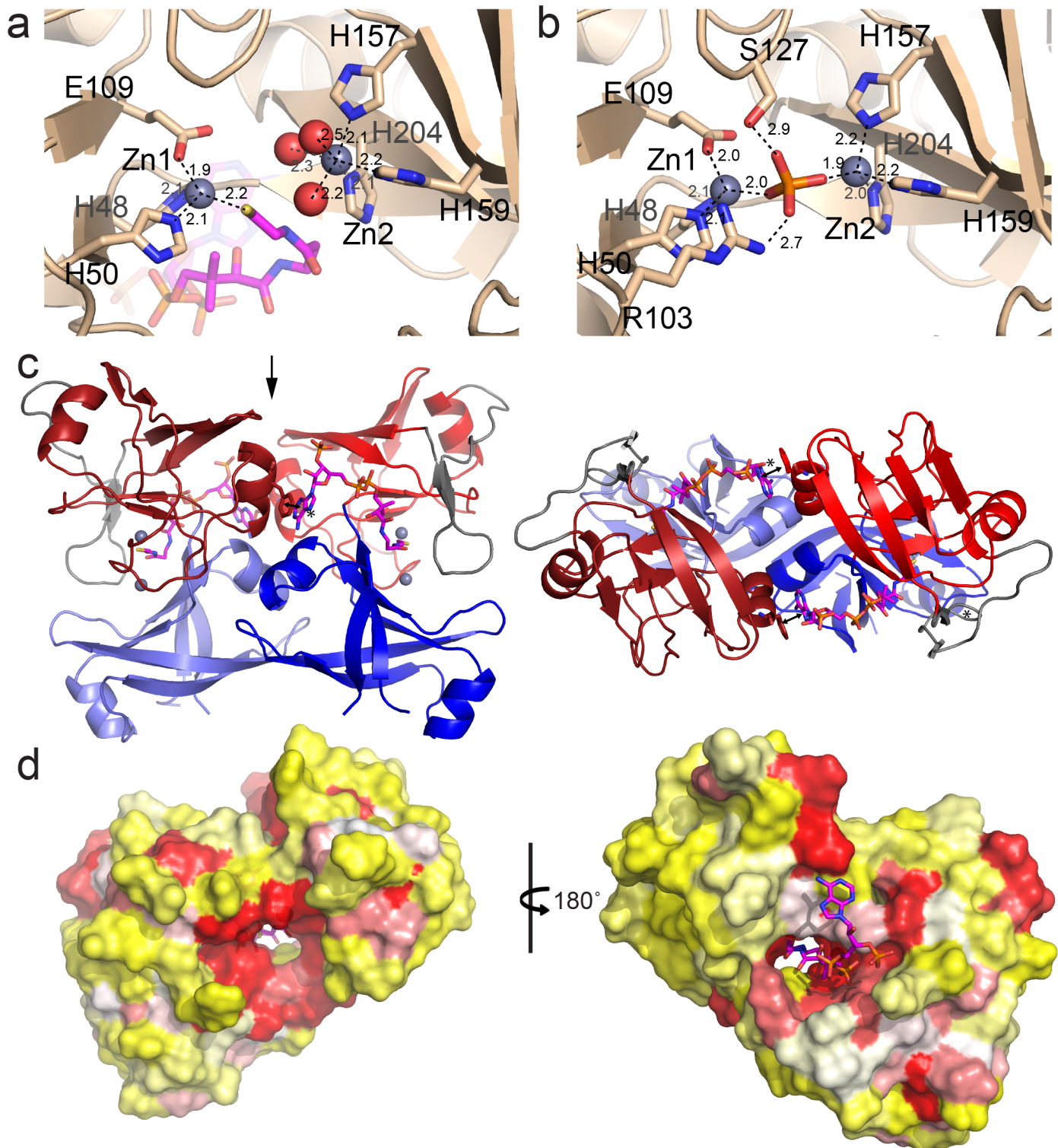
CoA and the metal ions bind between the two domains, presumably in the active site (Figs 2b and 4a). To identify the bound metals, we performed an X-ray fluorescence scan on the crystals at various wavelengths (corresponding to the K-edges of Mn, Fe, Co, Ni, Cu, and Zn). There was a large signal at the zinc edge, and we tested for the presence of zinc by collecting full data sets before and after the Zn K-edge (1.2861 and 1.2822 Å, respectively). The large differences between the anomalous signals confirm the presence of zinc at both metal sites (S3 Fig).

The first zinc ion (Zn1) is in a tetrahedral coordination state with His48, His50, Glu109, and the CoA sulfur (Fig 4a). The second (Zn2) is in octahedral coordination by three conserved histidine residues (His157, His159 and His204) as well as three water molecules (Fig 4a). The nitrogen atom coordinating the zinc is the Ne in each histidine residue, as is typical for this interaction [33]. When the crystals were soaked in a sodium phosphate solution for 2 d prior to data collection, the CoA dissociates, and density for a phosphate molecule is visible at the active site (Table 2, Fig 4b). The phosphate-bound structure aligns well with the CoA-bound structure (0.43 Å rmsd over 2,361 atoms for the monomer, 0.83 Å over 5,259 aligned atoms for the dimer). The phosphate contacts both zinc atoms (Fig 4b) and replaces the coordination by CoA at Zn1; the coordination for Zn2 changes from octahedral with three bound waters to tetrahedral with a phosphate ion as one of the ligands (Fig 4b). Conserved Arg103 seems to be



**Fig 3. Primary structure conservation of the PduL protein family.** Sequence logo calculated from the multiple sequence alignment of PduL homologs (see Materials and Methods), but not including putative EP sequences. Residues 100% conserved across all PduL homologs in our dataset are noted with an asterisk, and residues conserved in over 90% of sequences are noted with a colon. The sequences aligning to the PF06130 domain (determined by BLAST) are highlighted in red and blue. The position numbers shown correspond to the residue numbering of rPduL; note that some positions in the logo represent gaps in the rPduL sequence.

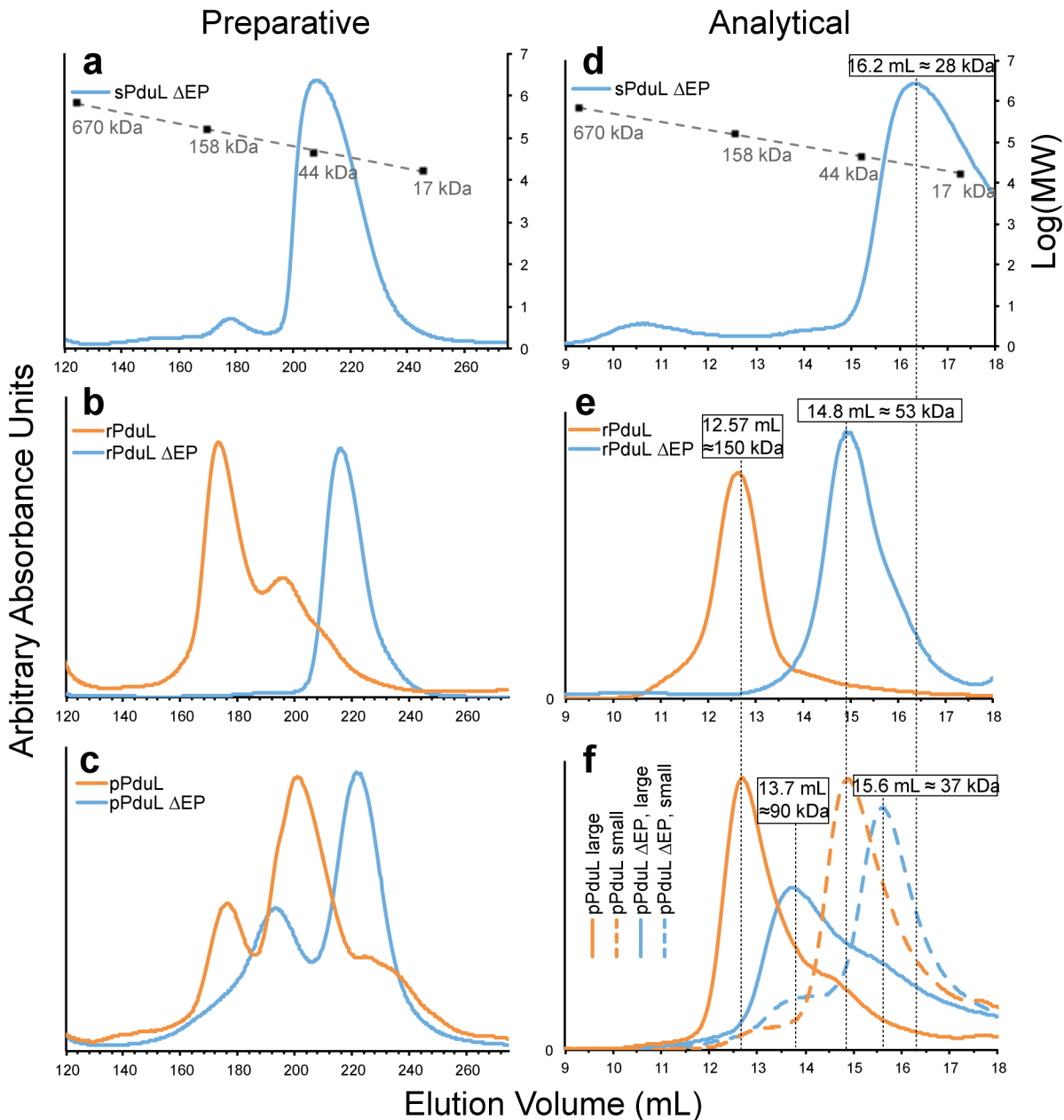
doi:10.1371/journal.pbio.1002399.g003



**Fig 4. Details of active site, dimeric assembly, and sequence conservation of PduL.** (a,b) Proposed active site of PduL with relevant residues shown as sticks in atom coloring (nitrogen blue, oxygen red, sulfur yellow), zinc as grey colored spheres and coordinating ordered water molecules in red. Distances between atom centers are indicated in Å. (a) Coenzyme A containing, (b) phosphate-bound structure. (c) View of the dimer in the asymmetric unit from the side, domains 1 and 2 colored as in Fig 2 and the two chains differentiated by blue/red versus slate/firebrick. The bottom panel shows a top view down the 2-fold axis as indicated by the arrow in the top panel. The asterisk and double arrow marks the location of the  $\pi$ - $\pi$  interaction between F116 and the CoA base of the other dimer chain. (d) Surface representation of the structure with indicated conservation (red: high, white: intermediate, yellow: low).

doi:10.1371/journal.pbio.1002399.g004





**Fig 5. Size exclusion chromatography of PduL homologs.** (a)–(c): Chromatograms of sPduL (a), rPduL (b), and pPduL (c) with (orange) or without (blue) the predicted EP, post-nickel affinity purification, applied over a preparative size exclusion column (see [Materials and Methods](#)). (d)–(f): Chromatograms of sPduL (d), rPduL (e), and pPduL (f) post-preparative size exclusion chromatography with different size fractions separated, applied over an analytical size exclusion column (see [Materials and Methods](#)). All chromatograms are cropped to show only the linear range of separation based on standard runs, shown in black squares with a dashed linear trend line. All y-axes are arbitrary absorbance units except the right-hand axes for panels (a) and (d), which is the  $\log_{10}$ (molecular weight) of the standards.

doi:10.1371/journal.pbio.1002399.g005

involved in maintaining the phosphate in that position. The two zinc atoms are slightly closer together in the phosphate-bound form (5.8 Å vs 6.3 Å), possibly due to the bridging effect of the phosphate. An additional phosphate molecule is bound at a crystal contact interface, perhaps accounting for the 14 Å shorter c-axis in the phosphate-bound crystal form ([Table 2](#)).

## Oligomeric States of PduL Orthologs Are Influenced by the EP

Interestingly, some of the residues important for dimerization of rPduL, particularly Phe116, are poorly conserved across PduL homologs associated with functionally diverse BMCs (Figs 4c and 3), suggesting that they may have alternative oligomeric states. We tested this hypothesis by performing size exclusion chromatography on both full-length and truncated variants (lacking the EP,  $\Delta$ EP) of sPduL, rPduL, and pPduL. These three homologs are found in functionally distinct BMCs (Table 1). Therefore, they are packaged with different signature enzymes and different ancillary proteins [2]. It has been proposed that the catabolic BMCs may assemble in a core-first manner, with the luminal enzymes (signature enzyme, aldehyde, and alcohol dehydrogenases and the BMC PTAC) forming an initial bolus, or prometabolosome, around which a shell assembles [1]. Given the diversity of signature enzymes (Table 1), it is plausible that PduL orthologs may adopt different oligomeric states that reflect the differences in the proteins being packaged with them in the organelle lumen.

We found that not only did the different orthologs appear to assemble into different oligomeric states, but that quaternary structure was dependent on whether or not the EP was present. Full-length sPduL was unstable in solution—precipitating over time—and eluted throughout the entire volume of a size exclusion column, indicating it was nonspecifically aggregating. However, when the putative EP (residues 1–27) was removed (sPduL  $\Delta$ EP), the truncated protein was stable and eluted as a single peak (Fig 5a) consistent with the size of a monomer (Fig 5d, blue curve). In contrast, both full-length rPduL and pPduL appeared to exist in two distinct oligomeric states (Fig 5b and 5c respectively, orange curves), one form of the approximate size of a dimer and the second, a higher molecular weight oligomer (~150 kDa). Upon deletion of the putative EP (residues 1–47 for rPduL, and 1–20 for pPduL), there was a distinct change in the elution profiles (Fig 5b and 5c respectively, blue curves). pPduL $\Delta$ EP eluted as two smaller forms, possibly corresponding to a trimer and a monomer. In contrast, rPduL $\Delta$ EP eluted as one smaller oligomer, possibly a dimer. We also analyzed purified rPduL and rPduL $\Delta$ EP by size exclusion chromatography coupled with multiangle light scattering (SEC-MALS) for a complementary approach to assessing oligomeric state. SEC-MALS analysis of rPduL $\Delta$ EP is consistent with a dimer (as observed in the crystal structure) with a weighted average ( $M_w$ ) and number average ( $M_n$ ) of the molar mass of 58.4 kDa  $\pm$  11.2% and 58.8 kDa  $\pm$  10.9%, respectively (S4a Fig). rPduL full length runs as  $M_w = 140.3$  kDa  $\pm$  1.2% and  $M_n = 140.5$  kDa  $\pm$  1.2%. This corresponds to an oligomeric state of six subunits (calculated molecular weight of 144 kDa). Collectively, these data strongly suggest that the N-terminal EP of PduL plays a role in defining the quaternary structure of the protein.

## Discussion

The hallmark attribute of an organelle is that it serves as a discrete subcellular compartment functioning as an isolated microenvironment distinct from the cytosol. In order to create and preserve this microenvironment, the defining barrier (i.e., lipid bilayer membrane or micro-compartment shell) must be selectively permeable. The BMC shell not only sequesters specific enzymes but also their cofactors, thereby establishing a private cofactor pool dedicated to the encapsulated reactions. In catabolic BMCs, CoA and NAD<sup>+</sup> must be continually recycled within the organelle (Fig 1). Homologs of the predominant cofactor utilizer (aldehyde dehydrogenase) and NAD<sup>+</sup> regenerator (alcohol dehydrogenase) have been structurally characterized, but until now structural information was lacking for PduL, which recycles CoA in the organelle lumen [12,34]. Curiously, while the housekeeping Pta could provide this function,

and indeed does so in the case of one type of ethanolamine-utilizing (EUT) BMC [2], the evolutionarily unrelated PduL fulfills this function for the majority of metabolosomes [2,4] using a novel structure and active site for convergent evolution of function.

## The Tertiary Structure of PduL Is Formed by Discontinuous Segments of Primary Structure

The structure of PduL consists of two  $\beta$ -barrel domains capped by short alpha helical segments (Fig 2b). The two domains are structurally very similar (superimposing with a rmsd of 1.34 Å (over 123 out of 320/348 aligned backbone atoms, S5a Fig). However, the amino acid sequences of the two domains are only 16% identical (mainly the RHxH motif,  $\beta$ 2 and  $\beta$ 10), and 34% similar. Our structure reveals that the two assigned PF06130 domains (Fig 3) do not form structurally discrete units; this reduces the apparent sequence conservation at the level of primary structure. One strand of the domain 1 beta barrel (shown in blue in Fig 2) is contributed by the N-terminus, while the rest of the domain is formed by the residues from the C-terminal half of the protein. When aligned by structure, the  $\beta$ 1 strand of the first domain (Fig 2a and 2b, blue) corresponds to the final strand of the second domain ( $\beta$ 9), effectively making the domains continuous if the first strand was transplanted to the C-terminus. Refined domain assignment based on our structure should be able to predict domains of PF06130 homologs much more accurately. The closest structural homolog of the PduL barrel domain is a subdomain of a multienzyme complex, the alpha subunit of ethylbenzene dehydrogenase [35] (S5b Fig, rmsd of 2.26 Å over 226 aligned atoms consisting of one beta barrel and one capping helix). In contrast to PduL, there is only one barrel present in ethylbenzene dehydrogenase, and there is no comparable active site arrangement. The PduL signature primary structure, two PF06130 domains, occurs in some multidomain proteins, most of them annotated as Acks, suggesting that PduL may also replace Pta in variants of the phosphotransacetylase-Ack pathway. These PduL homologs lack EPs, and their fusion to Ack may have evolved as a way to facilitate substrate channeling between the two enzymes.

## Implications for Metabolosome Core Assembly

For BMC-encapsulated proteins to properly function together, they must be targeted to the lumen and assemble into an organization that facilitates substrate/product channeling among the different catalytic sites of the signature and core enzymes. The N-terminal extension on PduL homologs may serve both of these functions. The extension shares many features with previously characterized EPs [24,26,36]: it is present only in homologs associated with BMC loci, and it is predicted to form an amphipathic  $\alpha$ -helix. Moreover, its removal affects the oligomeric state of the protein. EP-mediated oligomerization has been observed for the signature and core BMC enzymes; for example, full-length propanediol dehydratase and ethanolamine ammonia-lyase (signature enzymes for PDU and EUT BMCs) subunits are also insoluble, but become soluble upon removal of the predicted EP [27,28,11]. sPduL has also previously been reported to localize to inclusion bodies when overexpressed [4]; we show here that this is dependent on the presence of the EP. This propensity of the EP to cause proteins to form complexes (Fig 5) might not be a coincidence, but could be a necessary step in the assembly of BMCs. Structured aggregation of the core enzymes has been proposed to be the initial step in metabolosome assembly [1,37] and is known to be the first step of  $\beta$ -carboxysome biogenesis, where the core enzyme Ribulose Biphosphate Carboxylase/Oxygenase (RuBisCO) is aggregated by the CcmM protein [37]. Likewise, CsoS2, a protein in the  $\alpha$ -carboxysome core, also aggregates when purified and is proposed to facilitate the nucleation and encapsulation of RuBisCO molecules in the lumen of the organelle [36]. Coupled with protein-protein interactions with other luminal components, the aggregation of these enzymes could lead to a densely

packed organelle core. This role for EPs in BMC assembly is in addition to their interaction with shell proteins [24–26,36,38].

Moreover, the PduL crystal structures offer a clue as to how required cofactors enter the BMC lumen during assembly. Free CoA and NAD<sup>+</sup>/H could potentially be bound to the enzymes as the core assembles and is encapsulated. However, this raises an issue of stoichiometry: if the ratio of cofactors to core enzymes is too low, then the sequestered metabolism would proceed at suboptimal rates. Our PduL crystals contained CoA that was captured from the *Escherichia coli* cytosol, indicating that the “ground state” of PduL is in the CoA-bound form; this could provide an elegantly simple means of guaranteeing a 1:1 ratio of CoA:PduL within the metabolosome lumen.

## Active Site Identification and Structural Insights into Catalysis

The active site of PduL is formed at the interface of the two structural domains (Fig 2b). As expected, the amino acid sequence conservation is highest in the region around the proposed active site (Fig 4d); highly conserved residues are also involved in CoA binding (Figs 2a and 3, residues Ser45, Lys70, Arg97, Leu99, His204, Asn211). All of the metal-coordinating residues (Fig 2a) are absolutely conserved, implicating them in catalysis or the correct spatial orientation of the substrates. Arg103, which contacts the phosphate (Fig 4b), is present in all PduL homologs. The close resemblance between the structures binding CoA and phosphate likely indicates that no large changes in protein conformation are involved in catalysis, and that our crystal structures are representative of the active form. The native substrate for the forward reaction of rPduL and pPduL, propionyl-CoA, most likely binds to the enzyme in the same way at the observed nucleotide and pantothenic acid moiety, but the propionyl group in the CoA-thioester might point in a different direction. There is a pocket nearby the active site between the well-conserved residues Ser45 and Ala154, which could accommodate the propionyl group (S6 Fig). A homology model of sPduL indicates that the residues making up this pocket and the surrounding active site region are identical to that of rPduL, which is not surprising, because these two homologs presumably have the same propionyl-CoA substrate. The homology model of pPduL also has identical residues making up the pocket, but with a key difference in the vicinity of the active site: Gln77 of rPduL is replaced by a tyrosine (Tyr77) in pPduL. The physiological substrate of pPduL (Table 1) is thought to be lactyl-CoA, which contains an additional hydroxyl group relative to propionyl-CoA. The presence of an aromatic residue at this position may underlie the substrate preference of the PduL enzyme from the *pvm* locus. Indeed, in the majority of PduLs encoded in *pvm* loci, Gln77 is substituted by either a Tyr or Phe, whereas it is typically a Gln or Glu in PduLs in all other BMC types that degrade acetyl- or propionyl-CoA. A comparison of the PduL active site to that of the functionally identical Pta suggests that the two enzymes have distinctly different mechanisms. The catalytic mechanism of Pta involves the abstraction of a thiol hydrogen by an aspartate residue, resulting in the nucleophilic attack of thiolate upon the carbonyl carbon of acetyl-phosphate, oriented by an arginine and stabilized by a serine [31]—there are no metals involved. In contrast, in the rPduL structure, there are no conserved aspartate residues in or around the active site, and the only well-conserved glutamate residue in the active site is involved in coordinating one of the metal ions. These observations strongly suggest that an acidic residue is not directly involved in catalysis by PduL. Instead, the dimetal active site of PduL may create a nucleophile from one of the hydroxyl groups on free phosphate to attack the carbonyl carbon of the thioester bond of an acyl-CoA. In the reverse direction, the metal ion(s) could stabilize the thiolate anion that would attack the carbonyl carbon of an acyl-phosphate; a similar mechanism has been described for phosphatases where hydroxyl groups or hydroxide ions can act as a base when coordinated by a dimetal active site [39].

Our structures provide the foundation for studies to elucidate the details of the catalytic mechanism of PduL. Conserved residues in the active site that may contribute to substrate binding and/or transition state stabilization include Ser127, Arg103, Arg194, Gln107, Gln74, and Gln/Glu77. In the phosphate-bound crystal structure, Ser127 and Arg103 appear to position the phosphate (Fig 4b). Alternatively, Arg103 might act as a base to render the phosphate more nucleophilic. The functional groups of Gln74, Gln/Glu77, and Arg194 are directed away from the active site in both CoA and phosphate-bound crystal structures and do not appear to be involved in hydrogen bonding with these substrates, although they could be important for positioning an acyl-phosphate.

The free CoA-bound form is presumably poised for attack upon an acyl-phosphate, indicating that the enzyme initially binds CoA as opposed to acyl-phosphate. This hypothesis is strengthened by the fact that the CoA-bound crystals were obtained without added CoA, indicating that the protein bound CoA from the *E. coli* expression strain and retained it throughout purification and crystallization. The phosphate-bound structure indicates that in the opposite reaction direction phosphate is bound first, and then an acyl-CoA enters. The two high-resolution crystal structures presented here will serve as the foundation for mechanistic studies on this noncanonical PTAC enzyme to determine how the dimetal active site functions to catalyze both forward and reverse reactions.

## Functional, but Not Structural, Convergence of PduL and Pta

PduL and Pta are mechanistically and structurally distinct enzymes that catalyze the same reaction [4], a prime example of evolutionary convergence upon a function. There are several examples of such functional convergence of enzymes, although typically the enzymes have independently evolved similar, or even identical active sites; for example, the carbonic anhydrase family [40,41]. However, apparently less frequent is functional convergence that is supported by distinctly different active sites and accordingly catalytic mechanism, as revealed by comparison of the structures of Pta and PduL. One well-studied example of this is the  $\beta$ -lactamase family of enzymes, in which the active site of Class A and Class C enzymes involve serine-based catalysis, but Class B enzymes are metalloproteins [42,43]. This is not surprising, as  $\beta$ -lactamases are not so widespread among bacteria and therefore would be expected to have evolved independently several times as a defense mechanism against  $\beta$ -lactam antibiotics. However, nearly all bacteria encode Pta, and it is not immediately clear why the Pta/PduL functional convergence should have evolved: it would seem to be evolutionarily more resourceful for the Pta-encoding gene to be duplicated and repurposed for BMCs, as is apparently the case in one type of BMC—EUT1 (Table 1). There could be some intrinsic biochemical difference between the two enzymes that renders PduL a more attractive candidate for encapsulation in a BMC—for example, PduL might be more amenable to tight packaging, or is better suited for the chemical microenvironment formed within the lumen of the BMC, which can be quite different from the cytosol [44,45]. Further biochemical comparison between the two PTACs will likely yield exciting results that could answer this evolutionary question.

## Implications

BMCs are now known to be widespread among the bacteria and are involved in critical segments of both autotrophic and heterotrophic biochemical pathways that confer to the host organism a competitive (metabolic) advantage in select niches. As one of the three common metabolosome core enzymes, the structure of PduL provides a key missing piece to our structural picture of the shared core biochemistry (Fig 1) of functionally diverse catabolic BMCs. We have observed the oligomeric state differences of PduL to correlate with the presence of an

EP, providing new insight into the function of this sequence extension in BMC assembly. Moreover, our results suggest a means for Coenzyme A incorporation during metabolosome biogenesis. A detailed understanding of the underlying principles governing the assembly and internal structural organization of BMCs is a requisite for synthetic biologists to design custom nanoreactors that use BMC architectures as a template. Furthermore, given the growing number of metabolosomes implicated in pathogenesis [46–50], the PduL structure will be useful in the development of therapeutics. It is gradually being realized that the metabolic capabilities of a pathogen are also important for virulence, along with the more traditionally cited factors like secretion systems and effector proteins [51]. The fact that PduL is confined almost exclusively to metabolosomes can be used to develop an inhibitor that blocks only PduL and not Pta as a way to selectively disrupt BMC-based metabolism, while not affecting most commensal organisms that require PTAC activity.

## Materials and Methods

### Molecular Cloning

Genes for PduL homologs with and without the EP were amplified via PCR using the primers listed in [S1 Table](#). sPduL was amplified using *S. enterica* Typhimurium LT2 genomic DNA, and pPduL and rPduL sequences were codon optimized and synthesized by GenScript with the 6xHis tag. All 5' primers included EcoRI and BglII restriction sites, and all 3' primers included a BamHI restriction site to facilitate cloning using the BglBricks strategy. 5' primers also included the sequence TTTAAGAAGGAGATATACCATG downstream of the restriction sites, serving as a strong ribosome binding site. The 6x polyhistidine tag sequence was added to the 3' end of the gene using the BglBricks strategy and was subcloned into the pETBb3 vector, a pET21b-based vector modified to be BglBricks compatible.

### Protein Purification, Size Exclusion Chromatography, and Protein Crystallization

*E. coli* BL21(DE3) expression strains containing the relevant PduL construct in the pETBb3 vector were grown overnight at 37°C in standard LB medium and then used to inoculate 1L of standard LB medium in 2.8 L Fernbach flasks at a 1:100 dilution, which were then incubated at 37°C shaking at 150 rpm, until the culture reached an OD<sub>600</sub> of 0.8–1.0, at which point cultures were induced with 200 μM IPTG (isopropylthio-β-D-galactoside) and incubated at 20°C for 18 h, shaking at 150 rpm. Cells were centrifuged at 5,000 xg for 15 min, and cell pellets were frozen at –20°C.

For protein purifications, cell pellets from 1–3 L cultures were resuspended in 20–30 ml buffer A (50 mM Tris-HCl pH 7.4, 300 mM NaCl) and lysed using a French pressure cell at 20,000 lb/in<sup>2</sup>. The resulting cell lysate was centrifuged at 15,000 xg. 30 mM imidazole was added to the supernatant that was then applied to a 5 mL HisTrap column (GE Healthcare Bio-Sciences, Pittsburgh, PA). Protein was eluted off the column using a gradient of buffer A from 0 mM to 500 mM imidazole over 20 column volumes. Fractions corresponding to PduL were pooled and concentrated using Amicon Ultra Centrifugal filters (EMD Millipore, Billerica, MA) to a volume of no more than 2.5 mL. The protein sample was then applied to a HiLoad 26/60 Superdex 200 preparative size exclusion column (GE Healthcare Bio-Sciences, Pittsburgh, PA) and eluted with buffer B (20 mM Tris pH 7.4, 50 mM NaCl). Where applicable, fractions corresponding to different oligomeric states were pooled separately, leaving one or two fractions in between to prevent cross contamination. Pooled fractions were concentrated to 1–20 mg/mL protein as determined by the Bradford method [52] prior to applying on a

Superdex 200 10/300 GL analytical size exclusion column (GE Healthcare Bio-Sciences, Pittsburgh, PA). Size standards used were Thyroglobulin 670 kDa,  $\gamma$ -globulin 158 kDa, Ovalbumin 44 kDa, and Myoglobin 17 kDa. For light scattering, the proteins were measured in a Protein Solutions Dynapro dynamic light scattering instrument with an acquisition time of 5 s, averaging 10 acquisitions at a constant temperature of 25°C. The radii were calculated assuming a globular particle shape.

Size exclusion chromatography coupled with SEC-MALS was performed on full-length rPduL and rPduL- $\Delta$ EP similar to Luzi et al. 2015 [53]. A Wyatt DAWN Heleos-II 18-angle light scattering instrument was used in tandem with a GE AKTA pure FPLC with built in UV detector, and a Wyatt Optilab T-Rex refractive index detector. Detector 16 of the DAWN Heleos-II was replaced with a Wyatt Dynapro Nanostar QELS detector for dynamic light scattering. A GE Superdex S200 10/300 GL column was used, with 125–100  $\mu$ l of protein sample at 1 mg/ml concentration injected, and the column run at 0.5 ml/min in 20 mM Tris, 50 mM NaCl, pH 7.4.

Each detector of the DAWN-Heleos-II was plotted with the Zimm model in the Wyatt ASTRA software to calculate the molar mass. The molar mass was measured at each collected data point across the peaks at  $\sim$ 1 point per 8  $\mu$ l eluent. Both the  $M_w$  and  $M_n$  of the molar mass calculations, as well as percent deviations, were also determined using Wyatt software program ASTRA.

For preparing protein for crystallography, expression cells were grown as above, except were induced with 50  $\mu$ M IPTG. Harvested cells were resuspended in buffer B and lysed using a French Press. Cleared lysate was applied on a 5 ml HisTrap HP column (GE Healthcare) and washed with buffer A containing 20 mM imidazole. Pdul-His was eluted with 2 CV buffer B containing 300 mM imidazole, concentrated and then applied on a HiLoad 26/60 Superdex 200 (GE Healthcare) column equilibrated in buffer B for final cleanup. Protein was then concentrated to 20–30 mg/ml for crystallization. Crystals were obtained from sitting drop experiments at 22°C, mixing 3  $\mu$ l of protein solution with 3  $\mu$ l of reservoir solution containing 39%–35% MPD. Crystals were flash frozen in liquid nitrogen after being adding 5  $\mu$ l of a reservoir solution. For heavy atom derivatives, 0.2  $\mu$ l of 100 mM Thiomersal (Hampton Research) was added to the crystallization drop 36 h prior to freezing. For phosphate soaks, 5  $\mu$ l reservoir and 1.5  $\mu$ l 200 mM sodium phosphate solution (pH 7.0) were added 2 d prior to flash freezing.

## PTAC Activity Assay

Enzyme reactions were performed in a 2 mL cuvette containing 50 mM Tris-HCl pH 7.5, 0.2 mM 5,5'-dithiobis-2-nitrobenzoic acid (DTNB; Ellman's reagent), 0.1 mM acyl-CoA, and 0.5  $\mu$ g purified PTAC, unless otherwise noted. To initiate the reaction, 5 mM  $\text{NaH}_2\text{PO}_4$  was added, the cuvette was inverted to mix, and the absorbance at 412 nm was measured every 2 s over the course of four minutes in a Nanodrop 2000c, in the cuvette holder.  $14,150 \text{ M}^{-1}\text{cm}^{-1}$  was used as the extinction coefficient of DTNB to determine the specific activity.

## PduL Sequence Analysis

A multiple sequence alignment of 228 PduL sequences associated with BMCs [2] and 20 PduL sequences not associated with BMCs was constructed using MUSCLE [54]. PduL sequences associated with BMCs were determined from Dataset S1 of Reference [2], and those not associated with BMCs were determined by searching for genomes that encoded PF06130 but not PF03319 nor PF00936 in the IMG database [18]. The multiple sequence alignment was visualized in Jalview [55], and the nonconserved N- and C-terminal amino acids were deleted. This trimmed alignment was used to build the sequence logo using WebLogo [56].

## Diffraction Data Collection, Structure Determination and Visualization

Diffraction data were collected at the Advanced Light Source at Lawrence Berkeley National Laboratory beamline 5.0.2 (100 K, 1.0000 Å wavelength for native data, 1.0093 Å for mercury derivative, 1.2861 Å for Zn pre-edge and 1.2822 Å for Zn peak). Diffraction data were integrated with XDS [57] and scaled with SCALA (CCP4 [58]). The structure of PduL was solved using phenix.autosol [59], which found 11 heavy atom sites and produced density suitable for automatic model building. The model was refined with phenix.refine [59], with refinement alternating with model building using 2Fo-Fc and Fo-Fc maps visualized in COOT [60]. Statistics for diffraction data collection, structure determination and refinement are summarized in Table 2. Figures were prepared using pymol ([www.pymol.org](http://www.pymol.org)) and Raster3D [61].

## Homology Modeling

Models of *S. enterica* Typhimurium LT2 and *P. limnophilus* PduL were generated with Modeler using the align2d and model-default scripts [62].

## Supporting Information

### S1 Data. Data for S1a Fig.

(XLSX)

### S2 Data. Data for S2b Fig.

(XLSX)

**S1 Fig. Enzymatic activity of PduL homologs.** (a) sPduL $\Delta$ EP, rPduL hexamer and rPduL $\Delta$ EP, (b) pPduL hexamer and dimer—pPduL experiments were performed with 8  $\mu$ g purified protein.

(TIFF)

### S2 Fig. Simulated annealing omit map of Coenzyme A region (a) and the active site (b).

Fofc density in green at 1.8 rmsd from a simulated annealing refinement run omitting Coenzyme A, Zn atoms, and water molecules.

(TIF)

**S3 Fig. Metal identification by differential anomalous signals.** Anomalous map density contoured at 0.025  $e^-/\text{\AA}^3$  in the vicinity of the metal sites for data collected at 1.2822 Å (red, Zn peak) and 1.2861 Å (yellow map, Zn pre-edge) identifies the metals as zinc based on the large decrease of signal when collecting the data above the Zn edge.

(TIF)

### S4 Fig. SEC-MALS analysis using Wyatt ASTRA software indicating the calculated mass of the peaks for (A) rPduL $\Delta$ EP and (B) rPduL full length.

(TIF)

**S5 Fig. Analysis of the structural domains of PduL.** (a) Structural alignment of domain 1 (red) and domain 2 (blue), linker region colored grey. (b) Structural alignment of domain 1 of PduL with a subdomain (residues 872–936 and 967–976 of chain A) of the ethylbenzene dehydrogenase from *Aromatoleum aromaticum* (pdb ID 2IVF).

(TIF)

**S6 Fig. Potential propionyl binding pocket in PduL.** The area around conserved residues A154 and S45 forms a potential binding pocket for the propionyl group (shaded ellipsoid).

(TIF)



**S1 Table. Primers used in this study.**  
(DOCX)

**Acknowledgments**

We thank the entire staff at the Advanced Light Source, Lawrence Berkeley National Laboratory (supported by the Office of Basic Energy Sciences of the US Department of Energy (DOE) (<http://www.energy.gov>) under contract DEAC02-05CH11231), and Marc Allaire in particular, for excellent support during data collection. We also thank Dr. Jan Zarzycki for valuable comments on the manuscript, Dr. Rodney Burton and Aiko Turmo for SEC-MALS analysis, and Cristina Mora for assistance with enzyme assays.

**Author Contributions**

Conceived and designed the experiments: CAK OE MS. Performed the experiments: OE MS. Analyzed the data: OE MS. Wrote the paper: CAK OE MS.

**References**

1. Kerfeld CA, Erbilgin O. Bacterial microcompartments and the modular construction of microbial metabolism. *Trends in microbiology*. 2015; 23(1):22–34. doi: [10.1016/j.tim.2014.10.003](https://doi.org/10.1016/j.tim.2014.10.003) PMID: [25455419](https://pubmed.ncbi.nlm.nih.gov/25455419/)
2. Axen SD, Erbilgin O, Kerfeld CA. A taxonomy of bacterial microcompartment loci constructed by a novel scoring method. *PLoS Comput Biol*. 2014; 10(10):e1003898. doi: [10.1371/journal.pcbi.1003898](https://doi.org/10.1371/journal.pcbi.1003898) PMID: [25340524](https://pubmed.ncbi.nlm.nih.gov/25340524/)
3. Liu Y, Jorda J, Yeates TO, Bobik TA. The PduL phosphotransacylase is used to recycle coenzyme A within the Pdu microcompartment. *Journal of bacteriology*. 2015. doi: [10.1128/JB.00056-15](https://doi.org/10.1128/JB.00056-15) PMID: [25962918](https://pubmed.ncbi.nlm.nih.gov/25962918/).
4. Liu Y, Leal NA, Sampson EM, Johnson CLV, Havemann GD, Bobik Ta. PduL is an Evolutionarily Distinct Phosphotransacylase Involved in B12-dependent 1,2-propanediol degradation by *Salmonella enterica* serovar typhimurium LT2. *Journal of bacteriology*. 2007; 189(5):1589–96. doi: [10.1128/JB.01151-06](https://doi.org/10.1128/JB.01151-06) PMID: [17158662](https://pubmed.ncbi.nlm.nih.gov/17158662/)
5. Kerfeld CA, Sawaya MR, Tanaka S, Nguyen CV, Phillips M, Beeby M, et al. Protein structures forming the shell of primitive bacterial organelles. *Science (New York, NY)*. 2005; 309:936–8. doi: [10.1126/science.1113397](https://doi.org/10.1126/science.1113397)
6. Pang A, Liang M, Prentice MB, Pickersgill RW. Substrate channels revealed in the trimeric *Lactobacillus reuteri* bacterial microcompartment shell protein PduB. *Acta Crystallographica Section D Biological Crystallography*. 2012; 68(12):1642–52. doi: [10.1107/S0907444912039315](https://doi.org/10.1107/S0907444912039315)
7. Bobik TA, Havemann GD, Busch RJ, Williams DS, Aldrich HC. The Propanediol Utilization (pdu) Operon of *Salmonella enterica* serovar Typhimurium LT2 Includes Genes Necessary for Formation of Polyhedral Organelles Involved in coenzyme B12-Dependent 1, 2-Propanediol Degradation. *Journal of bacteriology*. 1999; 181(19):5967–75. PMID: [10498708](https://pubmed.ncbi.nlm.nih.gov/10498708/)
8. Abdul-Rahman F, Petit E, Blanchard JL. The Distribution of Polyhedral Bacterial Microcompartments Suggests Frequent Horizontal Transfer and Operon Reassembly. *Phylogenetics and Evolutionary Biology*. 2013; 1(4):1000118. doi: [10.4172/2329-9002.1000118](https://doi.org/10.4172/2329-9002.1000118)
9. Jorda J, Lopez D, Wheatley NM, Yeates TO. Using comparative genomics to uncover new kinds of protein-based metabolic organelles in bacteria. *Protein Science*. 2013; 22(2):179–95. doi: [10.1002/pro.2196](https://doi.org/10.1002/pro.2196) WOS:000314216100001. PMID: [23188745](https://pubmed.ncbi.nlm.nih.gov/23188745/)
10. Roof DM, Roth JR. Ethanolamine utilization in *Salmonella typhimurium*. *Journal of bacteriology*. 1988; 170(9):3855–63. PMID: [3045078](https://pubmed.ncbi.nlm.nih.gov/3045078/)
11. Leal NA, Havemann GD, Bobik TA. PduP is a coenzyme-a-acylating propionaldehyde dehydrogenase associated with the polyhedral bodies involved in B12-dependent 1,2-propanediol degradation by *Salmonella enterica* serovar Typhimurium LT2. *Archives of microbiology*. 2003; 180(5):353–61. doi: [10.1007/s00203-003-0601-0](https://doi.org/10.1007/s00203-003-0601-0) PMID: [14504694](https://pubmed.ncbi.nlm.nih.gov/14504694/)
12. Huseby DL, Roth JR. Evidence that a metabolic microcompartment contains and recycles private cofactor pools. *Journal of bacteriology*. 2013; 195(12):2864–79. doi: [10.1128/JB.02179-12](https://doi.org/10.1128/JB.02179-12) PMID: [23585538](https://pubmed.ncbi.nlm.nih.gov/23585538/)

13. Cheng S, Fan C, Sinha S, Bobik TA. The PduQ enzyme is an alcohol dehydrogenase used to recycle NAD<sup>+</sup> internally within the Pdu microcompartment of *Salmonella enterica*. *PLoS ONE*. 2012; 7(10): e47144–e. doi: [10.1371/journal.pone.0047144](https://doi.org/10.1371/journal.pone.0047144) PMID: [23077559](https://pubmed.ncbi.nlm.nih.gov/23077559/)
14. White D. *The Physiology and Biochemistry of Prokaryotes*. 2 ed. New York: Oxford University Press; 2000.
15. Lundie LL Jr., Ferry JG. Activation of acetate by *Methanosarcina thermophila*. Purification and characterization of phosphotransacetylase. *J Biol Chem*. 1989; 264(31):18392–6. PMID: [2808380](https://pubmed.ncbi.nlm.nih.gov/2808380/).
16. Ferry JG. Enzymology of the fermentation of acetate to methane by *Methanosarcina thermophila*. *Biofactors*. 1997; 6(1):25–35. PMID: [9233537](https://pubmed.ncbi.nlm.nih.gov/9233537/).
17. Tielens AG, van Grinsven KW, Henze K, van Hellemond JJ, Martin W. Acetate formation in the energy metabolism of parasitic helminths and protists. *Int J Parasitol*. 2010; 40(4):387–97. doi: [10.1016/j.ijpara.2009.12.006](https://doi.org/10.1016/j.ijpara.2009.12.006) PMID: [20085767](https://pubmed.ncbi.nlm.nih.gov/20085767/).
18. Markowitz VM, Chen IMA, Palaniappan K, Chu K, Szeto E, Grechkin Y, et al. IMG: the Integrated Microbial Genomes database and comparative analysis system. *Nucleic acids research*. 2012; 40(Database issue):D115–22. doi: [10.1093/nar/gkr1044](https://doi.org/10.1093/nar/gkr1044) PMID: [22194640](https://pubmed.ncbi.nlm.nih.gov/22194640/)
19. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. *Nucleic acids research*. 2014; 42(Database issue):D222–30. doi: [10.1093/nar/gkt1223](https://doi.org/10.1093/nar/gkt1223) PMID: [24288371](https://pubmed.ncbi.nlm.nih.gov/24288371/)
20. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database. *Nucleic acids research*. 2012; 40(Database issue):D290–301. doi: [10.1093/nar/gkr1065](https://doi.org/10.1093/nar/gkr1065) PMID: [22127870](https://pubmed.ncbi.nlm.nih.gov/22127870/)
21. Stadtman ER. The purification and properties of phosphotransacetylase. *J Biol Chem*. 1952; 196(2):527–34. PMID: [12980995](https://pubmed.ncbi.nlm.nih.gov/12980995/).
22. Rado TA, Hoch JA. Phosphotransacetylase from *Bacillus subtilis*: purification and physiological studies. *Biochim Biophys Acta*. 1973; 321(1):114–25. PMID: [4201530](https://pubmed.ncbi.nlm.nih.gov/4201530/).
23. Smith JE, Ng WS. Fluorometric determination of glycolytic intermediates and adenylates during sequential changes in replacement culture of *Aspergillus niger*. *Can J Microbiol*. 1972; 18(11):1657–64. PMID: [4263885](https://pubmed.ncbi.nlm.nih.gov/4263885/).
24. Aussignargues C, Paasch BC, Gonzales-Esquer R, Erbilgin O, Kerfeld CA. Bacterial Microcompartment Assembly: The Key Role of Encapsulation Peptides. *Communicative & Integrative Biology*. 2015; 8(3):e1039755.
25. Fan C, Cheng S, Sinha S, Bobik Ta. Interactions between the termini of lumen enzymes and shell proteins mediate enzyme encapsulation into bacterial microcompartments. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109(37):14995–5000. doi: [10.1073/pnas.1207516109](https://doi.org/10.1073/pnas.1207516109) PMID: [22927404](https://pubmed.ncbi.nlm.nih.gov/22927404/)
26. Kinney JN, Salmeen A, Cai F, Kerfeld CA. Elucidating the essential role of the conserved carboxysomal protein CcmN reveals a common feature of bacterial microcompartment assembly. *The Journal of biological chemistry*. 2012; 287(21):17729–36. doi: [10.1074/jbc.M112.355305](https://doi.org/10.1074/jbc.M112.355305) PMID: [22461622](https://pubmed.ncbi.nlm.nih.gov/22461622/)
27. Tobimatsu T, Kawata M, Toraya T, Obimatsu TT, Awata MK, Oraya TT. The N-Terminal Regions of  $\beta$  and  $\gamma$  Subunits Lower the Solubility of Adenosylcobalamin-Dependent Diol Dehydratase. *Bioscience, Biotechnology, and Biochemistry*. 2005; 69(3):455–62. PMID: [15784971](https://pubmed.ncbi.nlm.nih.gov/15784971/)
28. Shibata N, Tamagaki H, Hieda N, Akita K, Komori H, Shomura Y, et al. Crystal structures of ethanolamine ammonia-lyase complexed with coenzyme B12 analogs and substrates. *The Journal of biological chemistry*. 2010; 285(34):26484–93. doi: [10.1074/jbc.M110.125112](https://doi.org/10.1074/jbc.M110.125112) PMID: [20519496](https://pubmed.ncbi.nlm.nih.gov/20519496/)
29. Frank S, Lawrence AD, Prentice MB, Warren MJ. Bacterial microcompartments moving into a synthetic biological world. *Journal of biotechnology*. 2013; 163(2):273–9. doi: [10.1016/j.jbiotec.2012.09.002](https://doi.org/10.1016/j.jbiotec.2012.09.002) PMID: [22982517](https://pubmed.ncbi.nlm.nih.gov/22982517/)
30. Gonzalez-Esquer CR, Shubitowski TB, Kerfeld CA. Streamlined construction of the cyanobacterial CO<sub>2</sub>-fixing organelle via protein domain fusions. *Plant Cell*. in press.
31. Lawrence SH, Luther KB, Schindelin H, Ferry JG. Structural and functional studies suggest a catalytic mechanism for the phosphotransacetylase from *Methanosarcina thermophila*. *J Bacteriol*. 2006; 188(3):1143–54. doi: [10.1128/JB.188.3.1143-1154.2006](https://doi.org/10.1128/JB.188.3.1143-1154.2006) PMID: [16428418](https://pubmed.ncbi.nlm.nih.gov/16428418/); PubMed Central PMCID: PMC1347337.
32. Erbilgin O, McDonald KL, Kerfeld Ca. Characterization of a planctomycetal organelle: a novel bacterial microcompartment for the aerobic degradation of plant saccharides. *Applied and environmental microbiology*. 2014; 80(7):2193–205. doi: [10.1128/AEM.03887-13](https://doi.org/10.1128/AEM.03887-13) PMID: [24487526](https://pubmed.ncbi.nlm.nih.gov/24487526/)
33. Chakrabarti P. Geometry of interaction of metal ions with histidine residues in protein structures. *Protein engineering*. 1990; 4(1):57–63. PMID: [2290835](https://pubmed.ncbi.nlm.nih.gov/2290835/).

34. 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(14):2392–9. doi: [10.1128/JB.00056-15](https://doi.org/10.1128/JB.00056-15) PMID: [25962918](https://pubmed.ncbi.nlm.nih.gov/25962918/).
35. Kloer DP, Hagel C, Heider J, Schulz GE. Crystal structure of ethylbenzene dehydrogenase from *Aromatoleum aromaticum*. *Structure.* 2006; 14(9):1377–88. doi: [10.1016/j.str.2006.07.001](https://doi.org/10.1016/j.str.2006.07.001) PMID: [16962969](https://pubmed.ncbi.nlm.nih.gov/16962969/).
36. Cai F, Dou Z, Bernstein SL, Leverenz R, Williams EB, Heinhorst S, et al. Advances in Understanding Carboxysome Assembly in *Prochlorococcus* and *Synechococcus* Implicate CsoS2 as a Critical Component. *Life (Basel).* 2015; 5(2):1141–71. doi: [10.3390/life5021141](https://doi.org/10.3390/life5021141) PMID: [25826651](https://pubmed.ncbi.nlm.nih.gov/25826651/).
37. Cameron JC, Wilson SC, Bernstein SL, Kerfeld CA. Biogenesis of a Bacterial Organelle: The Carboxysome Assembly Pathway. *Cell.* 2013; 155(5):1131–40. doi: [10.1016/j.cell.2013.10.044](https://doi.org/10.1016/j.cell.2013.10.044) PMID: [24267892](https://pubmed.ncbi.nlm.nih.gov/24267892/)
38. Lawrence AD, Frank S, Newnham S, Lee MJ, Brown IR, Xue W-F, et al. Solution structure of a bacterial microcompartment targeting peptide and its application in the construction of an ethanol bioreactor. *ACS Synthetic Biology.* 2014; 140130103032007. doi: [10.1021/sb4001118](https://doi.org/10.1021/sb4001118)
39. Cleland WW, Hengge AC. Enzymatic mechanisms of phosphate and sulfate transfer. *Chem Rev.* 2006; 106(8):3252–78. doi: [10.1021/cr050287o](https://doi.org/10.1021/cr050287o) PMID: [16895327](https://pubmed.ncbi.nlm.nih.gov/16895327/).
40. Kimber MS, Pai EF. The active site architecture of *Pisum sativum* beta-carbonic anhydrase is a mirror image of that of alpha-carbonic anhydrases. *EMBO J.* 2000; 19(7):1407–18. doi: [10.1093/emboj/19.7.1407](https://doi.org/10.1093/emboj/19.7.1407) PMID: [10747009](https://pubmed.ncbi.nlm.nih.gov/10747009/); PubMed Central PMCID: PMC310211.
41. Tripp BC, Smith K, Ferry JG. Carbonic anhydrase: new insights for an ancient enzyme. *J Biol Chem.* 2001; 276(52):48615–8. doi: [10.1074/jbc.R100045200](https://doi.org/10.1074/jbc.R100045200) PMID: [11696553](https://pubmed.ncbi.nlm.nih.gov/11696553/).
42. Carfi A, Pares S, Duce E, Galleni M, Duez C, Frere JM, et al. The 3-D structure of a zinc metallo-beta-lactamase from *Bacillus cereus* reveals a new type of protein fold. *EMBO J.* 1995; 14(20):4914–21. PMID: [7588620](https://pubmed.ncbi.nlm.nih.gov/7588620/); PubMed Central PMCID: PMC394593.
43. Drawz SM, Bonomo RA. Three decades of beta-lactamase inhibitors. *Clin Microbiol Rev.* 2010; 23(1):160–201. doi: [10.1128/CMR.00037-09](https://doi.org/10.1128/CMR.00037-09) PMID: [20065329](https://pubmed.ncbi.nlm.nih.gov/20065329/); PubMed Central PMCID: PMC2806661.
44. Peña KL, Castel SE, de Araujo C, Espie GS, Kimber MS. Structural basis of the oxidative activation of the carboxysomal gamma-carbonic anhydrase, CcmM. *Proceedings of the National Academy of Sciences of the United States of America.* 2010; 107(6):2455–60. doi: [10.1073/pnas.0910866107](https://doi.org/10.1073/pnas.0910866107) PMID: [20133749](https://pubmed.ncbi.nlm.nih.gov/20133749/)
45. Chen AH, Robinson-Mosher A, Savage DF, Silver Pa, Polka JK. The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo. *PLoS ONE.* 2013; 8(9):e76127–e. doi: [10.1371/journal.pone.0076127](https://doi.org/10.1371/journal.pone.0076127) PMID: [24023971](https://pubmed.ncbi.nlm.nih.gov/24023971/)
46. Harvey PC, Watson M, Hulme S, Jones Ma, Lovell M, Berchieri A, et al. *Salmonella enterica* typhimurium colonizing the lumen of the chicken intestine grows slowly and upregulates a unique set of virulence and metabolism genes. *Infection and immunity.* 2011; 79(10):4105–21. doi: [10.1128/IAI.01390-10](https://doi.org/10.1128/IAI.01390-10) PMID: [21768276](https://pubmed.ncbi.nlm.nih.gov/21768276/)
47. Joseph B, Przybilla K, Stu C, Schauer K, Fuchs TM, Goebel W, et al. Identification of *Listeria monocytogenes* Genes Contributing to Intracellular Replication by Expression Profiling and Mutant Screening. *Journal of bacteriology.* 2006; 188(2):556–68. PMID: [16385046](https://pubmed.ncbi.nlm.nih.gov/16385046/)
48. Kendall MM, Gruber CC, Parker CT, Sperandio V. Ethanolamine controls expression of genes encoding components involved in interkingdom signaling and virulence in enterohemorrhagic *Escherichia coli* O157:H7. *mBio.* 2012; 3(3):1–10. doi: [10.1128/mBio.00050-12](https://doi.org/10.1128/mBio.00050-12)
49. Klumpp J, Fuchs TM. Identification of novel genes in genomic islands that contribute to *Salmonella typhimurium* replication in macrophages. *Microbiology.* 2007; 2:1207–20.
50. Maadani A, Fox KA, Mylonakis E, Garsin DA. *Enterococcus faecalis* mutations affecting virulence in the *Caenorhabditis elegans* model host. *Infection and immunity.* 2007; 75(5):2634–7. doi: [10.1128/IAI.01372-06](https://doi.org/10.1128/IAI.01372-06) PMID: [17307944](https://pubmed.ncbi.nlm.nih.gov/17307944/)
51. Mobley H. Redefining Virulence of Bacterial Pathogens. *Microbe.* 2015; 10(6):239–46.
52. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* 1976; 72:248–54. PMID: [942051](https://pubmed.ncbi.nlm.nih.gov/942051/).
53. Luzi S, Kondo Y, Bernard E, Stadler LK, Vaysburd M, Winter G, et al. Subunit disassembly and inhibition of TNFalpha by a semi-synthetic bicyclic peptide. *Protein Eng Des Sel.* 2015; 28(2):45–52. doi: [10.1093/protein/gzu055](https://doi.org/10.1093/protein/gzu055) PMID: [25614525](https://pubmed.ncbi.nlm.nih.gov/25614525/); PubMed Central PMCID: PMC4378371.
54. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research.* 2004; 32(5):1792–7. doi: [10.1093/nar/gkh340](https://doi.org/10.1093/nar/gkh340) PMID: [15034147](https://pubmed.ncbi.nlm.nih.gov/15034147/)
55. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics (Oxford, England).* 2009; 25(9):1189–91. doi: [10.1093/bioinformatics/btp033](https://doi.org/10.1093/bioinformatics/btp033)

56. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res.* 2004; 14(6):1188–90. doi: [10.1101/gr.849004](https://doi.org/10.1101/gr.849004) PMID: [15173120](https://pubmed.ncbi.nlm.nih.gov/15173120/); PubMed Central PMCID: PMC419797.
57. Kabsch W. XDS. *Acta crystallographica Section D, Biological crystallography.* 2010; 66(Pt 2):125–32. doi: [10.1107/S0907444909047337](https://doi.org/10.1107/S0907444909047337) PMID: [20124692](https://pubmed.ncbi.nlm.nih.gov/20124692/); PubMed Central PMCID: PMC2815665.
58. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. *Acta crystallographica Section D, Biological crystallography.* 2011; 67(Pt 4):235–42. Epub 2011/04/05. doi: [10.1107/S0907444910045749](https://doi.org/10.1107/S0907444910045749) PMID: [21460441](https://pubmed.ncbi.nlm.nih.gov/21460441/); PubMed Central PMCID: PMC3069738.
59. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D.* 2012; 68:352–67. doi: [10.1107/S0907444912001308](https://doi.org/10.1107/S0907444912001308) ISI:000302138400005. PMID: [22505256](https://pubmed.ncbi.nlm.nih.gov/22505256/)
60. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta crystallographica Section D, Biological crystallography.* 2004; 60(Pt 12 Pt 1):2126–32. Epub 2004/12/02. doi: [10.1107/S0907444904019158](https://doi.org/10.1107/S0907444904019158) PMID: [15572765](https://pubmed.ncbi.nlm.nih.gov/15572765/).
61. Merritt EA, Bacon DJ. Raster3D: photorealistic molecular graphics. *Methods in enzymology.* 1997; 277:505–24. PMID: [18488322](https://pubmed.ncbi.nlm.nih.gov/18488322/).
62. Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A. Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct.* 2000; 29:291–325. doi: [10.1146/annurev.biophys.29.1.291](https://doi.org/10.1146/annurev.biophys.29.1.291) PMID: [10940251](https://pubmed.ncbi.nlm.nih.gov/10940251/).