

# SCIENTIFIC REPORTS



OPEN

## *In Vitro* Characterization and Concerted Function of Three Core Enzymes of a Glycyl Radical Enzyme - Associated Bacterial Microcompartment

Received: 01 December 2016

Accepted: 13 January 2017

Published: 16 February 2017

Jan Zarzycki<sup>1</sup>, Markus Sutter<sup>2,3</sup>, Niña Socorro Cortina<sup>1</sup>, Tobias J. Erb<sup>1</sup> & Cheryl A. Kerfeld<sup>2,3,4,5,6</sup>

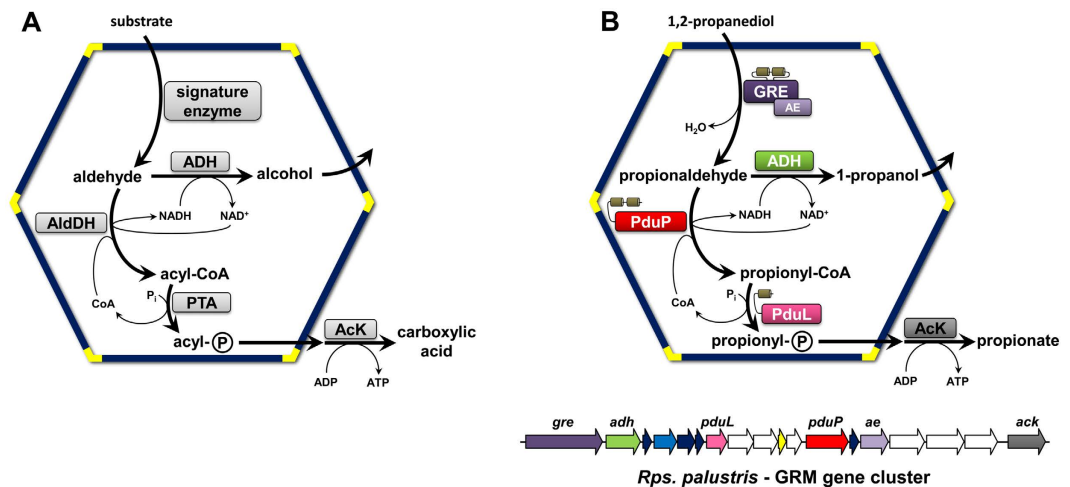
Many bacteria encode proteinaceous bacterial microcompartments (BMCs) that encapsulate sequential enzymatic reactions of diverse metabolic pathways. Well-characterized BMCs include carboxysomes for CO<sub>2</sub>-fixation, and propanediol- and ethanolamine-utilizing microcompartments that contain B<sub>12</sub>-dependent enzymes. Genes required to form BMCs are typically organized in gene clusters, which promoted their distribution across phyla by horizontal gene transfer. Recently, BMCs associated with glycyl radical enzymes (GREs) were discovered; these are widespread and comprise at least three functionally distinct types. Previously, we predicted one type of these GRE-associated microcompartments (GRMs) represents a B<sub>12</sub>-independent propanediol-utilizing BMC. Here we functionally and structurally characterize enzymes of the GRM of *Rhodopseudomonas palustris* BisB18 and demonstrate their concerted function *in vitro*. The GRM signature enzyme, the GRE, is a dedicated 1,2-propanediol dehydratase with a new type of intramolecular encapsulation peptide. It forms a complex with its activating enzyme and, in conjunction with an aldehyde dehydrogenase, converts 1,2-propanediol to propionyl-CoA. Notably, homologous GRMs are also encoded in pathogenic *Escherichia coli* strains. Our high-resolution crystal structures of the aldehyde dehydrogenase lead to a revised reaction mechanism. The successful *in vitro* reconstitution of a part of the GRM metabolism provides insights into the metabolic function and steps in the assembly of this BMC.

Bacterial microcompartments (BMCs) are polyhedral organelles found in a wide range of phyla<sup>1</sup>. BMCs are subcellular nanoreactors that typically encapsulate several enzymes catalyzing sequential steps of metabolic pathways. In contrast to membrane-delimited organelles, selective permeability of BMCs is conferred by a proteinaceous shell, which allows small polar and charged metabolites to pass, and appears to exclude passage of uncharged, non-polar molecules.

The majority of BMCs are involved in catabolism<sup>1,2</sup>. The general model<sup>2,3</sup> for the biochemical transformations within these ‘metabolosomes’ was derived from experimental characterization of a few catabolic BMCs<sup>3–6</sup> and bioinformatic analysis<sup>1,7</sup> (Fig. 1). Catabolic BMCs typically generate toxic and/or volatile aldehyde intermediates through the action of signature enzymes, which differ among functionally distinct BMCs. The subsequent conversion of the aldehyde is catalyzed by the other core enzymes: an acylating aldehyde dehydrogenase, an alcohol dehydrogenase, a phosphotransacylase and a carboxylic acid kinase<sup>1,2</sup> (Fig. 1).

A large and distinct group of the catabolic BMCs was recently discovered<sup>1</sup>. These BMCs contain glycyl radical enzymes (GREs) and their cognate activating enzymes (AEs) as signature enzymes. GRE-associated

<sup>1</sup>Max-Planck-Institute for Terrestrial Microbiology, Karl-von-Frisch-Str. 10, D-35043, Marburg, Germany. <sup>2</sup>MSU-DOE Plant Research Laboratory, Michigan State University, 612 Wilson Road, East Lansing, MI 48824, USA. <sup>3</sup>Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA. <sup>4</sup>Department of Biochemistry & Molecular Biology, Michigan State University, 603 Wilson Road, East Lansing, MI 48824, USA. <sup>5</sup>Berkeley Synthetic Biology Institute, Berkeley, CA, USA. <sup>6</sup>Department of Plant and Microbial Biology, University of California, 111 Koshland Hall, Berkeley, CA 94720, USA. Correspondence and requests for materials should be addressed to C.A.K. (email: ckerfeld@lbl.gov)



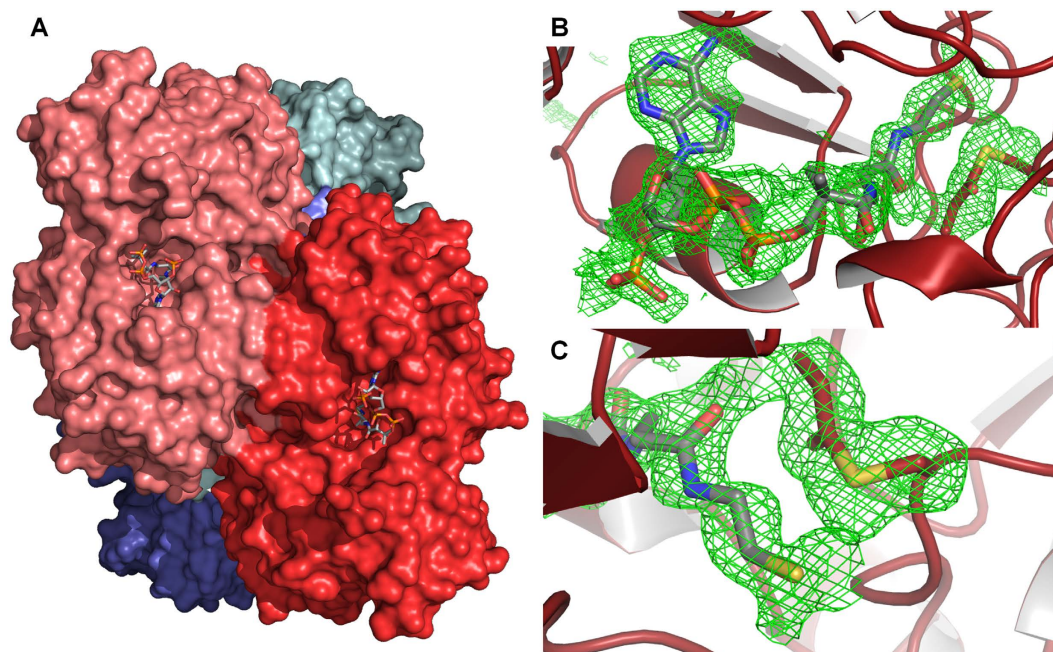
**Figure 1. Schematics of pathways in catabolic bacterial microcompartments and the propanediol utilizing GRMs.** (A) In the general model for metabolosome function (2, 3), the signature enzymes are involved in the formation of the aldehyde, which is subsequently oxidized to an acyl-CoA thioester by an acylating aldehyde dehydrogenase (AldDH), or reduced to an alcohol by an alcohol dehydrogenase (ADH). The acyl-CoA is then further converted to an acyl-phosphate by a phosphotransacylase (PTA) and finally to the free carboxylic acid by a kinase (AcK), generating ATP. The  $\text{NAD}^+$  and CoA required for the AldDH reaction are regenerated within the BMC by the ADH and PTA, respectively. (B) Specific enzymes encapsulated in GRM3: The GRE-type 1,2-propanediol dehydratase and its activating enzyme (dark and light purple, respectively) represent the signature enzymes. Encapsulation peptides are depicted as small beige cylinders. The GRM gene cluster of *Rps. palustris* is shown below. Genes encoding for enzymes that are involved in the encapsulated metabolic pathway are annotated and colored: aldehyde dehydrogenase (PduP), phosphotransacylase (PduL). Genes encoding different types of shell proteins are colored in dark blue, light blue, and yellow. Ancillary genes encoding proteins and enzymes that are not directly involved in the GRM metabolic pathway are white.

microcompartments (GRMs) appear to be diverse: at least three functionally different types were predicted by bioinformatic analysis<sup>1,7</sup>. GREs that function as choline trimethylamine-lyases<sup>8–12</sup>, producing trimethylamine (TMA) and acetaldehyde, are found associated with two microcompartment subtypes, GRM1 & GRM2 (nomenclature according to Zarzycki J. *et al.*<sup>7</sup>). Another type of GRM (GRM5) seems to function in fucose metabolism; the corresponding genes are expressed during the anaerobic degradation of fucose<sup>6</sup>. GRM5 was proposed to employ two additional enzymes, a fucose phosphate aldolase and a lactaldehyde reductase. Fucose phosphate can be cleaved into dihydroxyacetone phosphate and lactaldehyde. The latter would then be reduced to 1,2-propanediol, hence the associated GRE was proposed to function as a 1,2-propanediol dehydratase<sup>6</sup>. Indeed the GRE of a GRM5 from *Roseburia inulinivorans* has very recently been crystallized and shown to function as 1,2-propanediol dehydratase<sup>13</sup>, as was proposed for this type of GRM<sup>1,6,7</sup>.

A far more widespread subtype of the GRMs (GRM3), found in  $\alpha$ -proteobacteria,  $\gamma$ -proteobacteria, and firmicutes, was predicted to contain a GRE that also functions as a propanediol dehydratase<sup>1,7</sup>, but only shows 52% amino acid sequence identity to the GRM5 enzyme from *R. inulinivorans*. Here we systematically reconstruct a part of the metabolism of the GRM3 of *Rhodospseudomonas palustris* BisB18 by functionally and structurally characterizing three core enzymes, and by reconstituting the reaction sequence *in vitro*. We demonstrate that the GRE of *Rps. palustris* functions as a specific 1,2-propanediol dehydratase. Moreover, this signature enzyme is, to our knowledge, the first identified with an unusual intra-domain encapsulation peptide (EP) for packaging into the BMC. The *Rps. palustris* GRE binds to its AE, suggesting how the catalytic core of the organelle is assembled. Furthermore, we show that the acylating aldehyde dehydrogenase (AldDH), encoded within the GRM gene cluster, preferentially uses propionaldehyde as substrate. Moreover, we provide insights into the catalytic mechanism of AldDHs, by describing an acylated enzyme intermediate crystal structure, as well as clearly resolving the binding of CoA and propionyl-CoA in an acylating AldDH. By reconstitution of the core metabolism of the GRM3, we show that it functions analogously to the canonical PDU-BMCs, which employ  $\text{B}_{12}$ -dependent enzymes. This is the only known example of functional redundancy among distinct types of BMCs. Moreover, our results confirm that the general metabolic model for catabolic BMC function (Fig. 1A) also applies to the newly identified group of GRMs (Fig. 1B), the largest and most widespread family of BMCs detected to-date.

## Results

**Structural and Functional Analysis of the GRM3-associated Aldehyde dehydrogenase (AldDH).** After correcting the start site for the AldDH gene (see Methods), we found (using the criteria of Kinney J. N. *et al.*<sup>14</sup>) that the N-terminal 75 amino acids likely function as an EP<sup>7,15</sup>. The predicted EP of the AldDH comprises two amphipathic  $\alpha$ -helices, each 12–15 residues, separated by 20–25 unordered residues. The second helix connects to the enzymatic domain by another  $\sim 20$  amino acid, presumably flexible, linker<sup>7</sup>. Most of the experimentally identified or bioinformatically assigned EPs consist of one amphipathic  $\alpha$ -helix<sup>15</sup>.



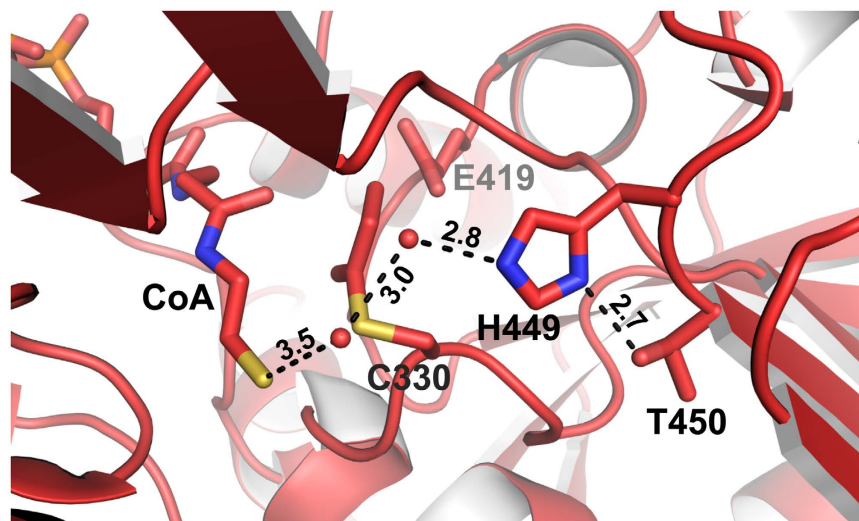
**Figure 2. The AldDH structure (PDB 5JFN) with bound CoA and acylated Cys330.** (A) Surface representation of the *Rps. palustris* AldDH, which forms a tetramer (dimer of dimers). A top view of a dimer interface with CoA molecules bound in the active site clefts is shown. (B) Superposition of an Fo-Fc electron density simulated annealing omit map at  $1.8 \sigma$  (green mesh) on refined CoA and S-propionylcysteine (Cys330) in the active site of the AldDH structure. (C) Close-up (from B) of the acylated Cys330.

Two different conditions were used to crystallize the AldDH with propionyl-CoA. At pH 8.1, AldDH crystallized in the monoclinic space group  $P 2_1$  with one tetramer per asymmetric unit (ASU). A tetrameric state was also reported for the AldDH homolog of *C. phytofermentans*<sup>16</sup>. This is consistent with our gel filtration analysis where AldDH eluted with an apparent mass of  $280 \pm 15$  kDa (55.5 kDa per monomer), the slightly higher molecular mass could be due to the flexible N-termini. The AldDH crystals diffracted to a resolution of 1.9 Å and the structure was solved by molecular replacement. The resulting model (PDB 5JFN) included all residues of the catalytic domains, as well as CoA or propionyl-CoA bound in the active site (Fig. 2A). No density was observed for the first 85 N-terminal residues comprising the putative EP and the Strep-tag. The only exception is a short unordered stretch of seven residues between the two  $\alpha$ -helices of the putative EP in one of the four chains, which is involved in crystal contacts. The resolved N-terminus of each of the four subunits each point in different directions, suggesting there is no interactions among the EPs.

Interestingly, we observed additional electron density around a cysteine residue in each of the putative active sites, which indicated a covalent modification. This cysteine residue was shown to be involved in catalysis and proposed to be acylated and deacylated in the course of the reaction<sup>17–19</sup>. Mutation of this cysteine into serine or alanine results in a loss of catalytic activity<sup>16,20,21</sup>. Because the enzyme was crystallized in the presence of propionyl-CoA at pH 8.1, we assumed that the corresponding cysteine residue was propionylated and modeled this residue as S-propionylcysteine. The covalent modification fit well into the electron density (Fig. 2B,C), to our knowledge the only acyl-enzyme thioester intermediate captured in a crystal structure of an acylating AldDH (pfam00171) to date. In three of the four active sites of the tetramer we also found coenzyme A bound in close proximity to the modified cysteine (Fig. 2B,C). In the fourth active site we were able to fit propionyl-CoA, which probably resulted from an additional soaking step before freezing the crystals.

When AldDH was crystallized at pH 4.8 with propionyl-CoA, we obtained crystals with triclinic  $P 1$  symmetry, diffracting to a resolution of 2.5 Å. The resulting model (PDB 5JFM) comprised two tetramers in the ASU, all chains lack electron density for the putative EPs. In the active sites we found electron density that allowed the fitting of either propionyl-CoA or CoA (Fig. S1A). Modification of the active site cysteine residue was not as apparent in this crystal form; this is likely due to the low pH, inhibiting the formation of a thiolate intermediate form of the cysteine<sup>20</sup>.

We used the same condition to co-crystallize AldDH with NAD<sup>+</sup>. These crystals had monoclinic  $P 2_1$  symmetry, diffracted to a resolution of 2.3 Å and exhibited twinning (twin fraction of 37%). The resulting model (PDB 5JFL) comprises two tetramers in the ASU, again all chains lacking the putative EPs. Electron density in the active sites allowed the fitting of NAD<sup>+</sup> (Fig. S1B). Interestingly, both co-substrates, NAD<sup>+</sup> and (propionyl-) CoA, occupy a similar position in the two structures (Fig. S2) but with their adenine rings pointing in opposite directions. This was also reported for crystal structures of the AldDH homolog from *C. phytofermentans*<sup>16</sup>, and is consistent with a proposed bi-uni-uni-ping pong mechanism<sup>16–19</sup>, in which the aldehyde would bind to the active site together with NAD<sup>+</sup>. Subsequently NADH is released from the active site before CoA can bind and the final product, acyl-CoA, is released again<sup>16,22</sup>. In contrast to the AldDH structure from *C. phytofermentans*



**Figure 3. Representation of the active site of the acylating propionaldehyde dehydrogenase with bound CoA and S-propionylcysteine.** Depicted is the proposed proton transfer relay, comprising Thr450, His449 as well as two structurally conserved water molecules, likely responsible for the deprotonation of CoA. Numbers indicate distances in Å.

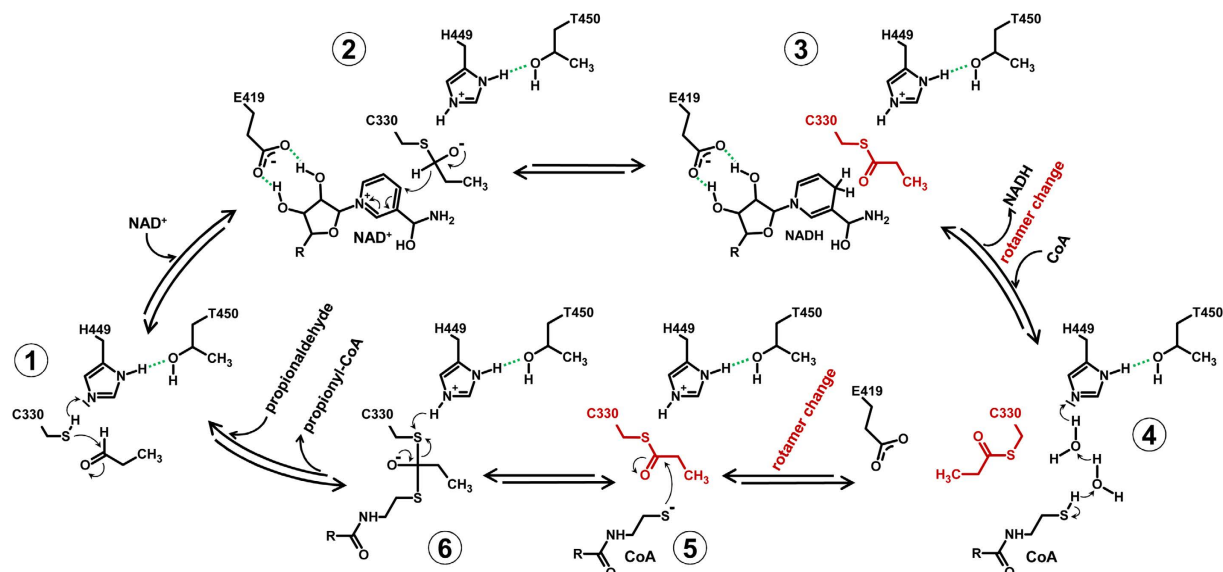
we were able to resolve the pantotheine moiety of the bound CoA as well as the propionyl thioester part of propionyl-CoA. Interestingly, we observed that the S-propionylcysteine assumes a different rotamer conformation than the non-acylated cysteine; this has implications for the reaction mechanism. The observed rotation prevents the interaction of the CoA with a conserved glutamate residue that was recently proposed to be responsible for the deprotonation of the CoA thiol<sup>16</sup>. The rotation of the S-propionylcysteine also opens up space for a proton transfer relay that could serve in the deprotonation of the CoA thiol (Fig. 3).

The proposed proton transfer relay involves Thr450 and His449, as well as two ordered water molecules. His449 is locked in the optimal tautomeric form for proton abstraction from the CoA thiol by hydrogen bonding to Thr450 (Figs 3 and 4). The corresponding histidine in the *C. phytofermentans* AldDH was previously proposed to catalyze the deprotonation of the catalytic cysteine residue<sup>16</sup>. We therefore suggest that this residue functions in two proton abstractions: from the catalytic cysteine (Cys330) and from the CoA thiol (Fig. 4). Both His449 and Thr450 are conserved among AldDHs, although the threonine is occasionally replaced by a serine (as in the *C. phytofermentans* AldDH). The overall reaction mechanism is discussed below. Data collection and refinement statistics are provided in Table 1.

To corroborate the hypothesis that the GRM3 of *Rps. palustris* functions in propanediol metabolism and produces propionaldehyde as an intermediate, we biochemically characterized the AldDH. We determined its kinetic parameters for propionaldehyde and acetaldehyde. The enzyme exhibited a higher maximal specific activity with propionaldehyde ( $V_{\max} = 242 \pm 14 \mu\text{mol mg}^{-1} \text{min}^{-1}$ ) than with acetaldehyde ( $V_{\max} = 117 \pm 8 \mu\text{mol mg}^{-1} \text{min}^{-1}$ ). Furthermore, the apparent  $K_m$ -value for propionaldehyde was about four times lower than for acetaldehyde ( $1.2 \pm 0.17 \text{ mM}$  and  $4.1 \pm 0.7 \text{ mM}$ , respectively). These values translate into a catalytic efficiency ( $k_{\text{cat}}/K_m$ ) that is almost one order of magnitude higher with propionaldehyde ( $745.8 \times 10^3 \text{ M}^{-1} \text{s}^{-1}$ ) compared to acetaldehyde ( $105.6 \times 10^3 \text{ M}^{-1} \text{s}^{-1}$ ). The enzyme also exhibited highest activity with propionyl-CoA when measuring the activity in the reverse direction (reduction of acetyl-CoA, propionyl-CoA, and 3-hydroxypropionyl-CoA) (Table 2).

### Characterization of the GRM3 Signature Enzymes: The GRE and its Activating enzyme (AE).

When we anaerobically expressed a His-tagged version of the activating enzyme (AE) in *E. coli* we noticed that the C-terminally His6-tagged AE eluted at very low imidazole concentrations (50 mM) from the Ni-NTA resin, suggesting that the His-tag was not completely exposed. In order to determine if the AE contained [4Fe-4S] clusters, as predicted from the primary structure, the UV-VIS features of purified AE were analyzed in the absence and presence of DTT, dithionite, and oxygen. Moreover, we compared AE samples after [Fe-S] cluster reconstitution to untreated (*i.e.*, as purified) samples. Both, reconstituted and non-reconstituted samples, showed an absorbance band at around 400 nm, indicative of [4Fe-4S] clusters, whereas absorbance bands indicative of [2Fe-2S] or [3Fe-4S] clusters were not observed (Fig. 5). Reconstitution of the [Fe-S] cluster did not have a significant effect on the AE. This was evident by estimating the iron content in the protein based on the assumption that the extinction coefficient at 400 nm for the non-reduced [4Fe-4S] cluster is about  $4,000 \text{ cm}^{-1} \text{M}^{-1}$  per iron atom<sup>23</sup>. For the reconstituted and the non-reconstituted AE, values of 13.1 and 12.3 iron atoms per monomer were determined. DTT (reduction potential of  $-330 \text{ mV}$ ) was not able to completely reduce the [Fe-S] clusters of the AE. A complete reduction was only achieved with dithionite (reduction potential of  $-660 \text{ mV}$ ) resulting in the loss of the UV-VIS features for the reconstituted and non-reconstituted samples. After exposing the reduced samples to ambient oxygen concentrations, the signal at 400 nm was recovered. However, after prolonged exposure to oxygen the signals were lost and the protein started to precipitate, underscoring the oxygen sensitivity of the AE.



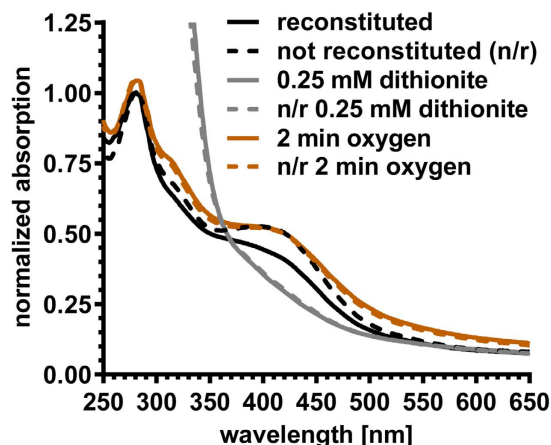
**Figure 4. Schematic of the proposed reaction mechanism for the acylating propionylaldehyde dehydrogenase.** Important hydrogen bonds are depicted in green. The switching between two different rotamer conformations of the acylated Cys330 is highlighted in red.

Ligands	CoA and S-propionylcysteine	Propionyl-CoA	NAD <sup>+</sup>
PDB ID	5JFN	5JFM	5JFL
Space group (No)	<i>P</i> 1 2 <sub>1</sub> 1 (4)	<i>P</i> 1 (1)	<i>P</i> 1 2 <sub>1</sub> 1 (4)
Cell dimensions			
<i>a b c</i> [Å]	83.19 106.45 125.73	92.11 105.78 126.43	107.39 154.80 126.59
$\alpha \beta \gamma$ [°]	90.00 109.03 90.00	89.52 71.00 68.92	90.00 90.13 90.00
Resolution [Å]	39.2–1.90 (2.00–1.90)	39.26–2.52 (2.65–2.52)	33.64–2.30 (2.42–2.30)
Number of observations			
total	716,571	308,204	1,228,225
Unique	162,389	135,429	181,023
Redundancy	4.4 (4.4)	2.3 (2.2)	7.1 (6.8)
<i>I</i> / $\sigma$ <sup>2</sup>	12.8 (2.1)	11.3 (2.3)	10.9 (2.7)
Completeness [%]	99.7 (98.3)	95.5 (91.8)	98.9 (98.0)
Refinement			
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	16.2/18.7	17.2/22.3	18.9/22.9
No. atoms	15,377	28,030	27,656
Protein	13,245	26,527	26,386
Ligand/ion	348	412	352
Water	1784	1091	827
Average B-factors	22.1	37.5	38.9
Protein	20.1	37.1	39.0
Ligand/ion	44.7	64.8	45.3
Water	33.1	36.3	35.6
R.m.s deviations			
Bond lengths [Å]	0.004	0.006	0.004
Bond angles [°]	0.69	0.78	0.62
Ramachandran			
favored	98.9	98.5	96.4
allowed	1.10	1.45	3.50
outliers	0.00	0.06	0.14

**Table 1. Data collection and refinement statistics for AldDH crystal structures.** Numbers in parentheses indicate statistics for highest resolution shell.

	$V_{\max}$ [ $\mu\text{mol min}^{-1} \text{mg}^{-1}$ ]	$k_{\text{cat}}$ [ $\text{s}^{-1}$ ]*	apparent $K_M$ [mM]	$k_{\text{cat}}/K_M$ [ $\text{M}^{-1} \text{s}^{-1}$ ]*
acetaldehyde	117 ± 8	433	4.1 ± 0.7	105.6 × 10 <sup>3</sup>
propionaldehyde	242 ± 14	895	1.2 ± 0.17	745.8 × 10 <sup>3</sup>
acetyl-CoA <sup>†</sup>	5.3 ± 0.3	19.6 ± 1.1	n.d.	n.d.
propionyl-CoA <sup>†</sup>	31.6 ± 1.9	116.9 ± 7.0	n.d.	n.d.
3OHP-CoA <sup>†</sup>	2.4 ± 0.05	8.9 ± 0.2	n.d.	n.d.

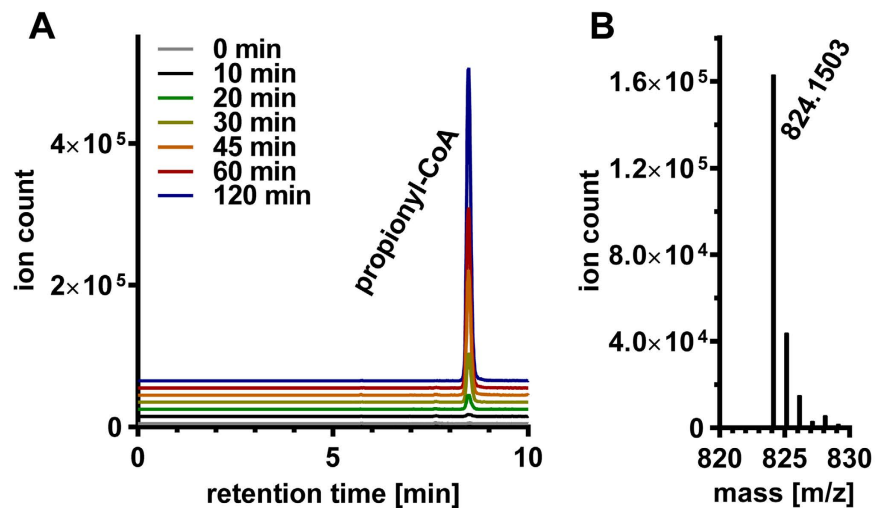
**Table 2. Catalytic properties of the recombinant aldehyde dehydrogenase from *Rps. palustris* (AldDH).**  
\*Per tetrameric enzyme, <sup>†</sup>at 250  $\mu\text{M}$ , n.d. – not determined.



**Figure 5. UV-VIS spectra of the [Fe-S] cluster containing activating enzyme.** Reconstituted (solid lines) and non-reconstituted (broken lines) samples of the activating enzyme exhibited identical spectra in the completely dithionite reduced forms (grey), as well as in the re-oxidized form after exposure to ambient oxygen (orange). The reconstituted sample (black solid line) differs from the as prepared sample (broken black line) due to partial reduction with DTT. The spectra were normalized to the absorbance at 280 nm of the reconstituted and non-reconstituted as prepared samples (black lines).

A recombinant His-tagged version of the GRE formed inclusion bodies when produced anaerobically in *E. coli*. Only a small fraction of the protein could be purified in a soluble form, which started to precipitate moments after elution from the nickel-affinity column. BMC-associated enzymes are often insoluble due to the EP; its removal frequently results in the production of soluble protein<sup>16,24–26</sup>. EPs are required for BMC assembly<sup>14,15,27,28</sup>, but in contrast to other BMC-associated GREs the GRM3 type does not possess an obvious N- or C-terminal EP. Inspection of the primary structure of GRM3 GRE suggested that the EP was instead located intramolecularly<sup>7</sup>. Bioinformatic analysis predicted an insertion domain that comprises two  $\alpha$ -helices (~15 residues each), which are connected to one another via a short loop (~5 residues) and are flanked by poorly conserved, unordered linkers (~15 residues on each side). One of the predicted helices shows the characteristic amphipathicity of EPs<sup>7</sup>. Interestingly, the expected topology of this EP is reminiscent of the putative EP of the GRM3 AldDH (see above and Zarzycki J. *et al.*<sup>7</sup>). The putative inter-domain EP of the GRM3 GRE is missing in the closely related B<sub>12</sub>-independent glycerol dehydratase<sup>7</sup> (GDH) of *Clostridium butyricum*<sup>22,29</sup>. The primary structure of this enzyme is 55% identical to the *Rps. palustris* GRE, but is not associated with any BMC. We therefore designed a truncated version of the *Rps. palustris* GRE lacking the two putative intramolecular EP-helices (Fig. S3). Indeed, expression of the EP-less version of the GRE yielded more soluble protein. Although the purified enzyme still precipitated over time, this occurred more slowly (after several days at 4 °C) relative to the full-length enzyme. Interestingly, the predicted intramolecular EP of the GRM5 GRE<sup>7</sup> from *R. inulinivorans* did not appear to cause precipitation in a recent study, but was, as noted by the authors, part of an oligomerization interface not previously observed in other GRE structures<sup>13</sup>. The amino acid sequences of the predicted intramolecular EPs of the GRM3 and GRM5 GREs are not strongly conserved and could therefore contribute to different characteristics in solution. However, the authors did not mention the GRE's involvement in a microcompartment and did not acknowledge the possible of this additional inter-protein domain being an EP, and thus could not assign any function to it<sup>13</sup>.

Nevertheless, when we combined the *E. coli* cell pellets from truncated GRE and AE overexpressions to co-purify both enzymes via NTA-affinity chromatography, we noticed that the AE no longer eluted at low imidazole concentrations as observed before when it was purified in isolation. The AE now co-eluted with the GRE only at high imidazole concentrations. The AE was bound to the GRE in solution, indicative of a 'piggy-back' recruitment of the AE to the BMC by the GRE, which likely becomes encapsulated via the proposed intrinsic inter-domain EP.



**Figure 6. Formation of propionyl-CoA followed by HRLC-MS.** (A) HRLC-MS extracted ion chromatogram of  $[M+H]^+$  following propionyl-CoA (calculated mass: 824.1487) formation (observed mass: 824.1503) due to the combined function of the GRE and the AldDH. (B) Mass-spectrum for propionyl-CoA.

Finally, to investigate the biochemistry of the GRE/AE complex, we examined whether the reductive cleavage of S-adenosylmethionine (SAM) during GRE activation resulted in the formation of 5'-deoxyadenosine (dA) or 5'-deoxy,5'-(methylthio)adenosine (MTA). In contrast to other SAM-dependent AEs, which produce dA, MTA formation had been reported for the AE of the GDH of *C. butyricum*<sup>30</sup>. In reconstituted GRM3 core reaction mixtures (see below) the formation of dA increased over time (Fig. S4). A background peak corresponding to the mass of MTA did not increase with time. Interestingly, the GRE of *Rps. palustris* shows a slightly higher percent identity in primary structure (55%) to the GDH of *C. butyricum* than to the 1,2-propanediol dehydratase from *R. inulinivorans* (52%). Despite all apparent similarities between the GRE and the AE of *Rps. palustris* as well as *R. inulinivorans* to their homologs in *C. butyricum* (GDH and its AE), there are distinct differences between the BMC-encapsulated and non-encapsulated enzymes with respect to structural motifs and biochemistry.

**Reconstitution of the GRM3 Core Reactions.** To demonstrate the metabolic function of the GRM3 core, a mixture of purified, truncated GRE, AE and AldDH was incubated in the presence of dithionite, 1,2-propanediol, CoA,  $NAD^+$ , and SAM. The mixtures were analyzed for the formation of propionyl-CoA by high resolution LC-MS. Propionyl-CoA was only detectable (Fig. 6A,B) when all (co-)substrates were present. No propionyl-CoA could be detected when 1,2-propanediol,  $NAD^+$ , or SAM were omitted from the reaction mixture. In these experiments propionyl-CoA could only have been formed by the AldDH from propionaldehyde; this was produced by the truncated GRE, which had been activated by AE. Moreover, when we substituted propanediol with glycerol as a substrate, no formation of the corresponding CoA-thioester, 3-hydroxypropionyl-CoA (3OHP-CoA) was detected. Because the AldDH was able to use 3OHP-CoA at fairly high rates when tested alone (Table 2), this indicates that the GRE, and thus the GRM3, specifically function in 1,2-propanediol, and not glycerol, degradation.

## Discussion

In conjunction with a recent structural and functional analysis of the novel BMC-associated phosphotransacylase of *Rps. palustris*<sup>24</sup> this study provides the comprehensive characterization of the catalytic core of a newly identified, widespread type of BMC. The functional characterization of the GRM3 increases the number of biochemically characterized BMCs to five and shows that GRMs also conform to the metabolic paradigm for catabolic BMCs<sup>2,7</sup> (Fig. 1). The *in vitro* reconstitution of the central reaction sequence of the metabolic pathway of the GRM3 of *Rps. palustris* established its function in 1,2-propanediol degradation (Fig. 1B), the same overall reactions as in the canonical PDU-BMC. This represents the only known example of functional redundancy among BMCs that employ distinctly different signature (aldehyde-generating) enzymes. To our knowledge, gene clusters encoding the GRM3 or PDU-BMC never co-occur in any genome<sup>7</sup>, underscoring their redundant functions. The fact that some organisms employ the PDU-BMC and others use the GRM3 for propanediol metabolism has led to the speculation that the mutually exclusive use of these two strategies might reflect different ecological niches occupied. In environments that are nutrient (e.g. cobalt) limited or energy restricted, the use of the GRMs over their  $B_{12}$ -dependent counterparts might be favored<sup>7</sup>.

Organisms that employ propanediol-utilizing GRMs comprise environmentally diverse microbes and pathogens, including *Clostridiales*, *Enterobacteriales*, *Vibrionales*, *Rhodospirillales*, *Rhodobacterales*, as well as *Rhizobiales*<sup>7</sup>, to which *Rps. palustris* belongs. It dwells in terrestrial and marine habitats and is able to live in anoxic sediments, where it likely uses 1,2-propanediol derived from the degradation of the plant and algal sugars fucose and rhamnose<sup>7</sup>. While *Rps. palustris* does not encode a conventional fucose phosphate aldolase, it may use a remote ortholog or be dependent on other organisms for the supply of propanediol<sup>7</sup>. Interestingly, genomes of at least eight different strains (GenBank) of *Rps. palustris* have been sequenced, however, only BisB18 harbors

the gene cluster encoding the BMC. Thus, the strain likely acquired the necessary genes in a relatively recent horizontal gene transfer event. This is also likely for other organisms. For example Some *E. coli* strains harbor a gene cluster encoding the canonical B<sub>12</sub>-dependent PDU BMC, whereas others encode the GRE-associated BMC, or neither. Indeed, BMC gene clusters appear to be very mobile within and across phyla<sup>1</sup>. Almost all of the *E. coli* strains that are able to employ one of these two types of BMCs are known pathogens<sup>31</sup>, indicating that both may contribute to pathogenicity and virulence in *Enterobacteriales* as has been suggested for other BMCs<sup>32</sup>. Therefore, components of these BMCs could be potential targets for antibiotic development.

Notably, the GRE characterized in this study shares a conserved active site architecture with GREs of the GRM5 subclass<sup>7</sup>. This close relationship is evidence of similar functions for GRM3 and GRM5. Besides GRE, AE and AldDH homologs, GRM5 additionally employs a canonical fuculose phosphate aldolase, as well as a putative lactaldehyde reductase<sup>6,7</sup>. Our data supports the proposed 1,2-propanediol metabolism of GRM5, in which the GRE converts propanediol produced by the sequential reactions of fuculose phosphate aldolase and lactaldehyde reductase. The propionaldehyde, is then further metabolized in the same way as in the canonical PDU-BMC and the GRM3<sup>6</sup>.

The capturing of the acyl-enzyme bound thioester intermediate in an AldDH crystal structure enables us to revisit the reaction mechanism for acylating AldDHs (Fig. 4). Upon binding of the aldehyde substrate to the active site of AldDH, a conserved histidine residue (here His449) abstracts the proton of the catalytic cysteine (Cys330), which can then attack the carbonyl carbon of the aldehyde (Step 1 in Fig. 4). In the next step, NAD<sup>+</sup> binds to the active site. A conserved glutamate (Glu419) forms hydrogen bonds to the nicotinamide ribose (Figs 3 and 4, Fig. S1B), which appears to facilitate the correct positioning of NAD<sup>+</sup> for the hydride transfer. After NADH leaves the active site, CoA can enter and bind while the acylated cysteine residue adopts another rotamer conformation (compare Figs 2 and 3 & Step 4 in Fig. 4). When rotated away, the acylated cysteine is located between Glu419 and the thiol group of CoA, which contradicts the proposed catalytic role of the conserved glutamate in CoA thiol proton abstraction<sup>16</sup>. Our observation is supported by two additional observations. First, mutation of a corresponding aspartate residue to an alanine in an AldDH from *Burkholderia xenovorans* did not abolish the enzyme's activity<sup>33</sup>. Second, the modeling of the CoA pantotheine moiety in the *C. phytofermentans* AldDH structure was speculative due the lack of electron density, which was pointed out by the authors<sup>16</sup>. Instead, we propose that His449 is also involved in the proton abstraction from CoA via two water molecules (water molecules 809 and 815 for the active site of chain B/5JFN, Fig. 3, Step 4 in Fig. 4). The rotamer switch of the cysteine provides the space for the water relay. Note that corresponding water molecules are also present in the *C. phytofermentans* AldDH structure with the bound CoA (PDB 5DBV), further evidence for a functional role. Hydrogen bonding between Thr450 and His449 ensures the correct tautomeric form for proton abstraction from the cysteine as well as from the first water of the proton relay (Figs 3 and 4). Supported by the position of the propionyl moiety in the structure with bound propionyl-CoA (Fig. S1A), the acylated cysteine would revert into the first rotamer conformation to be attacked by the activated CoA thiol, facilitating the acyl-CoA formation (Step 5 & 6 in Fig. 4) and completing the reaction cycle. In summary, characterization of the GRM3 AldDH leads to a revised and detailed reaction mechanism for the ubiquitous (present in all three domains of life) family of acylating aldehyde dehydrogenases. These enzymes take part in numerous pathways including fermentations, detoxifications, other catabolic and central carbon metabolic processes. Our findings apply beyond BMC function and will be of value for the bioengineering of these enzymes towards different substrates, particularly (toxic) aldehydes that may be formed in novel synthetic pathways for the microbial production of value-added chemicals<sup>34</sup>.

The addition of yet another function to the catalytic repertoire of GREs (propanediol dehydratases) in this as well as another recent study<sup>13</sup> underscores the versatility of this ancient and widespread enzyme family. Our findings demonstrate that GREs associated with BMCs can be divided into (at least two different types) choline TMA-lyases and 1,2-propanediol dehydratases. The apparent versatility of GREs is useful for understanding their potential roles in diverse microbes and for bioengineers striving to co-opt BMCs and particularly GRMs to fulfill novel, designed functions. A BMC shell may confer additional protection against molecular oxygen or reactive oxygen species for the extremely sensitive GREs and AEs and thereby extend the range of habitats wherein they may be useful to microbes. It is also important to note that the AldDH characterized in this study exhibits much higher rates of catalysis (~2 to 3 orders of magnitude) than other previously characterized orthologs<sup>16,34,35</sup>, and is able to accept not only acetyl-CoA but also 3-hydroxypropionyl-CoA, which is of interest to bioengineers aiming to produce 3-hydroxypropionate (a high value chemical building block) from glycerol<sup>34</sup>.

In addition to providing fundamental insights into the physiological function of GRMs, the structural and functional characterization presented here sheds light on novel aspects of the assembly and organization of the enzymatic core of the GRM3. It appears that the AldDH as well as the GRE contain novel types of EPs, which comprise two consecutive  $\alpha$ -helices. The exact mode of function of these EPs needs to be investigated further, as the putative intrinsic inter-domain EP of the GRM3 type GRE represents an intriguing and unprecedented example of an intra-protein EP and suggests that determinants for assembly into BMCs could be engineered between protein domains. The removal of the putative inter-domain EP resulted in production of an increased amount of soluble and yet catalytically active enzyme. GRM3 and perhaps other GRMs may assemble in an 'inside-out' fashion, where (some) of the encapsulated enzymes aggregate first, before the shell proteins are recruited<sup>2,15,36,37</sup>. The interaction between the AE and the GRE is likely responsible for the recruitment of the AE during BMC formation. In contrast, the EP of the AldDH did not lead to the aggregation of the enzyme, suggesting it may only interact with shell proteins. These results are consistent with the growing evidence that EPs can be involved in both core assembly and recruitment of the shell<sup>15,24,36</sup>. Collectively these observations are not only important for the fundamental understanding of BMC assembly, but also inform the design and construction of custom nanoreactors in synthetic biology<sup>38,39</sup>.



## Methods

**Chemicals.** Chemicals were purchased from Carl Roth and Sigma-Aldrich. Acetyl-CoA and propionyl-CoA were synthesized and purified from their respective carboxylic acid anhydrides<sup>40</sup>. 3-hydroxypropionyl-CoA was synthesized and purified from a mixed anhydride using carbonyldiimidazole as coupling agent<sup>40</sup>.

**Cloning of the aldehyde dehydrogenase, the glycol radical enzyme, and the activating enzyme from *Rps. palustris* BisB18.** We performed BLAST searches for each of the enzymes encoded in the BMC gene cluster of *Rps. palustris* BisB18. For AldDH, the aldehyde dehydrogenase, we noted that the primary structures of closely related homologs were about 50 residues longer, suggesting the start codon was mis-assigned in the GRM locus. Indeed, we found another start codon further up-stream, as well as an apparent Shine-Dalgarno sequence (AAGGAG) separated by 7 base pairs from that start codon. Instead of 464 amino acids (as in the GenBank record) the enzyme encoded by the revised full-length gene consisted of 512 amino acids.

The genes coding for the GRE, AE, and AldDH from *Rps. palustris* BisB18 were amplified using chromosomal DNA as the template. Standard PCRs were performed with Phusion<sup>®</sup> polymerase (ThermoFisher). The PCR products were cloned into vectors for expression in *Escherichia coli*. Used primers and the resulting plasmid constructs are listed in Table S1.

**Heterologous expression of the acylating aldehyde dehydrogenase (AldDH).** *E. coli* BL21 (DE3) cells containing the expression plasmid (pPduP\_Rp\_JZ73) for a Strep-tagged AldDH were grown in LB-medium. Ampicillin was added to cultures to concentrations of 100 µg/ml. Due to leaky expression, the cultures were grown for 24 h at 25 °C without IPTG induction.

**Heterologous expression of the glycol radical enzyme and its activating enzyme.** Cultures of *E. coli* BL21 (DE3) were grown in a MOPS/NaOH buffered (100 mM, pH 7.5) LB-medium. First, bottles with MOPS buffer were made anaerobic by sparging with nitrogen gas. The bottles were capped and transferred into an anaerobic glove box. The buffer was then supplemented with LB-broth (EMD Millipore). The bottles were sealed with rubber stoppers and autoclaved. The sterile media bottles were then anaerobically supplemented with 25 mM glucose, 25 mM sodium fumarate, and 100 µg/mL ampicillin from sterile stock solutions. For the activating enzyme, the medium was also supplemented with 1 mM L-cysteine, 1 mM ferric ammonium citrate. The bottles were then inoculated with 10 ml of an aerobically grown pre-culture of cells containing the expression plasmids for the glycol radical enzyme or the activating enzyme (pGRE\_Rp\_JZ15, pGRE\_Rp\_no-EP\_JZ88, or pGRE-AE\_Rp\_JZ16). The anaerobic cultures were grown at 30 °C with agitation (120 rpm). When the cultures reached an OD<sub>600nm</sub> of 1.0 the cultures were induced with 0.4 mM IPTG and grown for another 16 hours. The cells were harvested under anaerobic conditions.

**Purification of the aldehyde dehydrogenase (AldDH).** Cells were resuspended in a 3-fold volume of 50 mM MOPS/KOH buffer (pH 7.5) containing 150 mM KCl (buffer A) with 0.1 mg DNase I per ml. The cells were lysed by passage through a chilled French pressure cell at 137 MPa. The cell lysates were ultracentrifuged for 1 h at 100,000 × g at 4 °C and the supernatants were applied at a flow rate of 0.5 ml min<sup>-1</sup> to a 1 ml StrepTrap column (GE Healthcare) that had been equilibrated with buffer A. The column was washed with buffer A and the recombinant enzyme was eluted with the same buffer A with the addition of 2.5 mM desthiobiotin. Fractions containing AldDH activity were pooled and concentrated using a Amicon Ultra-4 Centrifugal Filter Unit (EMD Millipore) with a 30 kDa pore size. The sample was then applied to a 120 ml gel filtration column (HiLoad 16/600 Superdex 200 pg, GE Healthcare), equilibrated with 20 mM MOPS/KOH pH 7.5 buffer containing 50 mM KCl.

To render the enzyme anaerobic, it was applied to a 5 mL HiTrap Desalting column (GE Healthcare) that had been equilibrated with 100 mM MOPS/KOH at pH 7.5 containing 150 mM KCl under anaerobic conditions.

**Purification of the glycol radical enzyme and its activating enzyme.** Cells were resuspended in a 3-fold volume of 50 mM MOPS/KOH at pH 7.5 containing 250 mM KCl (buffer B) with 0.1 mg DNase I per mL and lysed using a French press (137 MPa) under anaerobic conditions. The lysates were centrifuged for one hour at 100,000 × g and 4 °C. All purification steps were performed in an anaerobic glove box. The 100,000 × g supernatants were applied to a 1 mL HisTrap HP column (GE Healthcare) that was equilibrated with buffer A. After loading of the sample the column was washed with buffer A.

For purification of the GRE the column was additionally washed with 80 mM imidazole in buffer B to remove unspecifically bound proteins. The enzyme was eluted with 500 mM imidazole in buffer A.

For the AE, the column was washed with 50 mM imidazole in buffer B to remove unwanted protein. The activating enzyme slowly eluted at the same concentration of imidazole in a second very broad peak.

To co-purify the GRE and AE, the harvested cell pellets were combined prior lysis and treated as above. After loading the sample to the HisTrap-column it was washed with 60 mM imidazole in buffer B. The GRE and AE co-eluted at 500 mM imidazole in buffer B.

Both the GRE and AE or the mixtures were concentrated using Amicon Ultra-4 Centrifugal Filter Units (EMD Millipore) with 10 kDa pore sizes. To remove the imidazole the enzymes were applied to a desalting column (HiTrap Desalting 5 mL, GE Healthcare) that had been equilibrated with 100 mM MOPS/KOH at pH 7.5 containing 150 mM KCl.

**Iron-sulfur cluster reconstitution for the activating enzyme.** After nickel affinity chromatography, half of the AE sample was incubated under anaerobic conditions with 5 mM DTT, 0.5 mM ammonium iron(II) sulfate, and 0.5 mM sodium sulfide for 12 h at 4 °C. To remove excess iron salts both the treated and untreated AE samples were applied to a HiLoad 16/600 Superdex 75 pg (GE Healthcare) gel filtration column equilibrated with

a 50 mM MOPS/KOH buffer (pH 7.5) containing 150 mM KCl. The purified proteins were then stored anaerobically at 4 °C until further use for UV-VIS photospectroscopy.

**UV-VIS photospectroscopy of the activating enzyme.** UV-VIS spectra of reconstituted and non-reconstituted AE were recorded anaerobically in rubber sealed quartz cuvettes using an Agilent Cary 60 spectrophotometer. Samples of reconstituted and non-reconstituted AE (in 50 mM MOPS/KOH buffer, pH 7.5, 150 mM KCl) were mixed with 2 mM DTT or 0.5 mM sodium dithionite and incubated for 2 min before recording the spectra. UV-VIS spectra were also recorded after exposing the AE samples to oxygen.

**Crystallization and structure determination of AldDH.** For co-crystallization of the purified Strep-tagged AldDH in the presence of propionyl-CoA two different conditions were used. 1) AldDH (2.4 mg mL<sup>-1</sup>) was mixed with 0.1 M HEPES/NaOH pH 7.0, 19% (w/v) PEG 4000, and 4 mM propionyl-CoA in a ratio of 2 μL:3 μL (enzyme to crystallization buffer). 2) AldDH (6 mg mL<sup>-1</sup>) was mixed with 50 μM sodium citrate pH 4.8, 4% (w/v) PEG 8000, and 5 mM propionyl-CoA in a ratio of 3 μL:6 μL (enzyme to crystallization buffer).

For co-crystallization in the presence of NAD<sup>+</sup> Strep-tagged AldDH (7.5 mg mL<sup>-1</sup>) was mixed with 50 μM sodium citrate pH 4.8, 4% (w/v) PEG 8000, and 5 mM NAD<sup>+</sup> in a ratio of 3 μL:6 μL (enzyme to crystallization buffer).

Crystals were soaked briefly with mother liquor supplemented with 35% (v/v) PEG 400, and 5–10 mM of propionyl-CoA or NAD<sup>+</sup>, respectively, before freezing in liquid nitrogen. X-ray diffraction data were collected at the Advanced Light Source (beamlines 5.0.2, 5.0.3) at Lawrence Berkeley National Laboratory.

The data was processed with the XDS software package<sup>41</sup>. The structures were solved by molecular replacement using the structure of a propionaldehyde dehydrogenase from *Clostridium phytofermentans*<sup>16</sup> (PDB 4C3S) as the search model. The molecular replacement was carried out using PhaserMR of the Phenix software package<sup>42</sup>. Initial models were built with Phenix.AutoBuild and refined with the phenix.refine. Additional modeling, manual refining and ligand fitting was done in COOT<sup>43</sup>. Final positional and B-factor refinements, as well as water-picking for the structures, were performed using phenix.refine.

**AldDH activity assay.** The activity of AldDH was routinely measured in photometric assays in the forward and reverse directions. For the forward direction the assay mixture (0.4 mL) contained 200 mM MOPS/KOH (pH 7.5), 5 mM NAD<sup>+</sup>, 0.5 mM CoA, and purified AldDH. The reaction was initiated by addition of propionaldehyde or acetaldehyde. The aldehyde and CoA-dependent reduction of NAD<sup>+</sup> was monitored at 340 nm. After completion of the reaction, 100 μL were withdrawn and enzyme was precipitated by addition of 5 μL 98% (v/v) formic acid. The precipitate was removed by centrifugation at 17,000 × g and 4 °C. The supernatants were analyzed by HRLC-MS to confirm the production of acetyl-CoA or propionyl-CoA, respectively.

For the reverse direction, the assay mixture (0.4 mL) contained 200 mM MOPS/KOH (pH 7.5), 0.4 mM NADH, 0.25 mM CoA-thioester (acetyl-CoA, propionyl-CoA or 3-hydroxypropionyl-CoA), and purified AldDH. The CoA-thioester dependent oxidation of NADH was monitored at 365 nm.

**Enzyme activity assay for the glycol radical enzyme under anaerobic conditions.** Co-purified truncated GRE and AE was incubated in 100 mM MOPS/KOH at pH 7.5 containing 150 mM KCl (final volume of 400 μL) in the presence of 50 mM (R/S)-1,2-propanediol (or 50 mM glycerol), 0.5 mM sodium dithionite, 1 mM SAM, 0.5 mM CoA, 2.5 mM NAD<sup>+</sup>, and anaerobic propionaldehyde dehydrogenase (AldDH). The reaction mixture was incubated at 16 °C for 2 hours. Samples of 50 μL were withdrawn after 0, 10, 20, 30, 45, 60, and 120 min. The reactions were stopped by addition of 5 μL of 98% (v/v) formic acid. Precipitated protein was removed by centrifugation at 17,000 × g and 4 °C. The supernatants were analyzed for the presence of propionyl-CoA, 3-hydroxypropionyl-CoA, and SAM cleavage products by HRLC-MS.

**High Resolution Liquid Chromatography-Mass Spectrography (HRLC-MS).** CoA, propionyl-CoA, and 3-hydroxypropionyl-CoA, SAM, MTA, and dA were analyzed using an Agilent 6550 iFunnel Q-TOF LC-MS system equipped with an electrospray ionization source set to positive ionization mode.

Compounds were separated on a RP-18 column (50 mm × 2.1 mm, particle size 1.7 μm, Kinetex XB-C18, Phenomenex) using a mobile phase system comprised of 50 mM ammonium formate pH 8.1 and methanol. Chromatographic separation was carried out using the following gradient condition at a flow rate of 250 μL/min: 0 min 0% methanol; 1 min 0% methanol; 3 min 2.5% methanol; 9 min 23% methanol; 14 min 80% methanol; 16 min 80% methanol.

Capillary voltage was set at 3.5 kV and nitrogen gas was used for nebulizing (20 psig), drying (13 l/min, 225 °C) and sheath gas (12 l/min, 400 °C). The TOF was calibrated using an ESI-L Low Concentration Tuning Mix (Agilent) before measurement (residuals less than 2 ppm for five reference ions) and was recalibrated during a run using 922 m/z as reference mass. MS data were acquired with a scan range of 200–1200 m/z.

LC-MS data were analyzed using MassHunter Qualitative Analysis software (Agilent).

## References

- Axen, S. D., Erbilgin, O. & Kerfeld, C. A. A taxonomy of bacterial microcompartment loci constructed by a novel scoring method. *PLoS Comp. Biol.* **10**, e1003898 (2014).
- Kerfeld, C. A. & Erbilgin, O. Bacterial microcompartments and the modular construction of microbial metabolism. *Trends Microbiol.* **23**, 22–34 (2015).
- Erbilgin, O., McDonald, K. L. & Kerfeld, C. A. Characterization of a planctomycetal organelle: a novel bacterial microcompartment for the aerobic degradation of plant saccharides. *Appl. Environ. Microbiol.* **80**, 2193–2205 (2014).
- Bobik, T. A., Havemann, G. D., Busch, R. J., Williams, D. S. & Aldrich, H. C. The propanediol utilization (*pdu*) operon of *Salmonella enterica* serovar Typhimurium LT2 includes genes necessary for formation of polyhedral organelles involved in coenzyme B<sub>12</sub>-dependent 1,2-propanediol degradation. *J. Bacteriol.* **181**, 5967–5975 (1999).

5. 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.* **181**, 5317–5329 (1999).
6. Petit, E. *et al.* Involvement of a bacterial microcompartment in the metabolism of fucose and rhamnose by *Clostridium phytofermentans*. *PLoS One* **8**, e54337 (2013).
7. Zarzycki, J., Erbilgin, O. & Kerfeld, C. A. Bioinformatic Characterization of Glycyl Radical Enzyme-Associated Bacterial Microcompartments. *Appl. Environ. Microbiol.* **81**, 8315–8329 (2015).
8. Craciun, S. & Balskus, E. P. Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. *Proc. Natl. Acad. Sci. USA* **109**, 21307–21312 (2012).
9. Craciun, S., Marks, J. A. & Balskus, E. P. Characterization of choline trimethylamine-lyase expands the chemistry of glycyl radical enzymes. *ACS Chem. Biol.* **9**, 1408–1413 (2014).
10. Martinez-del Campo, A. *et al.* Characterization and detection of a widely distributed gene cluster that predicts anaerobic choline utilization by human gut bacteria. *MBio* **6** (2015).
11. Kalnins, G. *et al.* Structure and function of CutC choline lyase from human microbiota bacterium *Klebsiella pneumoniae*. *J. Biol. Chem.* **290**, 21732–21740 (2015).
12. Bodea, S., Funk, M. A., Balskus, E. P. & Drennan, C. L. Molecular Basis of C-N Bond Cleavage by the Glycyl Radical Enzyme Choline Trimethylamine-Lyase. *Cell Chem. Biol.* **23**, 1206–1216 (2016).
13. LaMattina, J. W. *et al.* 1,2-Propanediol dehydration in *Roseburia inulinivorans*: Structural basis for substrate and enantiomer selectivity. *J. Biol. Chem.* **291**, 15515–15526 (2016).
14. Kinney, J. N., 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.* **287**, 17729–17736 (2012).
15. Aussignargues, C., Paasch, B. C., Gonzalez-Esquer, R., Erbilgin, O. & Kerfeld, C. A. Bacterial microcompartment assembly: The key role of encapsulation peptides. *Commun. Integr. Biol.* **8**, e1039755 (2015).
16. Tuck, L. R. *et al.* Insight into coenzyme A cofactor binding and the mechanism of acyl-transfer in an acylating aldehyde dehydrogenase from *Clostridium phytofermentans*. *Sci. Rep.* **6**, 22108 (2016).
17. Yan, C. C. & Fromm, H. J. Steady-state and pre-steady-state kinetics of coenzyme A linked aldehyde dehydrogenase from *Escherichia coli*. *Biochemistry* **20**, 7494–7501 (1981).
18. Smith, L. T. & Kaplan, N. O. Purification, properties, and kinetic mechanism of coenzyme A-linked aldehyde dehydrogenase from *Clostridium kluyveri*. *Arch. Biochem. Biophys.* **203**, 663–675 (1980).
19. Yan, R. T. & Chen, J. S. Coenzyme A-acylating aldehyde dehydrogenase from *Clostridium beijerinckii* NRRL B592. *Appl. Environ. Microbiol.* **56**, 2591–2599 (1990).
20. Cobessi, D., Tete-Favier, F., Marchal, S., Branlant, G. & Aubry, A. Structural and biochemical investigations of the catalytic mechanism of an NADP-dependent aldehyde dehydrogenase from *Streptococcus mutans*. *J. Mol. Biol.* **300**, 141–152 (2000).
21. Farres, J., Wang, T. T., Cunningham, S. J. & Weiner, H. Investigation of the active site cysteine residue of rat liver mitochondrial aldehyde dehydrogenase by site-directed mutagenesis. *Biochemistry* **34**, 2592–2598 (1995).
22. O'Brien, J. R. *et al.* Insight into the mechanism of the B<sub>12</sub>-independent glycerol dehydratase from *Clostridium butyricum*: preliminary biochemical and structural characterization. *Biochemistry* **43**, 4635–4645 (2004).
23. Sweeney, W. V. & Rabinowitz, J. C. Proteins containing 4Fe-4S clusters: an overview. *Annu. Rev. Biochem.* **49**, 139–161 (1980).
24. Erbilgin, O., Sutter, M. & Kerfeld, C. A. The structural basis of coenzyme A recycling in a bacterial organelle. *PLoS Biol.* **14**, e1002399 (2016).
25. Shibata, N. *et al.* Crystal structures of ethanolamine ammonia-lyase complexed with coenzyme B<sub>12</sub> analogs and substrates. *J. Biol. Chem.* **285**, 26484–26493 (2010).
26. Tobimatsu, T., Kawata, M. & Toraya, T. The N-terminal regions of beta and gamma subunits lower the solubility of adenosylcobalamin-dependent diol dehydratase. *Biosci. Biotechnol. Biochem.* **69**, 455–462 (2005).
27. Lawrence, A. D. *et al.* Solution structure of a bacterial microcompartment targeting peptide and its application in the construction of an ethanol bioreactor. *ACS Synth. Biol.* **3**, 454–465 (2014).
28. Fan, C., Cheng, S., Sinha, S. & Bobik, T. A. Interactions between the termini of lumen enzymes and shell proteins mediate enzyme encapsulation into bacterial microcompartments. *Proc. Natl. Acad. Sci. USA* **109**, 14995–15000 (2012).
29. Raynaud, C., Sarcabal, P., Meynial-Salles, I., Croux, C. & Soucaille, P. Molecular characterization of the 1,3-propanediol (1,3-PD) operon of *Clostridium butyricum*. *Proc. Natl. Acad. Sci. USA* **100**, 5010–5015 (2003).
30. Demick, J. M. & Lanzilotta, W. N. Radical SAM activation of the B<sub>12</sub>-independent glycerol dehydratase results in formation of 5'-deoxy-5'-(methylthio)adenosine and not 5'-deoxyadenosine. *Biochemistry* **50**, 440–442 (2011).
31. Jakobson, C. M. & Tullman-Ercek, D. Dumpster Diving in the Gut: Bacterial Microcompartments as Part of a Host-Associated Lifestyle. *PLoS Pathog.* **12**, e1005558 (2016).
32. Brzuszkiewicz, E. *et al.* How to become a uropathogen: comparative genomic analysis of extraintestinal pathogenic *Escherichia coli* strains. *Proc. Natl. Acad. Sci. USA* **103**, 12879–12884 (2006).
33. Baker, P., Carere, J. & Seah, S. Y. Substrate specificity, substrate channeling, and allostery in BphJ: an acylating aldehyde dehydrogenase associated with the pyruvate aldolase BphI. *Biochemistry* **51**, 4558–4567 (2012).
34. Luo, L. H. *et al.* Identification and characterization of the propanediol utilization protein PduP of *Lactobacillus reuteri* for 3-hydroxypropionic acid production from glycerol. *Appl. Microbiol. Biotechnol.* **89**, 697–703 (2011).
35. Leal, N. A., Havemann, G. D. & Bobik, T. A. PduP is a coenzyme-A-acylating propionaldehyde dehydrogenase associated with the polyhedral bodies involved in B<sub>12</sub>-dependent 1,2-propanediol degradation by *Salmonella enterica* serovar Typhimurium LT2. *Arch. Microbiol.* **180**, 353–361 (2003).
36. Lee, M. J., Brown, I. R., Juodeikis, R., Frank, S. & Warren, M. J. Employing bacterial microcompartment technology to engineer a shell-free enzyme-aggregate for enhanced 1,2-propanediol production in *Escherichia coli*. *Metab. Eng.* **36**, 48–56 (2016).
37. Cameron, J. C., Wilson, S. C., Bernstein, S. L. & Kerfeld, C. A. Biogenesis of a bacterial organelle: the carboxysome assembly pathway. *Cell* **155**, 1131–1140 (2013).
38. Gonzalez-Esquer, C. R., Newnham, S. E. & Kerfeld, C. A. Bacterial microcompartments as metabolic modules for plant synthetic biology. *Plant J.* **87**, 66–75 (2016).
39. Frank, S., Lawrence, A. D., Prentice, M. B. & Warren, M. J. Bacterial microcompartments moving into a synthetic biological world. *J. Biotechnol.* **163**, 273–279 (2013).
40. Peter, D. M., Vögeli, B., Cortina, N. S. & Erb, T. J. A chemo-enzymatic road map to the synthesis of CoA esters. *Molecules* **21** (2016).
41. Kabsch, W. Xds. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132 (2010).
42. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 (2010).
43. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).

## Acknowledgements

This work was supported by the National Institutes of Health, National Institute of Allergy and Infectious Diseases (NIAID) grant 1R01AI114975-01. J.Z. was additionally supported by FET686330. We thank the entire staff at the Advanced Light Source, Lawrence Berkeley National Laboratory, which is supported by the Director, Office of Science, Office of Basic Energy Sciences of the US Department of Energy (DOE) under contract DE-AC02-05CH11231. We thank Dr. Seigo Shima at the Max-Planck-Institute for Terrestrial Microbiology in Marburg for letting us use his anaerobic chamber and FPLC.

## Author Contributions

J.Z. performed the cloning of genes, expressed and purified the enzymes, performed enzyme assays, crystallized the AldDH, and determined its structure. N.S.C. performed liquid chromatography - high resolution mass spectroscopy and analyzed the data. M.S. collected X-ray diffraction data and refined the AldDH structure models. C.A.K. and J.Z. designed the study, interpreted the data and wrote the manuscript, with contributions from M.S., N.S.C. and T.J.E. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Zarzycki, J. *et al.* *In Vitro* Characterization and Concerted Function of Three Core Enzymes of a Glycyl Radical Enzyme - Associated Bacterial Microcompartment. *Sci. Rep.* **7**, 42757; doi: 10.1038/srep42757 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017