Title

A metabolic pathway for catabolizing levulinic acid in bacteria

Journal

Nature Microbiology, 2(12)

ISSN

2058-5276

Authors

Rand, Jacqueline M

Pisithkul, Tippapha

Clark, Ryan L

et al.

Publication Date

2017-12-01

DOI

10.1038/s41564-017-0028-z

Identification and characterization of the levulinic acid catabolic pathway in *Pseudomonas putida*

Jacqueline M. Rand[a], Tippapha Pisithkul[b], Ryan L. Clark[a], Joshua M. Thiede[a], Daniel E. Agnew[a], Candace E. Campbell[a], Andrew L. Markley[a], Morgan N. Price[c], Kelly M. Wetmore[c,e], Yumi Suh, Jayashree Ray, Adam P. Arkin[c,d], Adam M. Deutschbauer[c], Daniel Amador-Noguez[b,f], Brian F. Pfleger[a,b,#]

[a]Department of Chemical and Biological Engineering, University of Wisconsin-Madison

[b]Microbiology Doctoral Training Program, University of Wisconsin-Madison

[c]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory

[d]Department of Bioengineering, University of California, Berkeley

[e]Program in Comparative Biochemistry, University of California, Berkeley

[f]Department of Bacteriology, University of Wisconsin-Madison

[#]Address correspondence to Brian F. Pfleger, pfleger@engr.wisc.edu

Abstract

Microorganisms have the ability to catabolize a wide range of organic compounds and therefore the potential to perform many industrially relevant bioconversions. One barrier to realizing the full potential of bio-refining strategies lies in the incomplete knowledge of metabolic pathways, including those that can be used assimilate naturally abundant or easily generated feedstocks. For instance, levulinic acid (LA) is a carbon source that is readily obtainable as a dehydration product of lignocellulosic biomass molecules and can serve as the sole source of carbon for some bacteria. Despite the importance of LA catabolism for growth on certain biomass hydrolysates, both the genetics and intermediates of this pathway have remained unknown. Here, we report the identification and characterization of an operon responsible for LA catabolism in *Pseudomonas putida* KT2440. The seven-gene operon, designated *lva*, is under the control of a *prpR*-like regulator, and encodes five enzymatic proteins alongside two transporters. *In vitro* reconstitution of the pathway using purified enzymes demonstrated that LA is converted to 3-hydroxyvaleryl-CoA through the intermediates 4-hydroxyvaleryl-CoA and 4-phosphovaleryl-CoA. This discovery will allow for more efficient utilization of carbon from biomass hydrolysates and engineering of bacteria to perform novel bioconversions using LA as a feedstock.

Introduction

Levulinic acid (LA) is a five carbon γ-keto acid that can be readily obtained from biomass through non-enzymatic, acid hydrolysis of a wide range of feedstocks[1,2]. LA was named one of the US Department of Energy's "Top 12 value-added chemicals from biomass"[3] because it can be used as a renewable feedstock for generating a variety of molecules, such as fuel additives [4–10], flavors, fragrances[11,12] and polymers [13,14], through chemical catalysis. In addition, microbes can use LA as a sole carbon source and have been shown to convert LA into polyhydroxyalkanoates[15–18], short chain organic acids[19–21], and trehalose[19]. All of these bioconversion studies were conducted with natural bacterial isolates because the enzymes comprising a LA assimilation pathway were unknown[19]. This knowledge gap limits metabolic engineering and the potential of creating novel LA bioconversions.

While the enzymes responsible for LA assimilation were unknown at the time of these bioconversion demonstrations, other studies identified putative intermediates and suggested pathways for LA catabolism. In a study where crude cell lysates of *Cupriavidus necator* were fed LA, the concentration of LA and free CoA decreased over time while acetyl-CoA and propionyl-CoA concentrations increased, suggesting that LA is catabolized via CoA thioesters like other short-chain organic acids[22]. In a second study, cultures of *Pseudomonas putida* KT2440 expressing a heterologous TesB thioesterase were fed LA. Here, 4-hydroxyvalerate (4HV) and 3-hydroxyvalerate (3HV) transiently accumulated extracellularly before ultimately disappearing[23]. This observation strongly suggested that 4HV and 3HV (or their CoA thioesters) were pathway intermediates. Lastly, a metabolomic study of rat livers suggested that LA is catabolized to acetyl-CoA and propionyl-CoA via a unique phosphorylated acyl-CoA[24,25]. In sum, these

observations suggest a relatively direct route from LA to beta-oxidation intermediates, but the enzymes comprising such a pathway remain unknown.

In order to better utilize LA as a substrate for microbial growth or bioconversion, a detailed understanding of the metabolic pathway and enzymes involved is necessary. In this work, we investigated the genetic and biochemical factors that allow *P. putida* KT2440 to catabolize LA. Using a loss of function screen of a transposon library, we identified a putative LA utilization operon. The operon consists of seven genes, two homologs for membrane transporters and five enzymatic proteins. We reconstituted the pathway *in vitro* and determined that all five enzymatic proteins are required for complete conversion of LA into 3HV-CoA, an intermediate in the β-oxidation of odd-chain fatty acids. A closer inspection of the CoA ligase encoded in the operon revealed a broad substrate promiscuity including C_4 to C_6 organic acids. A putative regulator proximal to the operon activated transcription of the LA catabolic genes in the presence of LA or 4HV. The induction tests revealed that while the CoA ligase might have nonspecific activity towards similar chain length acids, the promoter is only responsive when cells were provided LA or 4HV. Altogether, the catabolism of LA to acetyl-CoA and propionyl-CoA requires at least 2 ATP that likely come from respiration and the TCA cycle.

Results

Identification of Genes Involved in Levulinic Acid Metabolism

P. putida KT2440 is known to metabolize LA as a sole carbon source and demonstrates diauxic growth in the presence of glucose and LA (**Supplementary Figure 1**). Therefore, we initiated a genetic study to identify genes involved in LA catabolism. We constructed a mutant library with a Tn5 mini transposase and screened for *P. putida* mutants lacking the ability to grow on LA as

the sole carbon source[26]. Thirteen out of 7,000 colonies screened demonstrated LA growth deficiencies. The location of each transposon insertion was determined by sequencing PCR products created with a primer nested in the transposon paired with a degenerate random primer. **Table 1** shows the ten unique isolates from these thirteen hits and the putative function of the disrupted genes. Two mutants had disruptions in genes involved in propionate metabolism, supporting the hypothesis that LA is catabolized to the central metabolites, acetyl-CoA and propionyl-CoA. Three transposon mutants had disruptions in a putative operon that had not been previously characterized (disrupting genes PP_2791, PP_2793, and PP_2794). Other mutants had disruptions in genes with no obvious connection to LA catabolism (*bioH*, *gcvP*, a hypothetical zinc protease, *mrdA*, and *fpvA*). To confirm that we had screened a sufficient number of clones, we performed random bar code transposon-site sequencing (RB-TnSeq) for cultures enriched by growth on LA and 4HV relative to growth on glucose. RB-TnSeq is an efficient method for determining gene essentiality under different conditions with high genomic coverage[27]. This analysis identified additional genes involved in LA metabolism including an acetoacetyl-CoA transferase important for growth on LA, genes functioning in β-oxidation and propionyl-CoA metabolism, and 14 transcriptional regulators potentially involved in LA metabolism. The RB-TnSeq dataset also revealed that 3-hydroxybutyryl-CoA dehydrogenase and β-ketothiolase are also necessary for growth on LA and 4HV, supporting our hypothesis that LA metabolism terminates through β-oxidation. For a more complete summary and analysis of the fitness data, please see **Supplementary Table 1** and **Supplementary Note:RB-TnSeq**.

Table 1. P. putida Levulinic Acid Transposon Insertion Sites

Locus	Insertion Point*	Gene Name	Description/Homology
PP_0364	442685	*bioH*	pimeloyl-ACP methyl ester esterase
PP_0988	1128706	*gcvP-1*	glycine dehydrogenase
PP_2332	2660666	N/A	ATP-dependent zinc protease family
PP_2336	2666405	*acnA-II*	aconitate hydratase
PP_2337	2666944	*prpF*	aconitate isomerase
PP_2791	3181098	N/A	Phosphotransferase family
PP_2793	3182533	N/A	acyl-CoA dehydrogenase family protein
PP_2794	3183601	N/A	short chain dehydrogenase/reductase family
PP_3741	4271628	*mrdA-I*	transpeptidase
PP_4217	4765953	*fpvA*	TonB-dependent outer membrane ferripyoverdine receptor

*Insertion point based on location from *P. putida* KT2440 origin

Operon Characterization and Induction

Given the propensity of bacteria to cluster related genes into operons, we examined the putative seven-gene operon, PP_2791-PP_2797, which contained three of our transposon hits (PP_2791, PP_2793 and PP_2794). We analyzed the sequence homology of the seven genes in the operon using the basic local alignment search tool (BLAST[28]) and assigned predicted functions, listed in **Table 2**. We were unable to find any published studies about these genes beyond the automated sequence annotations. Therefore, we investigated the expression and function of these genes involved in LA catabolism. First, we isolated RNA from wild type *P. putida* grown in minimal media with LA as the carbon source and demonstrated that we could locate all seven genes by PCR amplification of cDNA created with a reverse primer specific to PP_2797 (**Figure 1 A,B,C**). The transcription start site (TSS) of the operon was isolated by 5' RACE[29] (**Figure 1D**) and implicated a different start codon for PP_2791, 72 bp downstream of the one originally reported[30,31]. A σ^{54} promoter sequence located upstream of PP_2791 was identified by comparing upstream of the new TSS with published σ^{54} promoter consensus sequences[32] (**Supplementary**

Figure 1). The data presented below suggests the proteins encoded by this operon are important in LA catabolism and we propose that the polycistronic genes be designated as *lvaABCDEFG*.

Upstream of *lvaABCDEFG*, we identified a gene oriented divergently from the operon (PP_2790) and predicted to encode a transcription factor with a σ^{54} interaction domain and homology to the propionate metabolism activator, *prpR*. The genomic organization strongly suggested that the gene encoded a regulator for the *lva* operon. Consequently, we deleted PP_2790 and evaluated growth of *P. putida* strains on both LA and a likely intermediate, 4HV. The ΔPP_2790 mutant was unable to grow on LA and 4HV suggesting that it acts as an activator for the operon. Expression of PP_2790 on a plasmid restored growth of the deletion strain on LA and 4HV. To identify compounds that activate *lvaABCDEFG* expression, we built a transcriptional reporter system that linked sfGFP to the σ^{54} promoter sequence located upstream of *lvaA*. The reporter cassette was cloned onto a broad host range vector (**Figure 2A**) and the resulting construct was transformed into wild type *P. putida*. We tested a variety of short and medium chain length acids by adding them to rich media and evaluating the corresponding sfGFP expression levels. We observed strong sfGFP fluorescence only when LA or 4HV were added to the system (**Figure 2B**). For these reasons, we suggest that PP_2790 encodes transcriptional regulator responsive to the LA pathway and should be designated *lvaR*.

Genetic and Biochemical Studies of *lvaABCDEFG* Operon

To confirm the involvement of the *lva* operon in LA catabolism, we created a deletion mutant of each *lva* gene predicted to encode an enzymatic protein and a corresponding complementation plasmid using the P_{araBAD} promoter. We tested the ability of the resulting strains to grow on LA and 4HV (**Table 2, Supplementary Figure 2**). In addition, we purified the five enzymes from

cultures of *E. coli* BL21 (DE3), reconstituted the enzymatic reactions *in vitro*, and used liquid chromatography/mass spectrometry (LC/MS) to identify reaction products. We used selective ion scanning to monitor the masses for likely intermediates based on prior studies[22,23,25] and the following hypothesized pathway (**Figure 3**, **Supplementary Figure 3**). First, LA is activated as a coenzyme A-thioester, levulinyl-CoA (LA-CoA). Second, LA-CoA is reduced to 4-hydroxyvaleryl-CoA (4HV-CoA). Third, 4HV-CoA is phosphorylated at the γ-position to yield 4-phosphovaleryl-CoA (4PV-CoA). Fourth, 4PV-CoA is dephosphorylated to yield a pentenoyl-CoA species (likely 3-pentenoyl-CoA). Last, pentenoyl-CoA is hydrated to yield 3-hydroxyvaleryl-CoA (3HV-CoA) which can be further oxidized via β-oxidation to yield acetyl-CoA and propionyl-CoA or incorporate 3HV-CoA into PHA polymers. The remainder of this manuscript will provide evidence supporting our hypothesized metabolic pathway for converting LA to 3-hydroxyvaleryl-CoA (3HV-CoA) and assignment of enzymes to each reaction.

Table 2. *P. putida* LA Operon Knockout and Complementation

		Growth on LA		Growth on 4HV	
Genotype	Predicted Function	EV	Complement	EV	Complement
WT		++	N/A	++	N/A
Δ*lvaR*	σ^{54} dependent sensory box protein	-	++	-	++
Δ*lvaA*	Phosphotransferase family	-	++	-	++
Δ*lvaB*	Hypothetical protein	-	++	-	++
Δ*lvaC*	acyl-CoA dehydrogenase family protein	-	++	+	++
Δ*lvaD*	short chain dehydrogenase/reductase family	-	++	++	++
Δ*lvaE*	Acyl-CoA synthetase	++	++	-	+

(EV) empty vector plasmid; (N/A) not applicable; (-) No growth; (+) Visible growth; (++) Robust growth

lvaE

The presence of an enzyme (encoded by *lvaE*) with homology to an acyl-CoA synthetase (including a putative CoA binding region and an AMP binding site) suggested that the degradation pathway acts on CoA thioesters and begins with the activation of acids to acyl-CoA's. The Δ*lvaE* strain grew on LA but not on 4HV, indicating that LA may also be activated by other CoA-synthetases in *P. putida*. We quantified the activity of purified LvaE (6x-His N-terminal fusion) on a variety of organic acid substrates using the EnzChek® Pyrophosphate Assay Kit which detects pyrophosphate released in the first half reaction to creating the acyl-AMP intermediate (**Figure 4A**). LvaE demonstrated activity on C_4-C_6 carboxylic acids, including LA and 4HV (**Figure 4B**), but showed minimal activity on other organic acids (lactate, pyruvate,

acetate, propionate, octanoate). Using LC/MS to detect reaction products (**Figure 5**), we demonstrated that LvaE was necessary and sufficient to catalyze the ligation of CoA to LA, generating levulinyl-CoA (LA-CoA). None of the other enzymes from the operon catalyzed this or any other reaction using LA as a substrate (**Supplementary Figure 4**), confirming that the pathway proceeds via acyl-CoA intermediates.

lvaD

The second step in our proposed pathway is the reduction of LA-CoA to 4HV-CoA which we predicted to be catalyzed by *lvaD*. *lvaD* is annotated as an oxidoreductase containing an NADH binding domain and was found to be required for growth on LA but not necessary for growth on 4HV (**Table 2**). We purified LvaD in a similar manner to LvaE but used an N-terminal maltose binding protein (MBP) tag to increase the solubility of the enzyme[33]. The *in vitro* reaction containing LvaD and LvaE verified that LvaD is involved in the production of 4HV-CoA (**Figure 5A**). Furthermore, LvaDE was the only enzyme combination capable of generating 4HV-CoA *in vitro* (**Figure S2**). LvaD can catalyze the reduction of LA-CoA with either NADH or NADPH (**Supplementary Figure 4**).

lvaAB

We hypothesized that the third intermediate would be 4-phospho-valeryl-CoA (4PV-CoA) based off its observation in LA degradation in rat livers[24,25]. The first gene in the operon, *lvaA*, has putative homology regions, including an ATP binding site, that associated it with the kinase superfamily and phosphotransferase family of enzymes. The second protein in the operon (LvaB) has no listed function and is predicted to be only 12 kDa in size. Orthologous sequence alignments of *lvaB* reveal that in all other organisms this hypothetical protein is located

immediately downstream of an *lvaA* ortholog. Therefore, a pull down experiment was used to determine if the two proteins interact[34,35]. LvaA was N-terminally tagged with MBP and cloned into a pET expression vector. LvaB was cloned directly downstream of LvaA as it is found in *P. putida's* native genome sequence. The recombinant proteins were expressed in *E. coli* BL21 (DE3) and purified using the MBP tag. An SDS-page gel of the eluent contained two bands at 85 kDa and 12 kDa, closely matching the predicted sizes of MBP-LvaA and untagged LvaB respectively (**Supplementary Figure 5**). We performed MALDI-TOF-TOF mass spectrometry on a trypsin digest of the 12 kDa band and identified the protein sequence to be LvaB (**Supplementary Figure 5)**.

Growth studies of deletion mutants revealed that *lvaA* and *lvaB* are both required for growth on either LA or 4-HV. This supports the hypothesis that they are involved in a reaction after the conversion of LA-CoA to 4HV-CoA. To confirm that the association between LvaA and LvaB is important for enzymatic activity, we tested the following enzymatic combinations: i) LvaA, LvaD and LvaE, ii) LvaB, LvaD, and LvaE, iii) LvaAB, LvaD and LvaE. We observed a decrease of 4HV-CoA and an increase of the predicted 4PV-CoA intermediate only when all four of the enzymes were present (**Figure 5A**, **Supplementary Figure 4**).

To verify the identity of 4PV-CoA, we performed tandem mass spectrometry (**Figure 6, Supplementary Figure 6, Supplementary Table 2**). We compared the MS/MS spectra of 4HV-CoA and 4PV-CoA and detected major ion fragments at *m/z* 786.191, 537.106 and 519.095 (4HV-CoA) and 866.158, 617.072 and 599.061 (4PV-CoA). For each compound, these fragments can be assigned to the cleavage of a P-O bond, an O-C bond and the dehydration of O-C cleaved product, respectively. Both compounds are fragmenting at the same bonds, but the resulting *m/z*

values for the daughter ions differ by 79.967. This mass corresponds to the m/z of PO_3H^-, supporting the existence of the phosphorylated 4HV-CoA species, 4PV-CoA.

lvaC

The final step in the hypothesized pathway is the formation of 3HV-CoA. Given that the combination of LvaABDE was responsible for generating 4PV-CoA and no 3HV-CoA was detected in these reactions, we postulated that LvaC was responsible for the final conversion steps. LvaC has homology to the dehydrogenase family of enzymes and 30% amino acid sequence identity to the *E. coli* acyl-CoA dehydrogenase protein. The Δ*lvaC* strain was unable to grow on LA, but grew weakly on 4HV. LvaC was purified as an MBP fusion and the resulting protein pellet displayed a yellow hue. This is often indicative of a co-purified flavoprotein and an absorbance scan of the protein revealed absorbance maxima that are consistent with a flavin co-factor (**Supplementary Figure 5**). When the LvaC sample was treated with trichloroacetic acid and centrifuged[36], a white protein pellet and a yellow hued supernatant were observed (data not shown). This indicates that the co-factor was not covalently bound to LvaC.

When LvaC was added to the *in vitro* reaction mixture, the concentrations of reaction intermediates (LA-CoA, 4HV-CoA, 4PV-CoA) were reduced while the abundance of 3HV-CoA and a pentenoyl-CoA species increased (**Figure 5A**). This species is likely either 2-pentenoyl-CoA and/or 3-pentenoyl-CoA, which could not be resolved with our methods. Both compounds eluted at the same retention time with the same molecular mass. To test if LvaC is solely responsible for the conversion of 4PV-CoA to 3HV-CoA, we ran a two-step reaction. First, we performed the LvaABDE reaction with LA, CoA, ATP, NAD(P)H and separated the CoA products from the enzymes. To the enzyme-free mixture, we added LvaC without additional co-

factors. After 30 min, we observed signals for both pentenoyl-CoA and 3HV-CoA. This indicated that the putative oxidoreductase, LvaC, is responsible for both the removal of the phosphate group to produce the enoyl-CoA and the hydration of the enoyl to the 3-hydroxyl compound. To reconstitute the whole pathway, we set-up a time course reaction with all five Lva enzymes and LA as the starting substrate. Over time, we observed a rapid increase in pentenoyl-CoA followed by a slow disappearance that mirrored the increase in the 3HV-CoA signal (**Figure 5B**). This suggests that the hydration reaction may be the limiting step in the overall pathway.

lvaFG

Based on homology alignments, *lvaG* is predicted to encode a protein with 95% amino acid sequence identity to a *Pseudomonas aeruginosa* cation acetate symporter and *lvaF* shares 33% amino acid sequence identity with the *E. coli* inner membrane protein Yhjb[28]. Sequence alignments of *lvaF* orthologs indicate that *lvaF* and *lvaG* are found with the same spatial relationship to each other in many organisms (data not shown). These proteins are likely involved in organic acid transport but are unlikely to be involved in the catabolism of LA given that they were not necessary for the enzymatic conversion of LA to 3HV-CoA *in vitro*.

Discussion

The work described herein identified an operon responsible for assimilating LA into the β-oxidation pathway of *P. putida.* Through an integrated genetic and *in vitro* biochemistry study, we demonstrated that the genes *lvaABCDE* are upregulated in the presence of LA and are sufficient for the conversion of LA to 3HV-CoA, an intermediate of native β-oxidation. Removing any enzyme from the reaction mixture abolished 3HV-CoA production, indicating all 5 genes are necessary for this pathway. The biochemical assays confirmed the presence of 4PV-

CoA, an intermediate previously observed in the metabolism of LA in rat livers. In sum, the pathway consumes at least 2 ATP and one reducing equivalent to produce 3HV-CoA (**Figure 3**). β-oxidation of 3HV-CoA to acetyl-CoA and propionyl-CoA would recover the reducing equivalent. Given the energy demands of the pathway, growth on LA must be performed aerobically or in the presence of an alternative electron acceptor to enable ATP synthesis via respiration.

Like many catabolic pathways, expression of the *lva* operon is regulated by the presence of the pathway substrates. Using a transcriptional reporter assay, we demonstrated that the *lva* operon is upregulated by a transcriptional activator encoded by the divergent *lvaR* gene. Additionally, we suspect that the *lva* operon is also regulated by Crc, a global carbon catabolite repressor. Crc is an mRNA binding protein that prevents protein translation when bound to a specific mRNA sequence in *P. putida*, AAnAAnAA[37–40]. This sequence pattern is found immediately upstream of *lvaE* (**Supplementary Figure 1**), which encodes an acyl-CoA synthetase that initiates the pathway. The presence of the Crc target sequence suggests that the operon is also subject to *P. putida's* carbon catabolite repression system which may explain the diauxic growth curves observed for mixtures of glucose and LA.

While LvaB was shown to be essential for LA catabolism, its exact role remains unclear. LvaB is a small protein (~100 amino acids) that is unlikely to contain enzymatic activity by-itself. Furthermore, LvaB co-purifies with LvaA, is essential for the phosphorylation of 4HV-CoA, and its orthologs are consistently found adjacent to orthologs of *lvaA* in the genomes of other organisms. Similar examples where small proteins provide critical support to a variety of biological functions, including metabolism and enzymatic function, have been indentified[41,42]. For example, nonribosomal peptide synthetase gene clusters often contain a small protein that

belongs to the MbtH-like protein family, a family of proteins that are known to bind adenylation domains and enable catalytic activity. MbtH-like proteins form the necessary complexes required for domain activation but are not predicted to interact directly with the catalytic site[43–45]. Although *lvaB* does not share significant sequence homology with known MbtH-like proteins, we speculate that it could be playing a similar role with LvaA, where the presence of LvaB is required to form an active LvaAB complex. Without a crystal structure, the specific interaction between LvaA and LvaB and its role in catalysis will be difficult to unlock.

Interestingly, the isomerization of 4HV-CoA to 3HV-CoA in *P. putida* proceeds through a phosphorylated intermediate, 4PV-CoA, a compound also observed in a study of LA metabolism in rat livers[25]. This study suggested the 3HV-CoA was generated via a pathway comprised of complex phosphorylated intermediates. We did not detect MS peaks corresponding to any of these compounds in our *in vitro* reaction mixtures. Instead, based on changes we observed in total ion abundance over time, we propose that 4PV-CoA is dephosphorylated to an enoyl-CoA and subsequently rehydrated to 3HV-CoA. We suspect that the phosphorylation of 4HV-CoA by LvaAB generates a better leaving group and makes the subsequent dehydration more thermodynamically favorable. However, the mechanism for these last steps remains unclear.

Previous groups studying the nonmevalonate pathway have identified phosphate elimination steps for the formation of a double bond that is reminiscent of the intermediates we observed[46,47], but these reactions do not include a rehydration step. The time course measurements that we collected for the full reaction indicate that the formation of the pentenoyl-CoA happens fairly quickly, but the transition from the pentenoyl-CoA to the 3HV-CoA is a much slower reaction (**Figure 5B**). Our tests indicate that LvaC is capable of converting 4PV-CoA to 3HV-CoA, but

those reactions still contain a higher abundance of pentenoyl-CoA compared to 3HV-CoA. A more detailed mechanistic study of the final steps may clarify the specific role of *lvaC*.

Understanding how LA metabolism works is important because LA is a common byproduct of biomass hydrolysis and is often present in the final feedstock. High concentrations of LA in the feedstock can lead to microbial inhibition and represents an underused source of carbon in traditional sugar fermentations. By discovering the catabolic pathway, microbes can be engineered to detoxify the media and/or utilize LA as a source of carbon, maximizing the overall carbon conversion from biomass into high value products. Additionally, identifying the structure of LA metabolism will improve metabolic models and enable pathway design for novel LA-based bioconversions.

Acknowledgements

Work in the Pfleger lab was funded by the National Science Foundation (CBET-114678) and the William F. Vilas Trust. Work in the Deutschbauer lab was funded by ENIGMA, a Scientific Focus Area Program, supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research and Genomics:GTLFoundational Science through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.. Work in the Amador-Noguez lab was funded by the HHMI International Student Research Fellowship. R.L.C. was supported by the NIH NHGRI Genomic Sciences Training Program (T32 HG002760). A.L.M. was supported by NSF SEES fellowship (GEO-3171215871).

We want to thank J. Escalante for providing the plasmid pK18*mobsacB* and J. Altenbuchner for providing the strain *P. putida* KTU and the plasmid pJOE6261.2. We would also like to thank the Mass Spectrometry/ Proteomics Facility at the UW-Madison Biotechnology Center for performing the in-gel digest and providing the MALDI-TOF-TOF results.

Author Contribution

J.M.R. D.E.A. and B.F.P. conceived the study. J.M.R. designed and performed the experiments and analyzed the data with the following exceptions. T.P. and D.A.M. designed the LC/MS/MS experiments and T.P. performed the LC/MS and LC/MS/MS experiments. D.E.A. and J.M.T. performed the transposon library screen. C.E.C. assisted with the promoter and CoA ligase assay.

A.L.M. proposed, and helped design and perform, the pull-down experiment. Y.S. and J.R. prepared the RB-TnSeq mutant library of P. putida KT2440 (Putida_ML5). K.M.W., R.L.C, and A.M.D. performed the fitness assays with the Putida_ML5 library. M.N.P. performed the data analysis to determine fitness values. R.L.C. prepared the supplementary analysis of the Putida_ML5 fitness experiments. A.M.D. and A.P.A. managed the Bar-Seq experiments. J.M.R. and B.F.P wrote the manuscript.

Methods

Please see the **Supplementary Methods** for additional details.

Chemicals, Strains, and Media

All chemicals were obtained from Sigma-Aldrich or Fisher Scientific. 4-hydroxyvalerate was made through the saponification of γ-valerolactone (GVL)[23]. Bacterial strains and plasmids used in this study are summarized in **Table 3**. Plasmid sequences are listed in the **Supplementary Material**. *E. coli* strains were grown at 37°C and P. putida strains were grown at 30°C unless otherwise noted. Kanamycin was used at final concentration of 50 μg/ml. 5-Fluorouracil was used at a final concentration of 20 μg/mL.

Table 3. Strains and Plasmid List

Strain/Plasmid	Relevant genotype/property	Source or Reference
Strains		
Pseudomonas putida		
KT2440	Wild Type	ATCC 47054
KTU	Δ*upp*	Altenbuchner et al[51]
Δ*lvaR*	Δ*upp* ΔPP_2970	This work
Δ*lvaA*	Δ*upp* ΔPP_2971	This work
Δ*lvaB*	Δ*upp* ΔPP_2972	This work
Δ*lvaC*	Δ*upp* ΔPP_2973	This work
Δ*lvaD*	ΔPP_2974	This work
Δ*lvaE*	Δ*upp* ΔPP_2975	This work
Escherichia coli		
CC118λpir	Δ*(ara-leu), araD,* Δ*lacX174, galE, galK, phoA, thi1, rpsE, rpoB, argE (Am), recA1,* lysogenic λ*pir*	de Lorenzo et al[26]
DH5α	F^- Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*) U169 *recA1 endA1 hsdR17* (r_k^-,m_k^+) *phoA supE44 thi-1 gyrA96 relA1* λ^-	Invitrogen
Plasmids		
pBAM1	*tnpA*, Amp^R, Kan^R, *ori*R6K	de Lorenzo et al[26]
pJOE6261.2	*upp* (from *P. putida*), Kan^R, ColE1 origin	Altenbuchner et al[51]
pJOE-lvaR	pJOE6261.2 with up- and downstream regions of *lvaR*	This work
pJOE-lvaA	pJOE6261.2 with up- and downstream regions of *lvaA*	This work
pJOE-lvaB	pJOE6261.2 with up- and downstream regions of *lvaB*	This work
pJOE-lvaC	pJOE6261.2 with up- and downstream regions of *lvaC*	This work
pJOE-lvaE	pJOE6261.2 with up- and downstream regions of *lvaE*	This work
pBAD35	P_{BAD} promoter, Kan^R, pBBR1 origin	Lennen et al[52]
pBAD-*lvaA*	pBAD35 carrying *lvaA*	This work
pBAD-*lvaB*	pBAD35 carrying *lvaB*	This work
pBAD-*lvaC*	pBAD35 carrying *lvaC*	This work
pBAD-*lvaD*	pBAD35 carrying *lvaD*	This work
pBAD-*lvaE*	pBAD35 carrying *lvaE*	This work
pK18*mobsacB*	*sacB*, Kan^R, pMB1 origin	Schafer et al[53]
pK18-*lvaD*	pK18*mobsacB* containing up- and downstream regions of *lvaD*	This work
pJMR74	pBAD35 with P_{BAD} promoter and *araC* replaced with *lvaA* promoter and *lvaR* (*P. putida*) carrying sfGFP	This work

Transposon Library and Screening

The *P. putida* transposon library was created following a protocol adapted from Martinez-Garcia et al[26]. Suicide vector delivery was achieved through bi-parental mating. The transposon library was screened by replica plating colonies from the M9 citrate plates onto LB, M9 glucose and M9 LA plates supplemented with kanamycin. Positive hits were identified as colonies that exhibited growth on LB and glucose plates but not on LA plates.

RNA Experiments

RNA was extracted from P. putida using a protocol adapted from Pinto *et al*[54]. The transcription start site for genes *lvaR* and *lvaA* were isolated using an adapted 5' Race protocol from Schramm *et al*[29]. Reverse transcription PCR was performed using primers listed in **Supplementary Table 3**.

P. putida Knockouts

The genetic knockout of *lvaD* was performed following the protocol from Schafer et al[53]. Knockouts of the remaining genes in *P. putida* were performed following the protocol from Graf et al[51]. Knockout cassettes were designed with 500 bp of homology up and down stream of the deletion site flanking kanR and 5-FU markers. Cassettes were introduced to *P. putida* via a suicide vector. Disruption mutants were selected on kanamycin and clean deletions were selected on 5-FU. Colonies were then screened by colony PCR to confirm deletion strains.

Transcriptional Reporter Assay

P. putida KT2440 was transformed with a broad host range plasmid (pJMR74) containing a kan^R marker, *lvaR, and* sfGFP cloned under the native promoter for *lvaA*. *P. putida* KT2440 harboring an empty vector (no sfGFP) was used as a negative control. The fluorescence of cultures exposed to 20 mM of the appropriate carboxylic acid (acetate, propionate, butyrate, valerate, LA, 4HV, or hexanoate) was measured in a Tecan inifinite m1000 (ex 485 nm/em 510 nm).

CoA Ligase Assay

A CoA ligase activity assay was performed with the EnzChek® Pyrophosphate Assay Kit. The final reaction volume of 100 µL contained 0.1 mM ATP, 0.1 mM CoA, 0.2 µM LvaE, 0.2 mM MESG, 1 U purine nucleoside phosphorylase, 0.01 U pyrophosphatase, 50 mM Tris-HCL, 1 mM $MgCl_2$ and 0.1 mM substrate (sodium acetate, sodium propionate, butyric acid, valeric acid, levulinic acid, hexanoic acid, octanoic acid, 4HV, 2-pentenoic acid, 3-pentenoic acid, 3HV, γ-valerolactone, pyruvate, lactic acid, L-carnitine). All substrate stocks were adjusted to pH 7 before use. Reactions were incubated at 25°C for 30 minutes before measuring the absorbance at 360 nm in a Tecan infinite m1000.

Protein Production and Purification

Expression vectors for each *lva* gene were constructed using the pET28b backbone and expressed in *E. coli* BL21 (DE3). After induction with 1 mM IPTG at OD600 0.8, cultures were chilled on ice for 10 minutes and incubated at 16°C for 18 hours. Proteins were purified using a a GE Äkta Start System with a 1 mL HisTrap HP column (LvaE) or 1 mL MBPTrap HP column (LvaABCD). Eluted proteins were buffer exchanged using a GE PD-10 column and concentrated with an Amicon® Ultra 4 mL Centrifugal Filter with a 10 kDa cut-off size. Proteins were stored at -80°C until use.

Enzyme Assays and Metabolite Purification

All *in vitro* enzyme assays were performed in a 30° water bath at a pH of 6.5 and contained 50 mM Tris-HCL, 1 mM $MgCl_2$, and 2 mM DTT. Final reaction concentrations included the following components, depending on enzymes added: 0.5 mM LA, 0.55 mM CoA, 0.55 mM ATP (1.05 mM ATP when *lvaAB* were present), 0 mM NAD(P)H (0.55 mM NAD(P)H when *lvaD* was present). Final protein concentrations were: LvaA (0.2 µM), LvaB (0.8 µM), LvaAB (0.4 µM), LvaC (0.4 µM), LvaD (0.2 µM), and LvaE (0.2 µM) (**Supplementary Figure 5**). The *in vitro* enzyme assays were incubated for 30 minutes, excluding the timecourse which was incubated for various intervals up to 60 minutes. Reaction metabolites were purified following a modified protocol from Zhang *et al*[24] and stored at -80°C until LC/MS analysis.

Liquid Chromatography Mass spectrometry (LC/MS, LC/MS/MS)

Samples were analyzed using an HPLC-MS/MS system consisting of a VanquishTM UHPLC system (Thermo Scientific) coupled by electrospray ionization (ESI; negative polarity) to a hybrid quadrupole - high-resolution mass spectrometer (Q Exactive orbitrap, Thermo Scientific) operated in full scan mode for detection of targeted compounds based on their accurate masses. Fragmentation of CoA, 4HV-CoA, and phosphorylated 4HV-CoA was achieved using parameters indicated in the Supplementary Methods.

Figures

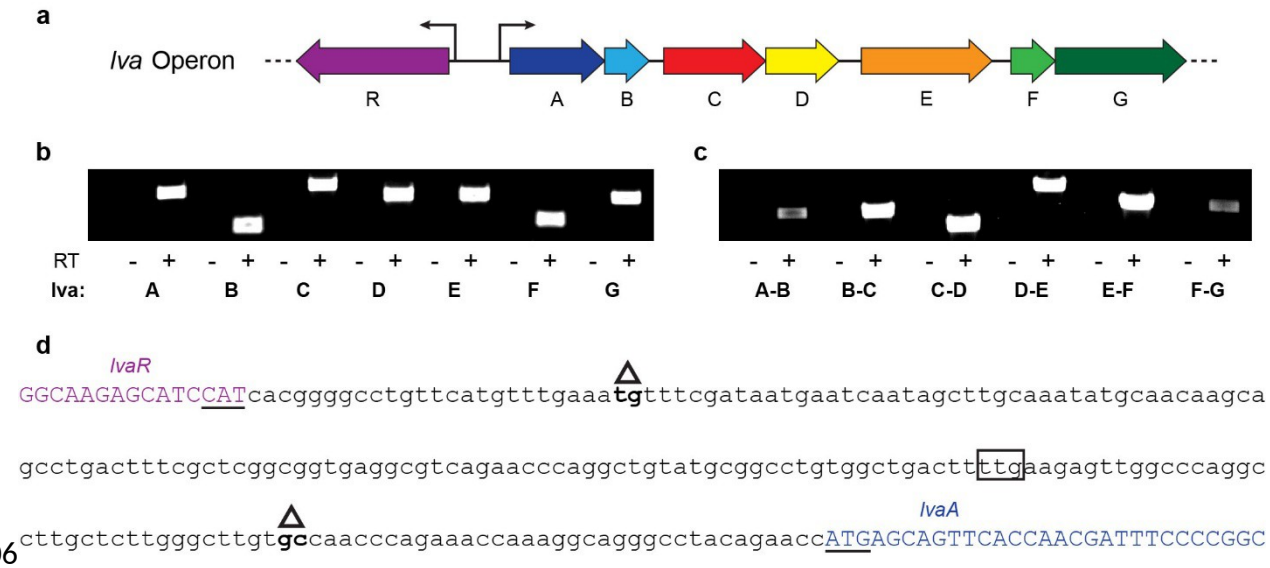


Figure 1. *P. putida lva* operon genetic characterization A. Organization of the *lvaRABCDEFG* (9,323 bp) operon. B. Reverse Transcriptase (RT) PCR demonstrates that each gene is expressed in cells grown on LA. Samples were compared with the negative control (-RT) where reverse transcriptase was omitted from the reaction. C. RT-PCR of cDNA created with primer JMR237 demonstrates that operon is polycistronic. Note that a product spanning each intergenic region was observed. D. The transcription start sites (TSS) of regulator *lvaR* and *lvaA* were identified by 5'-RACE. Underlined sequence indicates ATG start codon. Triangle highlights experimentally determined TSS. Boxed sequence indicates previously annotated translation start site for *lvaA*.

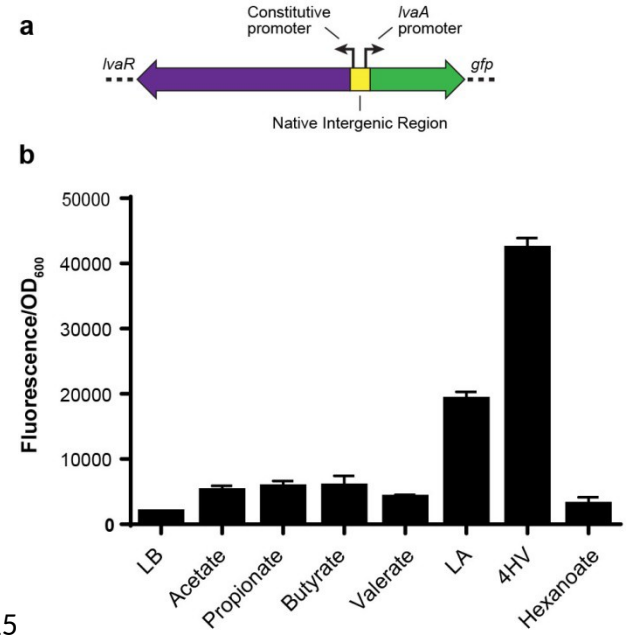


Figure 2. *lva* operon induction assay. A. Schematic of transcriptional GFP fusion used to test induction of the *lva* operon. *lvaR* was cloned onto a plasmid containing its native constitutive promoter and the native promoter region for *lvaA*. The fluorescent protein sfGFP was cloned in place of *lvaA*. B. GFP fluorescence was measured from LB-cultures supplemented with various organic acids (20 mM).

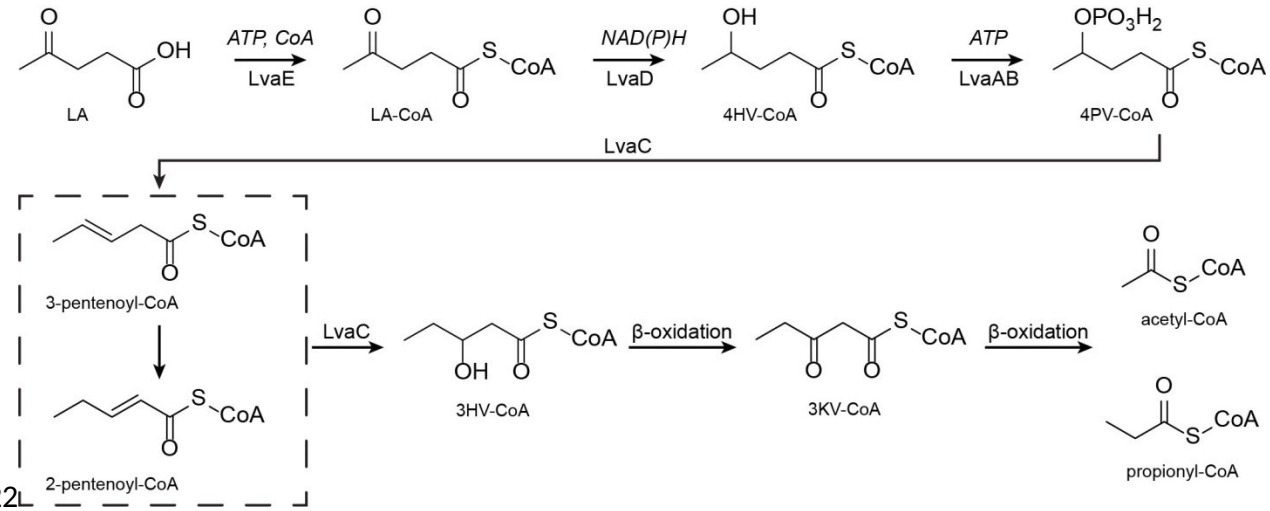


Figure 3. Proposed pathway for LA metabolism. LA, levulinic acid; 4HV, 4-hydroxyvalerate; 3HV, 3-hydroxyvalerate; LA-CoA, levulinyl-CoA; 4HV-CoA, 4-hydroxyvaleryl-CoA; CoA, coenzyme-A; ATP, adenosine triphosphate; 4PV-CoA, 4-phosphovaleryl-CoA; 3KV-CoA, 3-ketovaleryl-CoA; NAD(P)H, Nicotinamide adenine dinucleotide (phosphate) reduced.

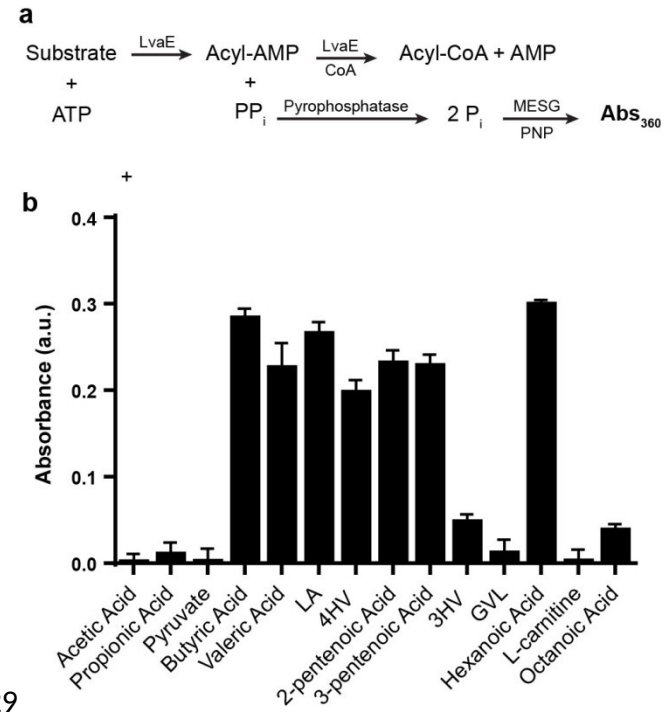
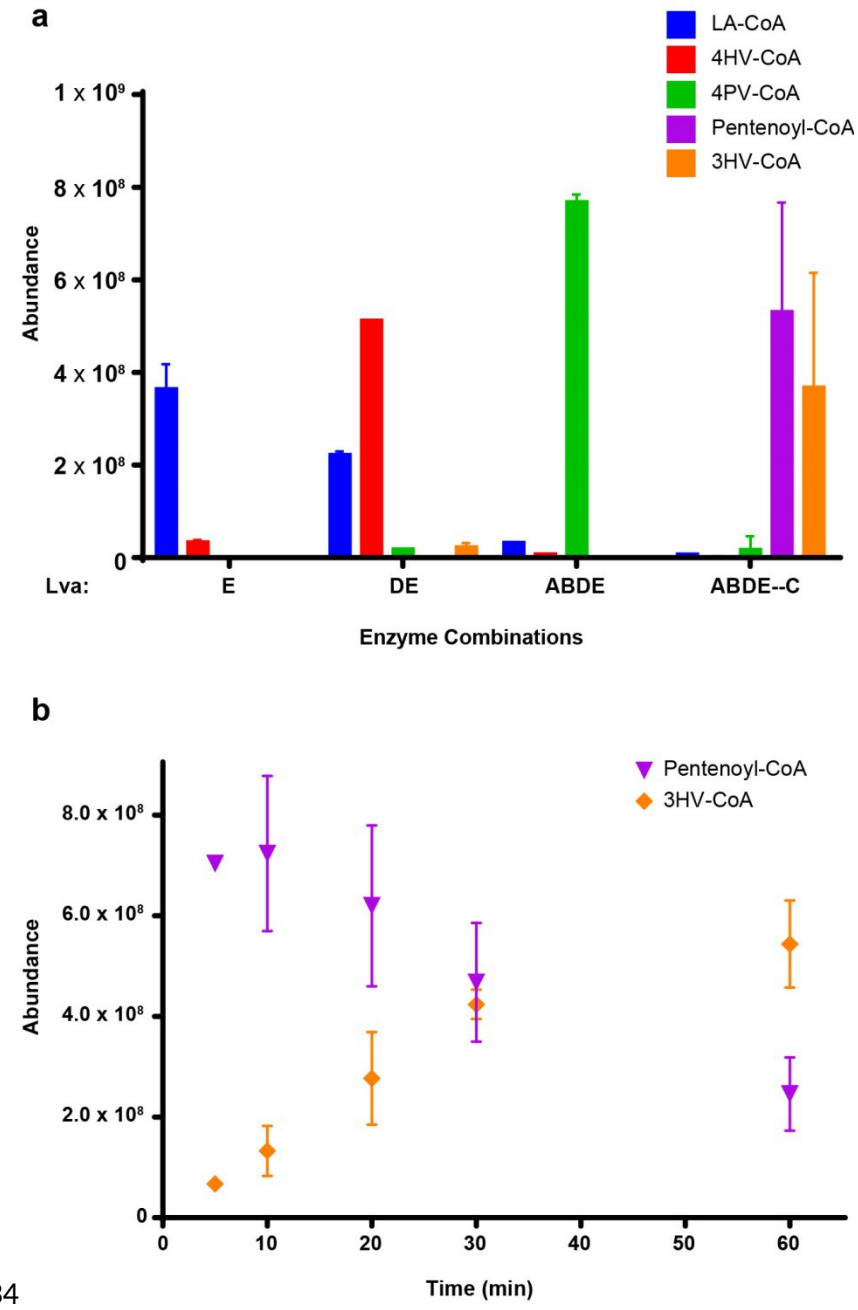


Figure 4. LvaE CoA-Ligase Activity A. Schematic of CoA-ligase activity assay. Using the Enzchek® Pyrophosphatase Assay kit, the amount of pyrophosphate released during the CoA ligase reaction was measured as an increase of absorbance at 360 nm. B. Activity of LvaE towards a variety of short and medium chain acids.



434

435**Figure 5**. CoA species abundance in LC/MS analysis of *in vitro* enzyme combinations. A. Abundance of CoA
436species created after 30 min of incubating LA, ATP, NAD(P)H with varying enzyme combinations. ABDE—C
437indicates that the LvaABDE reaction was performed first, metabolites were separated from LvaABDE, and the
438resulting solutions was supplemented with LvaC solely. The reaction confirms that LvaC is capable of converting
4394PV-CoA to 3HV-CoA. B. Abundance of CoA species over a 60 minute timecourse for a mixture of LvaABCDE,
440LA, ATP, and NAD(P)H.

441

442

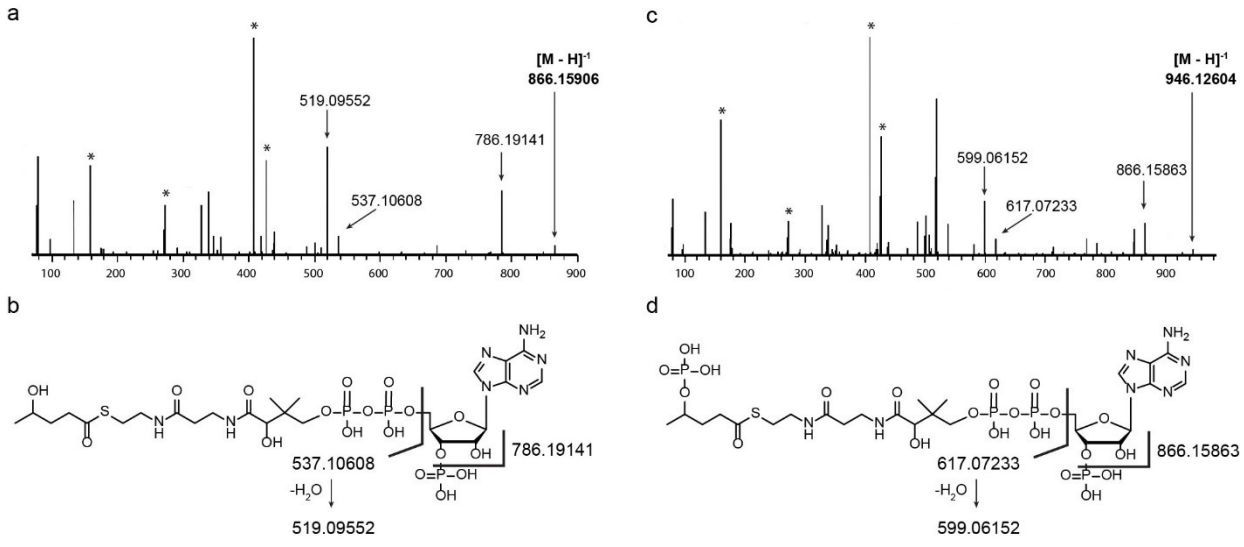


Figure 6. Comparison of 4HV-CoA and 4PV-CoA MS/MS spectra. A. MS/MS spectra for 4HV-CoA. B. Assignment of selected fragments from 4HV-CoA. C. MS/MS spectra for 4PV-CoA. D. Assignments of selected fragments from 4PV-CoA. The masses between the selected fragments of 4PV-CoA and 4HV-CoA differ by the mass of PO_3H^- (79.967), indicating 4PV-CoA contains a phosphate group not found in 4HV-COA. **Bold** values indicate the mass of the parent ion. Peaks identified with the symbol (*) are fragments resulting from coenzyme A. See **Supplementary Figure 6** and **Supplementary Table 2** for additional fragmentation information.

References

1. Pileidis, F. D. & Titirici, M. M. Levulinic Acid Biorefineries: New Challenges for Efficient Utilization of Biomass. *ChemSusChem* **9,** 562–582 (2016).

2. Bozell, J. J. *et al.* Production of levulinic acid and use as a platform chemical for derived products. *Resour. Conserv. Recycl.* **28,** 227–239 (2000).

3. Werpy, T. & Petersen, G. Top Value Added Chemicals from Biomass. *Program* 76 (2004). doi:10.2172/926125

4. Al-Shaal, M. G., Dzierbinski, A. & Palkovits, R. Solvent-free [gamma]-valerolactone hydrogenation to 2-methyltetrahydrofuran catalysed by Ru/C: a reaction network analysis. *Green Chem.* **16,** 1358–1364 (2014).

5. Geilen, F. M. A. *et al.* Selective and flexible transformation of biomass-derived platform chemicals by a multifunctional catalytic system. *Angew. Chemie - Int. Ed.* **49,** 5510–5514 (2010).

6. Lange, J. P. *et al.* Valeric biofuels: A platform of cellulosic transportation fuels. *Angew. Chemie - Int. Ed.* **49,** 4479–4483 (2010).

7. Chan-Thaw, C. E., Marelli, M., Psaro, R., Ravasio, N. & Zaccheria, F. New generation biofuels: γ-valerolactone into valeric esters in one pot. *RSC Adv.* **3,** 1302–1306 (2013).

8. Alonso, D. M., Wettstein, S. G. & Dumesic, J. A. Gamma-valerolactone, a sustainable platform molecule derived from lignocellulosic biomass. *Green Chem.* **15,** 584 (2013).

9. De, S., Saha, B. & Luque, R. Hydrodeoxygenation processes: Advances on catalytic transformations of biomass-derived platform chemicals into hydrocarbon fuels. *Bioresour. Technol.* **178,** 108–118 (2015).

10. Joshi, H., Moser, B. R., Toler, J., Smith, W. F. & Walker, T. Ethyl levulinate: A potential bio-based diluent for biodiesel which improves cold flow properties. *Biomass and Bioenergy* **35,** 3262–3266 (2011).

11. Zhang, Z., Dong, K. & Zhao, Z. Efficient conversion of furfuryl alcohol into alkyl levulinates catalyzed by an organic-inorganic hybrid solid acid catalyst. *ChemSusChem* **4,** 112–118 (2011).

12. Demolis, A., Essayem, N. & Rataboul, F. Synthesis and Applications of Alkyl Levulinates. *ACS Sustain. Chem. Eng.* **2,** 1338–1352 (2014).

13. Moore, J. A. & Tannahill, T. Homo- and co-polycarbonates and blends derived from diphenolic acid. *High Perform. Polym.* **13,** 305–316 (2001).

14. Guo, Y., Li, K., Yu, X. & Clark, J. H. Mesoporous H3PW12O40-silica composite: Efficient and reusable solid acid catalyst for the synthesis of diphenolic acid from levulinic acid. *Appl. Catal. B Environ.* **81,** 182–191 (2008).

48615. Chung, S. H., Choi, G. G., Kim, H. W. & Rhee, Y. H. Effect of Levulinic Acid on the
487 Production of Poly (3-hydroxybutyrate-co-3-hydroxyvalerate) by Ralstonia eutropha
488 KHB-8862. *Society* **39,** 79–82 (2001).

48916. Berezina, N. & Yada, B. Improvement of the poly(3-hydroxybutyrate-co-3-
490 hydroxyvalerate) (PHBV) production by dual feeding with levulinic acid and sodium
491 propionate in Cupriavidus necator. *N. Biotechnol.* **33,** 231–236 (2016).

49217. Valentin, H. E., Schönebaum, A. & Steinbüchel, A. Identification of 5-hydroxyhexanoic
493 acid, 4-hydroxyheptanoic acid and 4-hydroxyoctanoic acid as new constituents of
494 bacterial polyhydroxyalkanoic acids. *Appl. Microbiol. Biotechnol.* **46,** 261–267 (1996).

49518. Jang, J. H. & Rogers, P. L. Effect of levulinic acid on cell growth and poly-beta-
496 hydroxyalkanoate production by Alcaligenes sp SH-69. *J. Chem. Inf. Model.* **18,** 219–224
497 (1996).

49819. Habe, H. *et al.* Bacterial production of short-chain organic acids and trehalose from
499 levulinic acid: A potential cellulose-derived building block as a feedstock for microbial
500 production. *Bioresour. Technol.* **177,** 381–386 (2015).

50120. Martin, C. H., Wu, D., Prather, K. L. J. & Jones Prather, K. L. Integrated bioprocessing for
502 the pH-dependent production of 4-valerolactone from levulinate in pseudomonas putida
503 KT2440. *Appl. Environ. Microbiol.* **76,** 417–424 (2010).

50421. Yeon, Y. J., Park, H. Y. & Yoo, Y. J. Enzymatic reduction of levulinic acid by engineering
505 the substrate specificity of 3-hydroxybutyrate dehydrogenase. *Bioresour. Technol.* **134,**
506 377–380 (2013).

50722. Jaremko, M. & Yu, J. The initial metabolic conversion of levulinic acid in Cupriavidus
508 necator. *J. Biotechnol.* **155,** 293–298 (2011).

50923. Martin, C. H. & Prather, K. L. J. High-titer production of monomeric hydroxyvalerates
510 from levulinic acid in Pseudomonas putida. *J. Biotechnol.* **139,** 61–67 (2009).

51124. Zhang, G. F. *et al.* Catabolism of 4-hydroxyacids and 4-hydroxynonenal via 4-hydroxy-4-
512 phosphoacyl-CoAs. *J. Biol. Chem.* **284,** 33521–33534 (2009).

51325. Harris, S. R. *et al.* Metabolism of levulinate in perfused rat livers and live rats: Conversion
514 to the drug of abuse 4-hydroxypentanoate. *J. Biol. Chem.* **286,** 5895–5904 (2011).

51526. Martínez-García, E., Calles, B., Arévalo-Rodríguez, M. & de Lorenzo, V. pBAM1: an all-
516 synthetic genetic tool for analysis and construction of complex bacterial phenotypes. *BMC*
517 *Microbiol.* **11,** 38 (2011).

51827. Wetmore, K. M. M. *et al.* Rapid Quantification of Mutant Fitness in Diverse Bacteria by
519 Sequencing Randomly Bar-Coded Transposons. *MBio* **6,** 1–15 (2015).

52028. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment
521 search tool. *J. Mol. Biol.* **215,** 403–10 (1990).

52229. Schramm, G., Bruchhaus, I. & Roeder, T. A simple and reliable 5'-RACE approach.
523 *Nucleic Acids Res* **28,** E96 (2000).

52430. Espah Borujeni, A., Channarasappa, A. S. S. & Salis, H. M. M. Translation rate is
525 controlled by coupled trade-offs between site accessibility, selective RNA unfolding and
526 sliding at upstream standby sites. *Nucleic Acids Res.* **42,** 2646–2659 (2014).

52731. Salis, H. M. M., Mirsky, E. A. A. & Voigt, C. A. A. Automated design of synthetic
528 ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27,** 946–50 (2009).

52932. Barrios, H., Valderrama, B. & Morett, E. Compilation and analysis of sigma(54)-
530 dependent promoter sequences. *Nucleic Acids Res.* **27,** 4305–4313 (1999).

53133. Fox, J. D., Routzahn, K. M., Bucher, M. H. & Waugh, D. S. Maltodextrin-binding proteins
532 from diverse bacteria and archaea are potent solubility enhancers. *FEBS Lett.* **537,** 53–57
533 (2003).

53434. Striebel, F. *et al.* Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to
535 substrates by distinct but homologous enzymes. *Nat. Struct. Mol. Biol.* **16,** 647–651
536 (2009).

53735. Yamamoto, S. & Kutsukake, K. FliT acts as an anti-FlhD2C2 factor in the transcriptional
538 control of the flagellar regulon in Salmonella enterica serovar typhimurium. *J. Bacteriol.*
539 **188,** 6703–8 (2006).

54036. Dijkman, W. P. & Fraaije, M. W. Discovery and characterization of a 5-
541 hydroxymethylfurfural oxidase from Methylovorus sp. strain MP688. *Appl. Environ.*
542 *Microbiol.* **80,** 1082–1090 (2014).

54337. Moreno, R., Marzi, S., Romby, P. & Rojo, F. The Crc global regulator binds to an unpaired
544 A-rich motif at the Pseudomonas putida alkS mRNA coding sequence and inhibits
545 translation initiation. *Nucleic Acids Res.* **37,** 7678–7690 (2009).

54638. Moreno, R., Martínez-Gomariz, M., Yuste, L., Gil, C. & Rojo, F. The Pseudomonas putida
547 Crc global regulator controls the hierarchical assimilation of amino acids in a complete
548 medium: Evidence from proteomic and genomic analyses. *Proteomics* **9,** 2910–2928
549 (2009).

55039. Sonnleitner, E. *et al.* Novel Targets of the CbrAB/Crc Carbon Catabolite Control System
551 Revealed by Transcript Abundance in Pseudomonas aeruginosa. *PLoS One* **7,** (2012).

55240. Sonnleitner, E., Abdou, L. & Haas, D. Small RNA as global regulator of carbon catabolite
553 repression in Pseudomonas aeruginosa. (2009).

55441. Storz, G., Wolf, Y. I. & Ramamurthi, K. S. Small proteins can no longer be ignored. *Annu.*
555 *Rev. Biochem.* **83,** 753–77 (2014).

55642. Su, M., Ling, Y., Yu, J., Wu, J. & Xiao, J. Small proteins: Untapped area of potential
557 biological importance. *Front. Genet.* **4,** 1–9 (2013).

55843. Felnagle, E. A. *et al.* MbtH-like proteins as integral components of bacterial nonribosomal
559 peptide synthetases. *Biochemistry* **49,** 8815–8817 (2010).

56044. Herbst, D. A., Boll, B., Zocher, G., Stehle, T. & Heide, L. Structural basis of the
561 interaction of mbth-like proteins, putative regulators of nonribosomal peptide
562 biosynthesis, with adenylating enzymes. *J. Biol. Chem.* **288,** 1991–2003 (2013).

56345. Baltz, R. H. Function of MbtH homologs in nonribosomal peptide biosynthesis and
564 applications in secondary metabolite discovery. *J. Ind. Microbiol. Biotechnol.* **38,** 1747–
565 1760 (2011).

56646. Gräwert, T. *et al.* Structure of active IspH enzyme from escherichia coli provides
567 mechanistic insights into substrate reduction. *Angew. Chemie - Int. Ed.* **48,** 5756–5759
568 (2009).

56947. Hecht, S. *et al.* Studies on the nonmevalonate pathway to terpenes: the role of the GcpE
570 (IspG) protein. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 14837–14842 (2001).

57148. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred
572 kilobases. *Nat. Methods* **6,** 343–5 (2009).

57349. Sambrook, J. & Russell, D. W. (David W. *Molecular cloning : a laboratory manual*. (Cold
574 Spring Harbor Laboratory Press, 2001).

57550. Neidhardt, F. C., Bloch, P. L. & Smith, D. F. Culture medium for enterobacteria. *J.*
576 *Bacteriol.* **119,** 736–747 (1974).

57751. Graf, N. & Altenbuchner, J. Development of a method for markerless gene deletion in
578 Pseudomonas putida. *Appl. Environ. Microbiol.* **77,** 5549–52 (2011).

57952. Lennen, R. M., Braden, D. J., West, R. M., Dumesic, J. A. & Pfleger, B. F. A process for
580 microbial hydrocarbon synthesis: Overproduction of fatty acids in Escherichia coli and
581 catalytic conversion to alkanes. *Biotechnol. Bioeng.* **106,** 193–202 (2010).

58253. Schäfer, A. *et al.* Small mobilizable multi-purpose cloning vectors derived from the
583 Escherichia coli plasmids pK18 and pK19: selection of defined deletions in the
584 chromosome of Corynebacterium glutamicum. *Gene* **145,** 69–73 (1994).

58554. Pinto, F. L., Thapper, A., Sontheim, W. & Lindblad, P. Analysis of current and alternative
586 phenol based RNA extraction methodologies for cyanobacteria. *BMC Mol. Biol.* **10,** 79
587 (2009).

588