



## Initiation of fatty acid biosynthesis in *Pseudomonas putida* KT2440

Kevin J. McNaught<sup>a,b,1</sup>, Eugene Kuatsjah<sup>a,c,1</sup>, Michael Zahn<sup>d</sup>, Érica T. Prates<sup>c,e</sup>, Huiling Shao<sup>f</sup>, Gayle J. Bentley<sup>a,b</sup>, Andrew R. Pickford<sup>d</sup>, Josephine N. Gruber<sup>a</sup>, Kelley V. Hestmark<sup>a,b</sup>, Daniel A. Jacobson<sup>c,e</sup>, Brenton C. Poirier<sup>b,h</sup>, Chen Ling<sup>a,b</sup>, Myrsini San Marchi<sup>f</sup>, William E. Michener<sup>a</sup>, Carrie D. Nicora<sup>h</sup>, Jacob N. Sanders<sup>f</sup>, Caralyn J. Szostkiewicz<sup>a,b</sup>, Dušan Veličković<sup>g</sup>, Mowei Zhou<sup>g</sup>, Nathalie Munoz<sup>b,g</sup>, Young-Mo Kim<sup>b,h</sup>, Jon K. Magnuson<sup>b,h</sup>, Kristin E. Burnum-Johnson<sup>b,g</sup>, K.N. Houk<sup>f</sup>, John E. McGeehan<sup>d</sup>, Christopher W. Johnson<sup>a,b</sup>, Gregg T. Beckham<sup>a,b,c,\*</sup>

<sup>a</sup> Renewable Resources and Enabling Sciences Center, National Renewable Energy Laboratory, Golden, CO, 80401, USA

<sup>b</sup> DOE Agile BioFoundry, Emeryville, CA, 94608, USA

<sup>c</sup> Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, 37830, USA

<sup>d</sup> Centre for Enzyme Innovation, School of Biological Sciences, Institute of Biological and Biomedical Sciences, University of Portsmouth, Portsmouth, PO1 2DY, UK

<sup>e</sup> Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, 37830, USA

<sup>f</sup> Department of Chemistry and Biochemistry, University of California Los Angeles, CA, 90095, USA

<sup>g</sup> Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, 99354, USA

<sup>h</sup> Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, 99352, USA

### ARTICLE INFO

#### Keywords:

Fatty acid biosynthesis  
Decarboxylase  
Hotdog fold  
*Pseudomonas putida*

### ABSTRACT

Deciphering the mechanisms of bacterial fatty acid biosynthesis is crucial for both the engineering of bacterial hosts to produce fatty acid-derived molecules and the development of new antibiotics. However, gaps in our understanding of the initiation of fatty acid biosynthesis remain. Here, we demonstrate that the industrially relevant microbe *Pseudomonas putida* KT2440 contains three distinct pathways to initiate fatty acid biosynthesis. The first two routes employ conventional  $\beta$ -ketoacyl-ACP synthase III enzymes, FabH1 and FabH2, that accept short- and medium-chain-length acyl-CoAs, respectively. The third route utilizes a malonyl-ACP decarboxylase enzyme, MadB. A combination of exhaustive *in vivo* alanine-scanning mutagenesis, *in vitro* biochemical characterization, X-ray crystallography, and computational modeling elucidate the presumptive mechanism of malonyl-ACP decarboxylation via MadB. Given that functional homologs of MadB are widespread throughout domain Bacteria, this ubiquitous alternative fatty acid initiation pathway provides new opportunities to target a range of biotechnology and biomedical applications.

### 1. Introduction

The production of membrane lipids is an essential process for all known forms of cellular life. Except for Archaea, which synthesize membrane lipids from isoprenoids (De Rosa and Gliozzi, 1986), bacteria, plants, and animals utilize fatty acids for membrane production. In animals, the suite of reactions required for fatty acid biosynthesis are

performed by a type I fatty acid synthesis (FASI) pathway (Smith et al., 2003). Canonical FASI consists of a single polypeptide chain that folds to form several catalytic domains capable of iteratively elongating the covalently attached acyl chain (Smith, 1994). In contrast, most bacteria have a type II fatty acid synthesis (FASII) pathway, which is given by a distributed system of enzymes that catalyze each step of this biosynthetic process (Rock and Jackowski, 2002). The modular nature of

**Abbreviations:** ACP, acyl carrier protein; DFT, density functional theory; FAS, fatty acid synthesis; GC-MS-FAME, gas chromatography-mass spectrometry analysis of fatty acid methyl esters; GNAT, Gcn5-related N-acetyl transferase; HDF, hot-dog fold; ITC, isothermal titration calorimetry; KAS,  $\beta$ -ketoacyl-ACP synthase; LC-MS, liquid chromatography-mass spectrometry; mad, malonyl-ACP decarboxylase; MD, molecular dynamics; RMSD, root-mean-square deviation.

\* Corresponding author. Renewable Resources and Enabling Sciences Center, National Renewable Energy Laboratory, Golden, CO, 80401, USA.

E-mail address: [gregg.beckham@nrel.gov](mailto:gregg.beckham@nrel.gov) (G.T. Beckham).

<sup>1</sup> K.J.M. and E.K. contributed equally to this work.

<https://doi.org/10.1016/j.ymben.2023.02.006>

Received 1 December 2022; Received in revised form 9 February 2023; Accepted 13 February 2023

Available online 15 February 2023

1096-7176/© 2023 The Authors. Published by Elsevier Inc. on behalf of International Metabolic Engineering Society. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

FASII, in part, underlies the diversity of fatty-acid derived products bacteria can generate (Lu et al., 2004).

FASII can be broadly partitioned into three sets of reactions: initiation, elongation, and termination (Fig. 1). FASII initiation culminates in the production of a  $\beta$ -ketoacyl group covalently attached to acyl carrier protein (ACP) (Fig. 1A) (Lai and Cronan, 2003). In the case of *Escherichia coli*, *EcFabH* ( $\beta$ -ketoacyl-ACP synthase III (KASIII)) catalyzes the Claisen condensation of acetyl-CoA with malonyl-ACP to produce the  $\beta$ -acetoacetyl-ACP primer (Tsay et al., 1992). Moreover, the KASIII condensation reaction may also be catalyzed by a variant of the KASI/II domain containing protein which is exemplified by *FabY* from *P. aeruginosa* (Yuan et al., 2012b). Notably, homologs of *EcFabH* from other species have also been shown to accept alternative acyl-CoAs, including branched-chain acyl-CoAs and medium-chain-length acyl-CoAs (Qiu et al., 2005; Yuan et al., 2012a). Regardless of their identity, these varied  $\beta$ -ketoacyl-ACP primers all subsequently proceed through a series of reactions that elongate their chain length by two carbons in each catalytic cycle (Fig. 1B) (Alberts et al., 1964; Bergler et al., 1994; Garwin et al., 1980; Kass et al., 1967; Mohan et al., 1994). The resultant acyl-ACPs terminate the fatty acid elongation cycle when used as substrates in the production of membrane phospholipids or other fatty acid-derived products, such as polyhydroxyalkanoates (Fig. 1C) (Choi et al., 2020). Modulating different stages of fatty acid biosynthesis can

alter the profile of fatty acid-derived products generated and in extreme cases can potentially lead to inviability (Wright and Reynolds, 2007).

More recently an alternate KAS-independent FASII initiation pathway was discovered in *E. coli* in the form of malonyl-ACP decarboxylase (MadA, also referred to as YiiD) (Sanyal et al., 2019; Whaley et al., 2021). The Mad system was first identified in  $\Delta fabH$  strains which exhibited growth and colony morphology defects, yet remained viable, thereby suggesting the presence of a FabH-independent initiation pathway (Sanyal et al., 2019; Whaley et al., 2021). The catalytic domain of Mad is housed by a hot-dog fold (HDF) domain that may exist as a fusion with a Gcn5-related *N*-acetyl transferase (GNAT) domain (MadA), such as that discovered in *E. coli*, or as a stand-alone protein (MadB) (Whaley et al., 2021). Ultimately, the Mad system produces acetyl-ACP that can enter the elongation step via condensing enzymes such as of FabB/F (Garwin et al., 1980).

Despite landmark discoveries in model bacterial species such as *E. coli* and *P. aeruginosa*, there are significant gaps in our understanding of fatty acid biosynthesis in non-traditional model organisms. For example, although numerous publications reported the engineering of the industrially relevant microbe *Pseudomonas putida* KT2440 (Demling et al., 2021) to produce polyhydroxyalkanoates (Mezzina et al., 2021), a fatty acid-derived product, a comprehensive, mechanistic understanding of how fatty acids are produced in this species is lacking and thus prevent rationale engineering of the system. To that end, here we elucidate the three pathways for the initiation of fatty acid biosynthesis in *P. putida*. We find that *P. putida* harbors two KASIII enzymes, *FabH1* (PP\_4379) and *FabH2* (PP\_4545), that accept short- and medium-chain-length acyl-CoAs, respectively. We also demonstrate that  $\Delta fabH1 \Delta fabH2$  double mutants are not only viable, but exhibit wild type-like growth rates due to an undescribed gene, PP\_0262, that we identified independently in a forward-genetic screen. Biochemical and phylogenetic characterization of PP\_0262 established that it is a MadB enzyme. Whole-protein alanine-scanning mutagenesis, isothermal titration calorimetry (ITC), X-ray crystallography, and structural and computational studies elucidate the putative mechanism of action of MadB. In addition, we demonstrate that functional homologs of MadB are present in a variety of bacteria, highlighting the widespread nature of this enzyme in fatty acid biosynthesis.

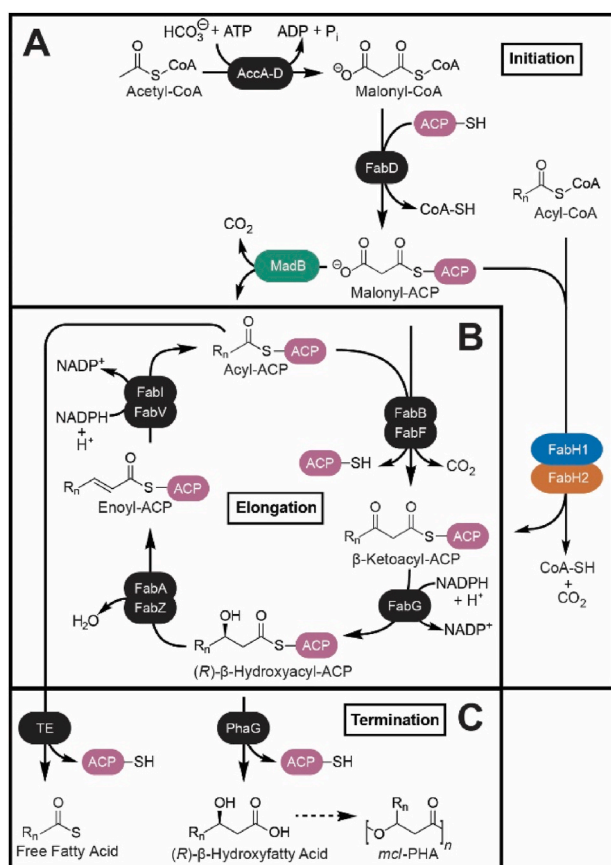
## 2. Material & Methods

All materials and methods are described in full in the Supplementary Data.

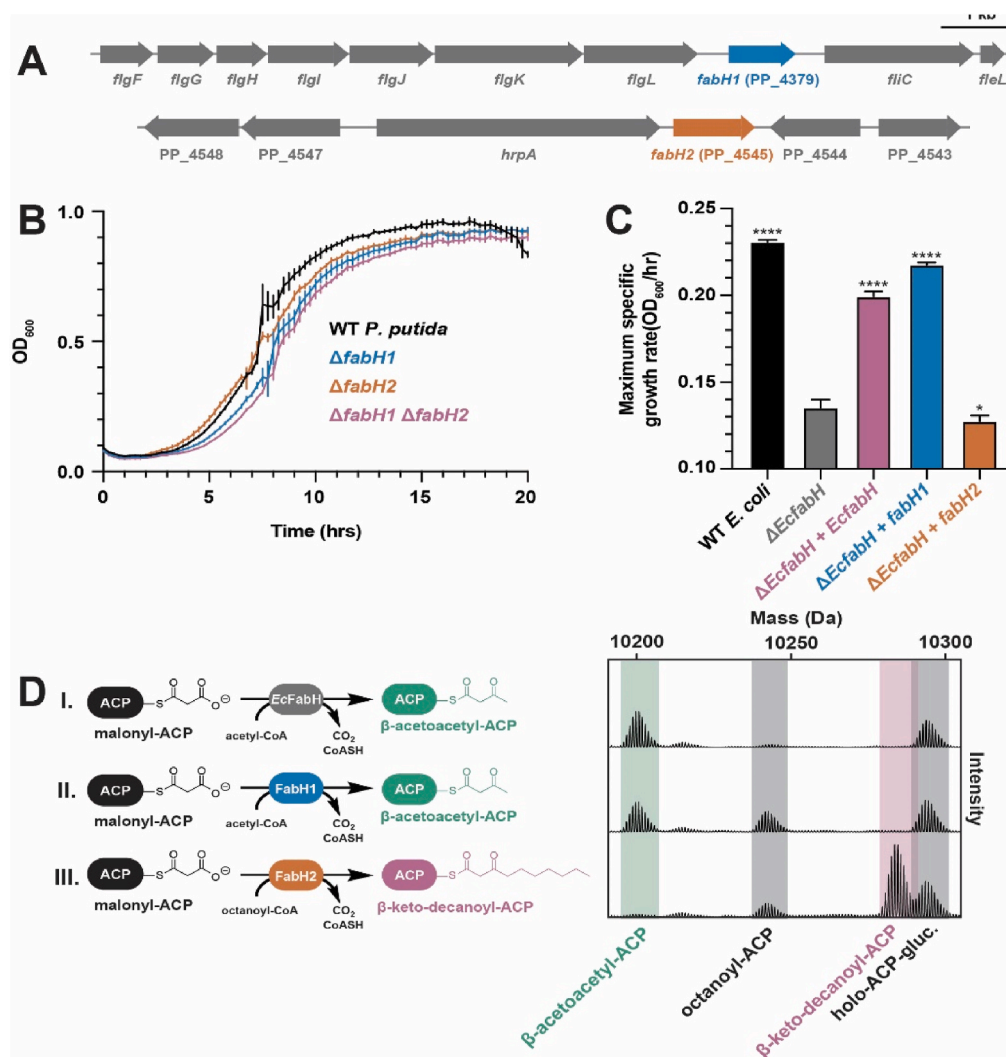
## 3. Results

***P. putida* strains lacking KASIII homologs are viable.** In an effort to identify candidate KASIII enzymes in *P. putida*, we performed a BLAST (Altschul et al., 1990) search against the *P. putida* proteome using *E. coli* *FabH* (*EcFabH*, hereafter) (Tsay et al., 1992) as the query sequence. We identified two predicted proteins that we deemed *FabH1* (PP\_4379) and *FabH2* (PP\_4545), which exhibit 29% and 26% identity to *EcFabH*, respectively (Fig. S1). Unlike the gene encoding *EcFabH*, *FabH1* and *FabH2* are not embedded within an operon that contains other fatty acid biosynthesis genes (My et al., 2013) (Fig. 2A). Characteristic of known KASIII enzymes, however, *FabH1* and *FabH2* both contain the canonical catalytic triad of residues important for enzymatic function (Davies et al., 2000) (Fig. S1). To determine what role, if any, *FabH1* and *FabH2* play in fatty acid initiation in *P. putida* grown in a modified M9 media (see *Material & Methods* for the exact composition) with glucose as the carbon source, we generated *P. putida* strains that lacked these corresponding genes both singly and in combination. Growth curve analyses of these deletion strains demonstrated that they did not exhibit a marked growth defect, but rather resembled wild type *P. putida* (Fig. 2B).

***FabH1* and *FabH2* exhibit different acyl-CoA substrate specificities.** Considering the lack of overt phenotype observed in the double



**Fig. 1. Proposed fatty acid biosynthesis pathway in *P. putida* KT2440.** (A) Initiation culminates either with the synthesis of a  $\beta$ -ketoacyl-ACP species ( $\beta$ -acetoacetyl-ACP via *FabH1* or  $\beta$ -keto-decanoyl-ACP via *FabH2*), or generation of acetyl-ACP from decarboxylation of malonyl-ACP by MadB. (B) Each catalytic cycle extends the growing acyl-ACP chain by two carbons donated from malonyl-ACP. (C) Fatty acid biosynthesis is terminated when one of the elongation intermediates is consumed to generate a fatty acid-derived product. The production of medium-chain-length polyhydroxyalkanoates from (R)- $\beta$ -hydroxyacyl-ACP and the liberation of free fatty acid by a thioesterase (TE) are shown as examples.



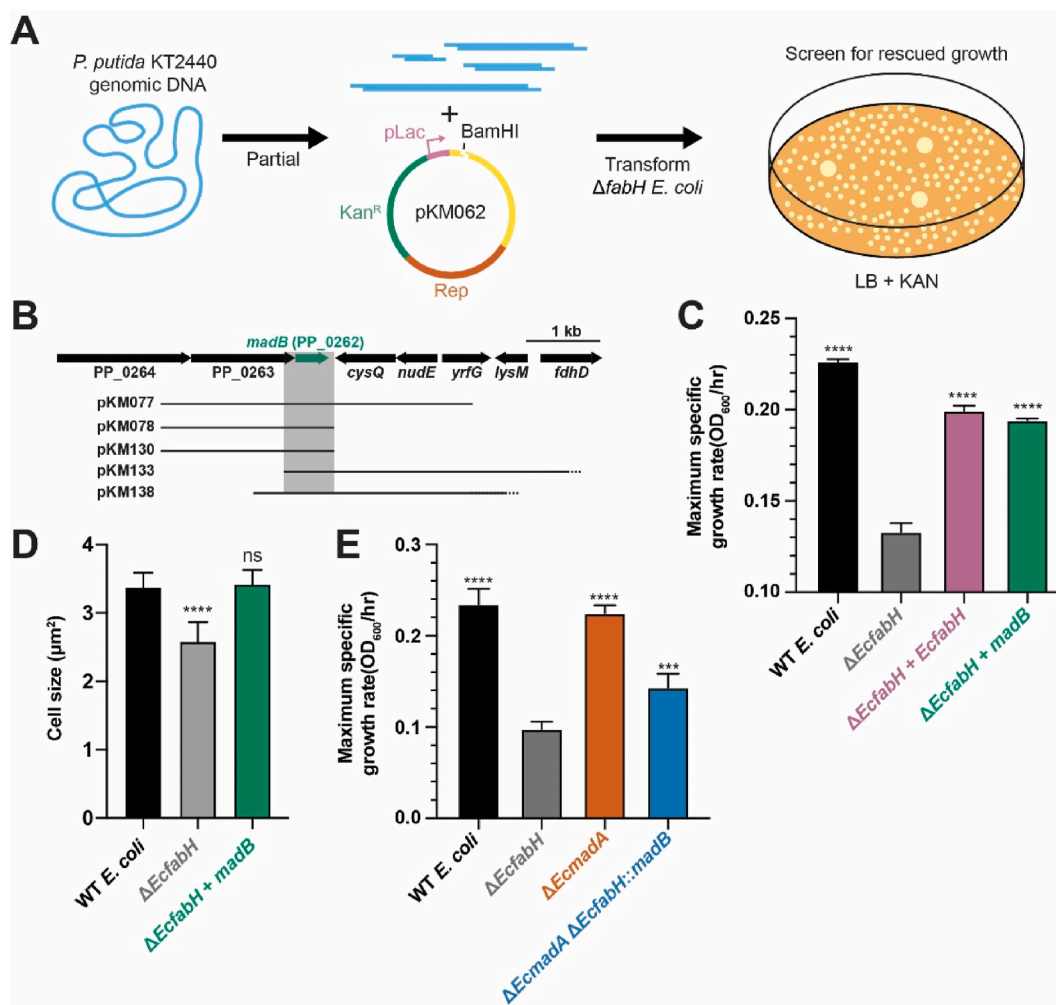
**Fig. 2.** *P. putida* contains two non-essential KASIII homologs with different acyl-CoA preferences. (A) Genomic context of identified KASIII homologs FabH1 (PP\_4379) and FabH2 (PP\_4545). (B) Growth curves of wild-type *P. putida* and indicated deletion mutants ( $n = 5$ ) grown in a modified M9 media (see *Material & Methods* for the composition) with glucose as the carbon source. (C) Maximum specific growth rates of wild-type *E. coli* and  $\Delta fabH$  *E. coli* strains transformed with pBTL-2 plasmid bearing indicated genes ( $n = 5$ ). (D) Enzymatic reactions (I, II, III) demonstrating the preferred acyl-CoA donor of EcFabH, FabH1, and FabH2 as detected by LC-MS.

deletion strain, we were curious if either FabH1 or FabH2 exhibited any fatty acid initiation activity at all. To examine the *in vivo* activity of both proteins, we expressed them heterologously in a  $\Delta fabH$  *E. coli* strain we generated using CRISPR-Cas9 (Jiang et al., 2015). While the  $\Delta fabH$  *E. coli* strain displayed a reduced maximal growth rate compared to wild type, plasmid-based expression of *EcFabH* or *fabH1* partially restored wild type-like growth in this genetic background (Fig. 2C and Fig. S2). Expression of *fabH2* did not improve the growth of the  $\Delta fabH$  *E. coli* strain (Fig. 2C and Fig. S2). This suggested that while FabH1 may have EcFabH-like activity, FabH2 may catalyze a different reaction. Indeed, FabH2 exhibits 75% amino acid sequence identity with a *Pseudomonas aeruginosa* protein (PA3286) that accepts octanoyl-CoA instead of acetyl-CoA (Yuan et al., 2012a) (Fig. S3).

We next validated the activity and specificity of the *P. putida* EcFabH homologs *in vitro* by determining their production of  $\beta$ -ketoacyl-ACP species using a high mass accuracy intact protein liquid chromatography-mass spectrometry (LC-MS). In the following assays, malonyl-ACP was produced by EcFabD-catalyzed acyl transfer reaction between holo-ACP and malonyl-CoA. In the presence of malonyl-ACP and acetyl-CoA, purified EcFabH produced  $\beta$ -acetoacetyl-ACP (Fig. 2D; Reaction I). When supplied with malonyl-ACP and a 1:1 mixture of acetyl-CoA:octanoyl-CoA, purified FabH1 only performed the condensation reaction with acetyl-CoA and malonyl-ACP producing  $\beta$ -acetoacetyl-ACP (Fig. 2D; Reaction II). In contrast, FabH2 specifically utilized octanoyl-CoA, producing  $\beta$ -keto-decanoyl-ACP (Fig. 2D; Reaction III).

The octanoyl-ACP was likely produced due to the promiscuous acyl-transfer activity of activity of FabD between acyl-ACP and octanoyl-CoA (Marcella and Barb, 2017). Overall, these *in vitro* activities are consistent with the ability and inability of FabH1 and FabH2, respectively, to rescue growth of the  $\Delta fabH$  *E. coli* strain. In addition, gas chromatography-mass spectrometry analysis of fatty acid methyl esters (GC-MS-FAME) obtained from *P. putida*  $\Delta fabH1$  and  $\Delta fabH2$  strains was consistent with their demonstrated activities (Fig. S4). Strains lacking FabH1 displayed a reduction in C<sub>14</sub> and C<sub>16</sub> species and a concomitant increase in C<sub>18</sub> species (Fig. S4), as previously seen in *E. coli*  $\Delta fabH$  strains (Yao et al., 2012). Strains lacking FabH2 had a conspicuous accumulation of  $\beta$ -hydroxyoctanoic acid, not detected in either wild-type or  $\Delta fabH1$  strains (Fig. S4), consistent with octanoyl-CoA entering the fatty acid degradation pathway (Thompson et al., 2020) rather than being shunted into fatty acid biosynthesis by FabH2. These data suggest that the viability of the  $\Delta fabH1 \Delta fabH2$  double mutant is due to an additional KASIII-independent mechanism for fatty acid biosynthesis initiation. Additionally, these observations also provide a confirmation of FabH activity in *P. putida*.

**Expression of PP\_0262 rescues *E. coli*  $\Delta fabH$  growth and cell size defects.** To elucidate other genes in *P. putida* with fatty acid initiation activity, we utilized a forward genetic screen (Fig. 3A), taking advantage of the small colony phenotype exhibited by  $\Delta fabH$  *E. coli* strains (Yao et al., 2012). Briefly, we cloned *P. putida* genomic fragments into pKM062, a derivative of pBTL-2 (Prior et al., 2010), generating a



**Fig. 3. MadB is implicated in fatty acid initiation.** (A) Schematic outlining the production and utilization of a *P. putida* genomic library to screen for novel fatty acid initiation factors. (B) Genomic fragments contained within indicated plasmids that rescued the *E. coli*  $\Delta fabH$  colony size defect. Dotted line reflects ambiguity in fragment junction. (C) Maximum specific growth rates of wild-type *E. coli* and  $\Delta fabH$  *E. coli* strains transformed with pBTL-2 plasmid bearing indicated genes ( $n = 5$ ). (D) Quantification of cell size of wild-type *E. coli*,  $\Delta EcfabH$ , and  $\Delta EcfabH + madB$  dividing cells ( $n = 8, 13, 5$ , respectively). (E) Maximum specific growth rates of wild-type *E. coli* and indicated mutant strains at 37 °C ( $n = 5$ ).

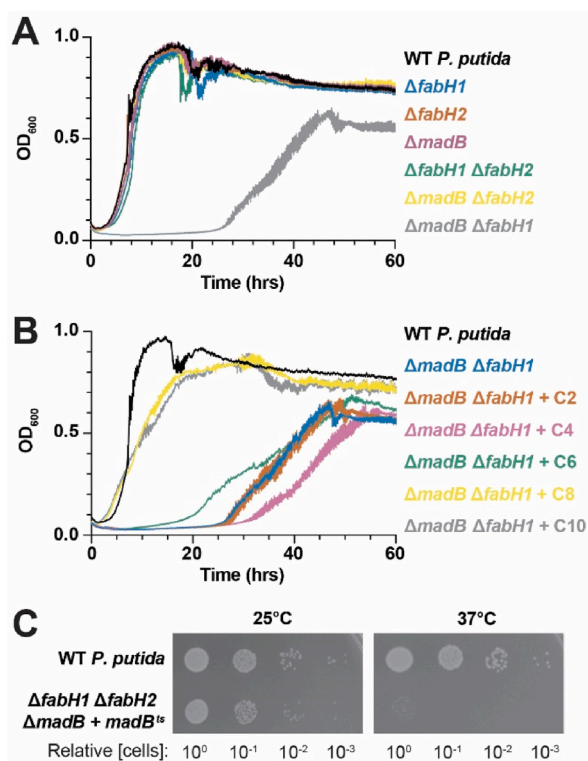
plasmid library. We then electroporated this plasmid library into a  $\Delta fabH$  *E. coli* strain, selected for kanamycin-resistance, and visually screened for large colonies corresponding to rescue of the  $\Delta EcfabH$  growth defect. Using this method, we isolated 52 large-colony transformants and subjected their respective plasmids to Sanger sequencing. Most of the recovered plasmids contained non-overlapping fragments of the *P. putida* genome. Indeed, several recovered plasmids contained no insert at all. Despite the high level of apparent false positives, we isolated five plasmids that all shared the *P. putida* genomic region containing PP\_0262, which encodes a 151 amino acid protein that we hypothesized to be a MadB enzyme (Fig. 3B). To verify that *madB* alone can rescue the  $\Delta fabH$  *E. coli* growth defect, we cloned the open reading frame of *madB* into an expression plasmid and introduced it into the  $\Delta fabH$  *E. coli* strain. Heterologous expression of *madB* in the  $\Delta EcfabH$  mutant restored wild type-like growth, comparable to the *EcfabH* control (Fig. 3C and Fig. S5).

In addition to a small colony phenotype, *E. coli* strains lacking FabH have been reported to have reduced cell size (Whaley et al., 2021; Yao et al., 2012). To determine if expression of *madB* can restore wild type-like cell size in a  $\Delta EcfabH$  strain, we used microscopy to examine actively dividing cells from wild type,  $\Delta EcfabH$ , and  $\Delta EcfabH + madB$  strains. As previously reported (Yao et al., 2012), we confirmed  $\Delta EcfabH$  cells are significantly smaller than wild-type cells ( $P < 0.0001$ ) (Fig. 3D

and Fig. S5). Expression of *madB* in  $\Delta EcfabH$  cells, notably, rescued this small cell size defect (Fig. 3D and Fig. S5). The ability of *madB* to restore both the growth and cell size defects of  $\Delta fabH$  *E. coli* strains suggests *madB* has a role in fatty acid initiation.

**MadB supports growth of *E. coli* lacking both native FabH and MadA.** Although MadB appeared to have fatty acid initiation capabilities comparable to *EcFabH*, it remained a formal possibility that MadB does not have a direct role in fatty acid initiation. For example, *E. coli* strains lacking FabH can still grow, albeit poorly, due to the presence of MadA (YiiD) (Sanyal et al., 2019; Whaley et al., 2021). Furthermore, overexpression of *madA* has been demonstrated to rescue both the growth and cell size defects of  $\Delta fabH$  *E. coli* strains (Whaley et al., 2021). To rule out any indirect effects MadB may exert on *madA* expression, we generated an *E. coli* strain that lacked *madA* and had *fabH* replaced with *madB* using CRISPR-Cas9 (Jiang et al., 2015) (Fig. 3E and Fig. S5). The rescue of the synthetically lethal  $\Delta fabH \Delta madA$  *E. coli* strain with *madB* supports a direct role for MadB in fatty acid initiation.

**Synthetic growth defect of  $\Delta madB \Delta fabH1$  *P. putida* is rescued by medium-chain-length fatty acids.** To examine the role of MadB in *P. putida* fatty acid initiation, we generated strains lacking *madB* singly and in combination with  $\Delta fabH1$  and  $\Delta fabH2$  (Fig. S6). We found that all single and double mutants, except  $\Delta madB \Delta fabH1$ , displayed wild type-like growth (Fig. 4A). This phenotype contrasts  $\Delta fabH$  *E. coli*, with an



**Fig. 4.** MadB and FabH1 are the main fatty acid initiation factors in *P. putida*. (A) Growth curves of wild-type *P. putida* and indicated deletion mutants ( $n = 5$ ). (B) Growth curves of wild-type *P. putida* and  $\Delta madB \Delta fabH1$  supplemented with the indicated length straight-chain fatty acids at 100  $\mu\text{g}/\text{mL}$  ( $n = 5$ ). (C) Serial dilutions of wild-type *P. putida* and  $\Delta fabH1 \Delta fabH2 \Delta madB$  bearing a temperature-sensitive plasmid with *madB* grown at 25 °C and 37 °C on LB agar plates for 16 h.

intact *madA*, that displayed a stunted growth rate (Fig. 3E) and diminished colony size (Fig. S5). The *P. putida* strain lacking both MadB and FabH1 exhibited a pronounced lag-phase and a reduced final optical density (Fig. 4A). GC-FAME of a  $\Delta madB$  single mutant, however, did not display overt fatty acid biosynthesis defects like the mutants lacking *fabH1* (Fig. S4). These data suggest that MadB and FabH1 are redundant in maintaining a wild type-like growth rate, but FabH1 contributes more to wild-type fatty acid biosynthesis.

Since the  $\Delta madB \Delta fabH1$  double mutant still retained FabH2, which we demonstrated can accept a medium-chain-length acyl-CoA (Fig. 2D; Reaction III), we surmised that supplementation of the minimal growth medium with medium-chain-length fatty acids may rescue the growth defect. Indeed, we observed a dramatic improvement of growth when  $\Delta madB \Delta fabH1$  was supplemented with C<sub>8</sub> and C<sub>10</sub> straight-chain saturated fatty acids (Fig. 4B). Supplementation with C<sub>6</sub> also appeared to improve growth, albeit less significantly, while C<sub>4</sub> and C<sub>2</sub> inhibited and had no effect on growth, respectively (Fig. 4B). This *in vivo* evidence supports the specificity of FabH2 towards medium-chain-length acyl-CoA substrates.

**Loss of FabH1, FabH2, and MadB is synthetically lethal.** To determine if we had identified all the fatty acid initiation enzymes in *P. putida*, we attempted to generate a *P. putida* strain lacking *fabH1*, *fabH2*, and *madB*. These initial efforts were unsuccessful. To test the hypothesis of synthetic lethality more definitively, we generated a *P. putida* strain lacking *fabH1*, *fabH2*, and *madB*, but containing a temperature-sensitive plasmid (Elmore et al., 2020) that expressed *madB*. Whereas wild-type *P. putida* showed similar growth at 25 °C and 37 °C, the triple mutant bearing the temperature-sensitive plasmid displayed a dramatic loss of viability at the restrictive temperature (Fig. 4C). We did, however, observe some growth at higher cell densities

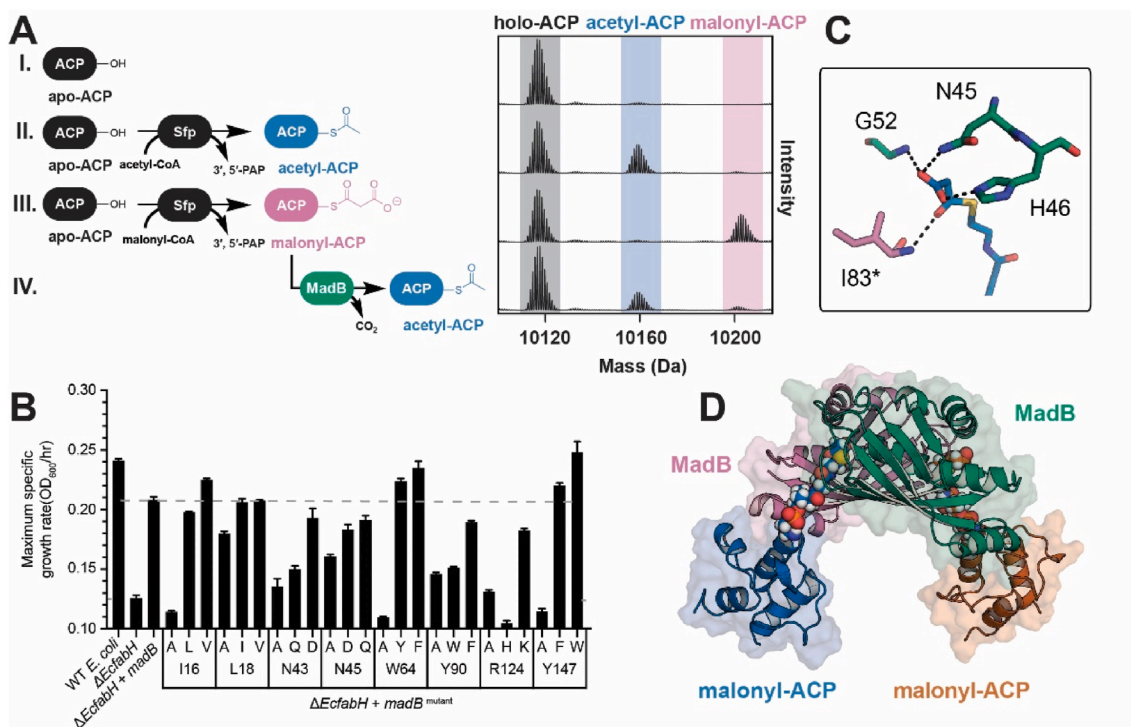
for the triple mutant strain (Fig. 4C). We reasoned that these rare cells may have retained the plasmid bearing *madB*, even at the restrictive temperature of 37 °C. Indeed, all triple mutant isolates capable of growing at 37 °C retained resistance to apramycin (35/35 colonies), which is conferred by the temperature-sensitive plasmid. These data suggest that triple mutants lacking FabH1, FabH2, and MadB are not viable and that these proteins constitute the entirety of fatty acid initiation enzymes in *P. putida*.

**MadB is a malonyl-ACP decarboxylase.** There are three potential routes to generate  $\beta$ -acetoacetyl-ACP for the initiation of fatty acid biosynthesis (Tsay et al., 1992). The first route is the direct Claisen-condensation of acetyl-CoA with malonyl-ACP to yield  $\beta$ -acetoacetyl-ACP (Tsay et al., 1992). We have shown that FabH1 is capable of this activity (Fig. 2D; Reaction II). The second route begins with the transfer of the acetyl group from acetyl-CoA onto holo-ACP, yielding acetyl-ACP. This acetyl-ACP can be subsequently condensed with malonyl-ACP, by either FabB/F (Garwin et al., 1980), resulting in  $\beta$ -acetoacetyl-ACP. The third route similarly generates the acetyl-ACP precursor but via decarboxylation of malonyl-ACP. To determine which activity, if any, MadB exhibits, we purified MadB and monitored its ability to generate relevant ACP species *in vitro* by LC-MS. We found that MadB failed to catalyze either the Claisen-condensation reaction or the acetyltransferase reaction (Fig. S7). However, addition of MadB to a solution containing malonyl-ACP resulted in a depletion of malonyl-ACP and a concomitant increase in acetyl-ACP (Fig. 5A). We also found that MadB is capable of decarboxylating malonyl-CoA and methylmalonyl-CoA to produce acetyl-CoA and propionyl-CoA, respectively, but not free malonate, thereby highlighting the importance of a pantetheine arm and the flexibility of MadB to accept a minor substitution along the malonyl group (Fig. S8).

The nature of MadB-catalyzed decarboxylation of malonyl-CoA was also investigated using ITC. The mixing of malonyl-CoA and MadB emitted a sustained heat indicating an exothermic reaction ( $\Delta H_{\text{app}} = -14.9 \text{ kJ}/\text{mol}$ ) (Fig. S9). In contrast, no heat was detected upon injection of acetyl-CoA, other than the transient heat of dilution due to the injection event (Fig. S9). The kinetics of this reaction were investigated by a multi-injection ITC method (Fig. S9), which revealed a  $K_m$  of  $12 \pm 1 \text{ mM}$  and  $k_{\text{cat}}$  of  $1.4 \pm 0.1 \text{ s}^{-1}$  for malonyl-CoA at 30 °C, a catalytic efficiency that is likely too low to be physiologically relevant (Takamura and Nomura, 1988) (Fig. S9). Hence, MadB was appropriately deemed a malonyl-ACP decarboxylase in agreement with prior determination of MadA from *E. coli* and MadB from *Shewanella oneidensis* (Whaley et al., 2021).

**Alanine-scanning mutagenesis of MadB indicates residues important for function.** To define the amino acids essential for MadB activity and gain insight into the reaction mechanism, we screened a complete library of variants harboring MadB alanine-substituted mutants for their inability to rescue the colony size defect of a  $\Delta fabH E. coli$  strain, ultimately identifying eight such MadB variants. We subsequently verified the growth defect of these variants using a microplate reader (Fig. 5B and Fig. S10). In addition, we investigated the consequence of more conservative substitutions at these eight sites identified by the alanine scanning screen (Ile16, Leu18, Asn43, Asn45, Trp64, Tyr90, Arg124, and Tyr147). In general, the more conservative substitutions were less detrimental when compared to their alanine counterpart, and some substitutions even appeared to outperform wild-type MadB suggesting opportunities for further engineering. (Fig. 5B and Fig. S10).

**MadB belongs to the hotdog fold protein family.** To better understand the mechanism of MadB-catalyzed malonyl-ACP decarboxylation, we solved the structure of MadB via X-ray crystallography at 1.04 Å resolution (PDB 8AYV, Fig. S11). MadB is structurally closest to a homolog from *Chlorobaculum tepidum* (PDB 3LMB; 38% sequence identity with a root-mean-square deviation (RMSD) of 1.6 Å) (Kube et al., 2013). We found that MadB adopts a HDF configuration, which is common among enzymes involved in fatty acid metabolism (Pidugu



**Fig. 5.** *madB* encodes a malonyl-ACP decarboxylase. (A) Enzymatic reactions (I, II, III, IV) demonstrating the conversion of malonyl-ACP to acetyl-ACP by addition of MadB as detected by LC-MS. (B) Maximum specific growth rates of wild-type and  $\Delta\text{fabH}$  *E. coli* strains transformed with pBTL-2 plasmid bearing *madB* variants ( $n = 5$ ). Dashed gray horizontal line indicates maximum specific growth rate of  $\Delta\text{fabH}$  *E. coli* expressing *madB*, above which might indicate rescue. (C) Active site of MadB with malonyl-CoA computationally docked. Dashed lines indicate the hydrogen bonding network and the asterisk represent residue from the subunit B. (D) A docking model of a MadB dimer interacting with two molecules of malonyl-ACP shown in spacefill.

et al., 2009). MadB exists as a dimer and this arrangement was observed *in crystallo* and corroborated by size-exclusion chromatography (Fig. S12).

Differential scanning calorimetry analysis of MadB at 22.2  $\mu\text{M}$  (monomer concentration) revealed that the enzyme thermally unfolds with a single transition at 66.4  $^{\circ}\text{C}$  (Fig. S13); the lack of a second peak indicates that there is no significant monomer-dimer exchange at this concentration. This melting temperature remains constant upon two-fold dilution to 11.1  $\mu\text{M}$  and 5.55  $\mu\text{M}$  monomer (Fig. S13).

Next, we sought to generate a model of MadB in complex with malonyl-ACP. We first performed a structural similarity search (Holm, 2020) and found various structures of LnmK (4% sequence identity to MadB with RMSD of 3.0  $\text{\AA}$ ) in substrate analog-bound complexes (Stunkard et al., 2021). LnmK is a previously characterized double-HDF enzyme that is a bifunctional decarboxylase/acyltransferase associated with the biosynthesis of leinamycin, a potential anti-cancer drug (Lohman et al., 2013; Stunkard et al., 2021). The similar scaffold (HDF), substrate (methylmalonyl-CoA vs. malonyl-ACP), and activity (decarboxylation) of LnmK and MadB guided our efforts to first dock malonyl-CoA into the presumed active site of MadB. An initial model ( $M_{\text{initial}}$ ) was built via structural alignment of MadB with LnmK in complex with 2-nitronate-propionyl-CoA (PDB 6X7L) (Stunkard et al., 2021), which was easily modified *in silico* to generate the malonyl-CoA ligand. We refined  $M_{\text{initial}}$  with a molecular dynamics (MD) protocol that reproduces an induced-fit process (Heo et al., 2021), and generated two top models,  $M_{\text{MD1\_CoA}}$  and  $M_{\text{MD2\_CoA}}$ , with AutoDock Vina estimated binding affinities at  $-13.9$  and  $-13.6$  kcal/mol (Trott and Olson, 2010), respectively (Fig. S14). Unlike  $M_{\text{MD1\_CoA}}$ ,  $M_{\text{MD2\_CoA}}$  implicated residues Asn45 and Arg124, identified in the alanine-scanning mutagenesis screen, as interacting with malonyl-CoA at its reaction center. In addition,  $M_{\text{MD2\_CoA}}$  established a hydrogen bond network with the terminal malonyl group analogous to that observed in LnmK (Stunkard et al., 2021) (Fig. 5C and Fig. S14). Together, these data suggest that  $M_{\text{MD2\_CoA}}$

captures legitimate interactions between MadB and malonyl-CoA.

We further performed MD-based refinement to generate a model of MadB-malonyl-ACP, using  $M_{\text{initial}}$  after the replacement of CoA with ACP from *E. coli* (PDB 4KEH; 86% identity to PP\_1915 gene product, ACP from *P. putida*). Five conformations selected from the MD simulations suggested that a binding mode like  $M_{\text{MD2\_CoA}}$  is prevalent in the presence of ACP (Fig. 5D). Typical of other ACP-interacting enzymes, MadB exhibited a remarkable complementarity in charge and hydrophobicity where it contacted helix II of ACP, also referred to as the “recognition helix” (Yadav et al., 2018) (Fig. S15). This interface between MadB and ACP is spontaneously formed in our simulations, indicating the robustness of the model.

Among the functional residues identified via alanine-scanning, our model of MadB-malonyl-ACP seems to explain the importance of residues Asn45 and Arg124 as being directly involved in substrate binding. Specifically, Asn45 and Arg124 bind the terminal carboxyl group of the malonyl and the pantetheine moieties, respectively. In addition, the identification of aromatic residues Trp64, Tyr90, and Tyr147 is likely explained by their role in forming the hydrophobic core underlying the substrate binding site (Fig. S16). The remaining identified residues, Ile16, Leu18, and Asn43, are located near the dimer interface where the proposed decarboxylation occurs (Fig. S17). To determine if these residues play an outsized role in MadB dimerization, we used the machine learning-based predictor KFC2a (Zhu and Mitchell, 2011), which predicts residues that account for the majority of the binding affinity in a complex. We found that Ile16, Leu18, and Asn43, were among the eight residues predicted by KFC2a to be important for MadB dimerization (Fig. S18). The independent identification and correlation of crucial residues in MadB through *in vivo* and *in silico* analyses strengthens the validity of our structural binding model.

**MadB catalyzes malonyl decarboxylation through carbonyl stabilization.** Based on the precedence of the catalytic mechanism proposed for LnmK (Fig. S19) (Stunkard et al., 2019, 2021) and the data

presented in this study, we propose a catalytic mechanism for MadB (Fig. 6A). This proposed mechanism is equivalent to the first-half reaction of LnmK. Based on the MD docking results, we created a theozyme with Asn45, His46, Gly52, and Ile83. In IM0, hydrogen bonds with the Gly52 backbone and the N45 side chain stabilize the anionic charge of the terminal carboxylate of malonyl-ACP, while the Ile83 backbone stabilizes the carbonyl group. Alternatively, IM1 represents a different substrate orientation, in which hydrogen bonds with Gly52 and Asn45 stabilize the carbonyl, and the His46 stabilizes the carboxyl group of the intermediate. The lone-pair electrons are subsequently delocalized from the terminal carboxylate towards the  $\beta$ -keto site, resulting in C–C bond fragmentation (decarboxylation) and the formation of an enolate intermediate (IM2). The anionic enolate in IM2 is then stabilized by the hydrogen bonds with Gly52, Asn45, and His46. By modeling effects of each amino acid residue independently, we identify the amide N–H in Asn45 and the imidazole N–H in His46 as strong hydrogen bond donors that can best stabilize an anionic substrate (see Material & Methods for more detail). Finally, deprotonation of IM2 will then readily produce the acetyl-ACP product. It is unclear whether the final resolution of the enolate would occur spontaneously via a proton exchange with a solvent species or whether the process is assisted by a protein side-chain.

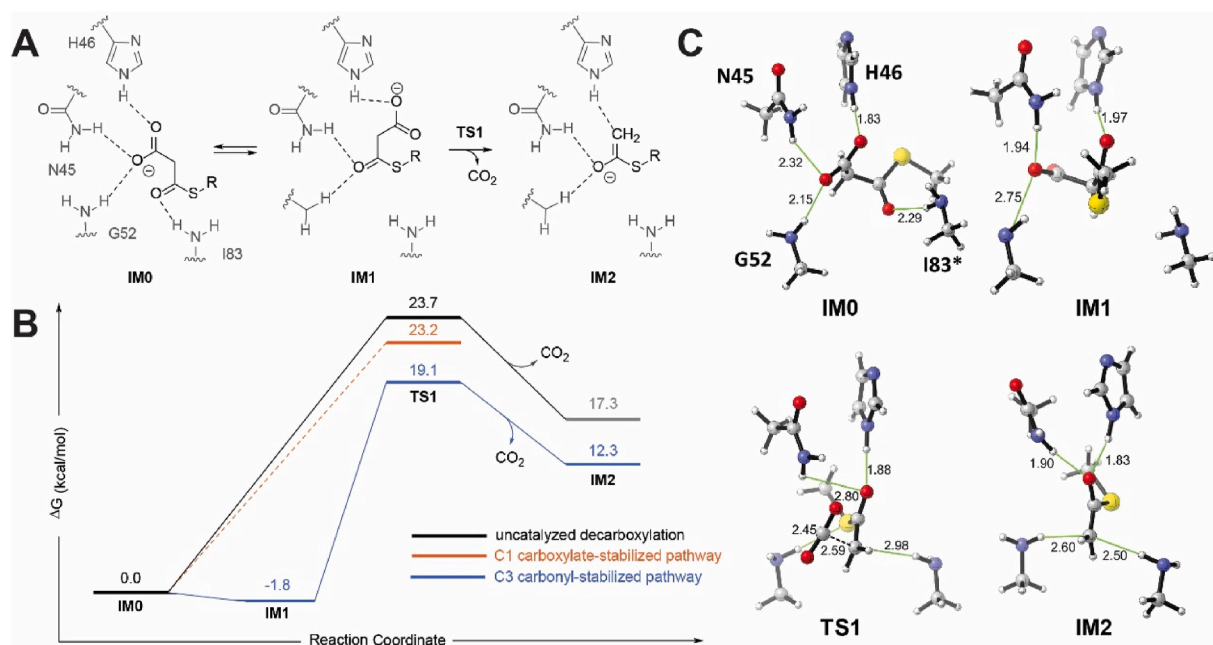
From the terminal carboxylate of malonyl-ACP (IM0), we modeled the reaction coordinate profile to form the enolate intermediate IM2 through the decarboxylation transition state TS1 (Fig. 6B). Our calculations indicated that without MadB present, the uncatalyzed decarboxylation (Fig. 6B, black) transition state has a 23.7 kcal/mol free energy of activation, which suggests that the reaction is likely to proceed slowly at room temperature. Next, we considered two possible pathways by which the amino acids (Gly52, Ile83, Asn45, and His46) can stabilize the decarboxylation transition state (TS1). These two catalyzed pathways are differentiated on whether the substrate undergoes rearrangement in the enzyme active to form a more stable orientation (IM1). Additionally, the two catalyzed pathways are also distinguished by the role of Asn45 on whether it stabilizes the C1 carboxyl (Fig. 6B, orange) or the C3 carbonyl groups (Fig. 6B, blue), of the substrate in the transition state (TS1). The transition state in the C1 carboxyl-stabilized pathway, (Fig. 6B, orange), has a kinetic barrier of 23.2 kcal/mol which is only slightly lower than the uncatalyzed reaction indicating a minor

contribution of the amino acid residues in facilitating the decarboxylation. Alternatively, MadB can more effectively promote decarboxylation along the C3 carbonyl-stabilized pathway (Fig. 6B, blue), as the transition state kinetic barrier was lowered to 19.1 kcal/mol. Closer evaluations of the DFT-optimized geometries along the pathways further highlight the effects of Asn45 and His46 in promoting the decarboxylation reaction (Fig. 6C). In the resting state (IM0), the anionic carboxylate group is stabilized by hydrogen bonding with Asn45 and Gly52, while the imidazole sidechain of His46 may interact with either the carboxylate or the carbonyl group of the substrate. In the transition state (TS1) of the C3 carbonyl-stabilized pathway, both the amide N–H in Asn45 and the imidazole N–H in His46 form hydrogen bonds with the carbonyl to stabilize the development of negative charge on the oxygen, a feature that is missing in the C1 carboxylate-stabilized pathway.

To further evaluate the relative contribution of amino acid side chains of Asn45 and His46 towards catalysis, we purified N45A and H46A variants of MadB and compared their catalytic efficiency for malonyl-CoA decarboxylation against the wild-type counterpart using the same ITC-based assay (Fig. S9). This *in vitro* experiment provides a clearer phenotype when compared to the prior *in vivo* alanine scan experiment as described previously. Similar analyses, however, are not available for Gly52 and Ile83 as their hydrogen bonding interactions are contributed by the amide main chain. Both variants were purified similarly to the wild-type MadB. As described above, the MadB-catalyzed decarboxylation of malonyl-CoA is not physiologically relevant; however, this scheme is useful for comparing the performance of the different variants. The kinetic results are summarized in Table 1.

**Table 1**  
Steady-state kinetic parameters from ITC analyses for MadB and select variants for malonyl-CoA decarboxylation.

MadB variant	$k_{cat}$ ( $s^{-1}$ )	$K_M$ (mM)	$k_{cat}/K_M$ ( $s^{-1} \cdot M^{-1}$ )
wild-type	$1.4 \pm 0.1$	$12 \pm 1$	$120 \pm 20$
N45A	$0.06 \pm 0.01$	$6.9 \pm 0.5$	$9 \pm 2$
H46A	$0.45 \pm 0.01$	$5.7 \pm 0.2$	$79 \pm 4$



**Fig. 6. Molecular mechanism of MadB-catalyzed decarboxylation of malonyl-ACP.** (A) Proposed decarboxylation mechanism. The R-group of the substrate is an ACP-bound pantetheine arm; however, a methyl group was used for the DFT calculation. (B) Free energy profile along the computed reaction coordinates. DFT-optimized key intermediates and transition states along the MadB reaction coordinates following (C) the lowest energy reaction C3 carbonyl-stabilized pathway.

While both variants have about half the Michaelis constant, their lowered turnover numbers dramatically affected the overall catalytic efficiency with N45A and H46A resulting in 13- and 1.5-fold lower catalytic efficiency, respectively when compared to the wild-type enzyme and are in line with our prior *in vivo* alanine scan result. The dramatic loss in turnover number in N45A variant is consistent with its purported role for transition state stabilization. Conversely, the less severe defect observed for H46A may stem from a less productive orientation of the bound substrate. Overall, along with the MD-derived model, our DFT calculations highlight the importance of multiple amino acids including Asn45, His46, Gly52, and Ile83 in promoting the decarboxylation reaction of the terminal carboxylate of malonyl-ACP at MadB.

**Functional homologs of MadB present in a diversity of bacterial lineages.** To determine the phylogenetic distribution of MadB homologs, we utilized an iterative profile-HMM search method, JackHMMER, using MadB as the search query (Johnson et al., 2010). We identified a total of 479 potential homologs with the same protein domain structure as MadB (Fig. 7A,B). In addition, we found 267 potential homologs containing a predicted GNAT domain appended to the N-terminus (Fig. 7A,B). One such dual-domain homolog is MadA from *E. coli* (EcMadA, hereafter Fig. 7A) (Sanyal et al., 2019; Whaley et al., 2021). We found that MadA-like proteins were restricted to Gammaproteobacteria, whereas MadB-like proteins were detected in 14 distinct phyla (Fig. 7B). To test if the homologs were functional, we surveyed the ability of 23 MadA/MadB-like proteins across 13 phyla for their ability to rescue the *E. coli*  $\Delta fabH$  growth defect (Fig. 7C and Fig. S20). We found that most of the examined homologs, both MadA- and MadB-like, conferred a growth advantage to  $\Delta fabH *E. coli* (Fig. 7C and Fig. S20). This suggests that MadA/MadB-like proteins can initiate fatty acid biosynthesis similarly in a multitude of bacteria species.$

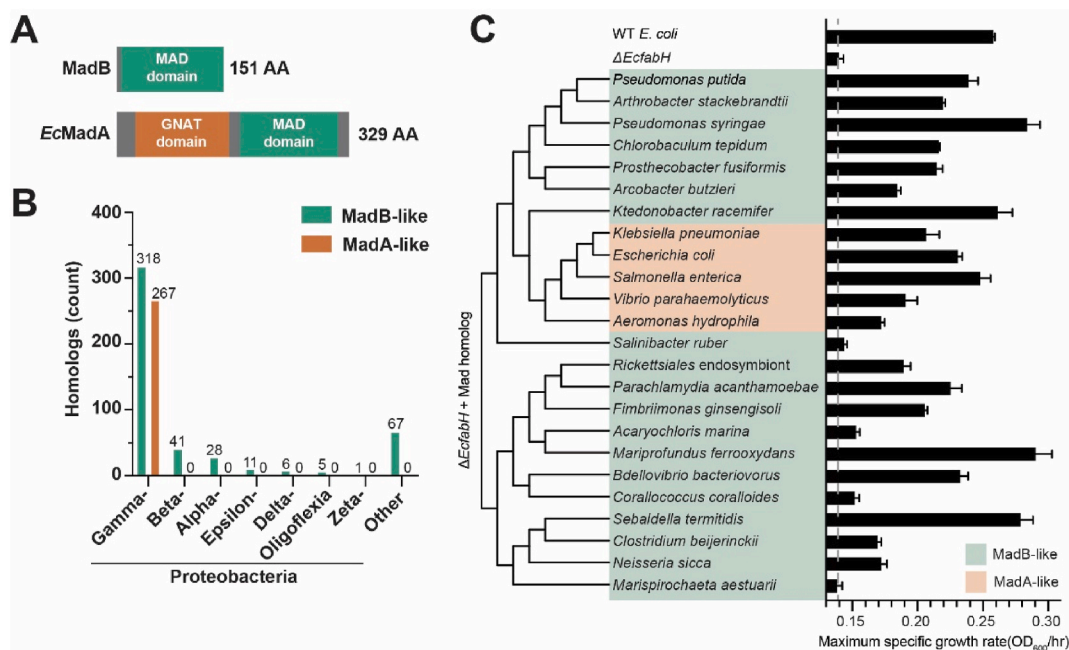
#### 4. Discussion & conclusions

Genetically dissecting the components of bacterial fatty acid biosynthesis has been traditionally difficult, in part, due to the essential

nature of this process. For example, supplementing *E. coli* growth medium with fatty acids does not guarantee the viability of a particular *fab* mutant (Lai and Cronan, 2003). Aided by genetic redundancy, here we identified and characterized the three enzymes, and three respective pathways, that initiate fatty acid biosynthesis in *P. putida* KT2440 (Fig. 1). Rather than being idiosyncratic to *P. putida*, these findings illuminate alternate fatty acid initiation in a wide variety of bacteria.

Our bioinformatics-driven approach identified FabH1 and FabH2 in *P. putida* KT2440 as potential KASIII enzymes. We found that FabH1, like EcFabH, catalyzes the Claisen condensation of acetyl-CoA with malonyl-ACP to produce  $\beta$ -acetoacetyl-ACP *in vitro* (Tsay et al., 1992). In addition, *P. putida*  $\Delta fabH1$  strains display a similar fatty acid biosynthesis defect observed in an *E. coli*  $\Delta fabH$  strain (Yao et al., 2012). For this reason, *fabH1* appears to be the gene conspicuously absent from the *P. putida* “*fabH*” operon (My et al., 2013). Sequence analysis of FabH2 revealed high identity with a *P. aeruginosa* protein, PA3286 gene product (Yuan et al., 2012a). Like PA3286 gene product, FabH2 prefers octanoyl-CoA over acetyl-CoA, and produces  $\beta$ -keto-decanoyl-ACP (Yuan et al., 2012a). In wild-type *P. putida*, FabH2 likely helps assimilate exogenous fatty acids and recycle endogenous fatty acids. Interestingly, a strain relying on FabH2 alone for growth ( $\Delta madB \Delta fabH1$ ) is viable in minimal medium, suggesting sufficient side-activity with acetyl-CoA to support *de novo* fatty acid initiation.

Despite the apparent ability of FabH1 and FabH2 to initiate fatty acid biosynthesis, we found that they were unnecessary for *P. putida* viability due to the presence of *madB*, which we identified in a forward genetic screen. Loss of MadB alone does not cause an overt defect in fatty acid biosynthesis, unlike loss of FabH1, but the  $\Delta madB \Delta fabH1$  double mutant exhibits a severe growth defect. This suggests that FabH1 and MadB to a lesser degree, are the main initiation factors in *P. putida*. We interrogated three possible routes for MadB-catalyzed fatty acid initiation and discovered that MadB is capable of malonyl-ACP decarboxylation. The acetyl-ACP species produced by MadB is presumably condensed with malonyl-ACP by FabB/F (Garwin et al., 1980), generating  $\beta$ -acetoacetyl-ACP. Whereas the FabH1 pathway only requires one



**Fig. 7. Distribution and *in vivo* activity of Mad homologs within the domain Bacteria.** (A) Protein domain structure of MadB and EcMadA. Green box indicates the HDF Malonyl-ACP Decarboxylase (MAD) domain, previously referred to as YiiD\_C (PF09500). Orange box indicates the GNAT domain (PF13673). (B) Count data for MadB/MadA-like proteins found in Proteobacteria species and beyond. (C) Maximum specific growth rates of wild-type *E. coli* and  $\Delta fabH$  *E. coli* strains transformed with pBTL-2 plasmid bearing homologs from indicated species ( $n = 5$ ). Dashed gray horizontal line indicates maximum specific growth rate of  $\Delta fabH$  *E. coli*, above which might indicate rescue. Media supplemented with 0.1 mM IPTG. Tree indicates relatedness of the protein sequences from each species generated by MUSCLE (Edgar, 2004).



malonyl-ACP and one acetyl-CoA to produce  $\beta$ -acetoacetyl-ACP, the MadB pathway requires two malonyl-ACP molecules: one to generate acetyl-ACP and another for the condensation reaction with FabB/F. This additional malonyl-ACP molecule, ultimately derived from acetyl-CoA, comes at the cost of ATP hydrolysis (Fig. 1A). This extra metabolic burden imposed by the MadB pathway may explain why the FabH1 pathway is the predominant initiation route. In *E. coli*, FabH similarly plays an outsized role in fatty acid initiation, whereas MadA is only employed under stress conditions (Sanyal et al., 2019). Future efforts in *P. putida* may elucidate whether comparable transcriptional regulation occurs with *madB*.

The decarboxylation of a malonyl group can occur through both abiotic and biotic processes. The non-enzymatic process requires a highly acidic condition and is initiated by the  $\beta$ -carbonyl group abstracting a proton from the terminal carboxylate ( $pK_a \sim 3$ ) (Hunt, 2006). Alternatively, an uncatalyzed reaction at neutral pH would encounter such a high activation energy barrier that the process would proceed at a glacial pace. By contrast, the MadB-catalyzed reaction enables FabH-independent FASII initiation such that no apparent growth defect is observed in *P. putida*  $\Delta fabH1 \Delta fabH2$  strains. In addition to HDF (Lohman et al., 2013), there are at least three additional enzyme scaffolds that have evolved means to catalyze the decarboxylation of a malonyl group at ambient conditions: crotonase (Benning et al., 2000; Maderbocus et al., 2017; Stunkard et al., 2019), GNAT (Froese et al., 2013), and the biotin dependent  $Na^+$  translocating (Buckel, 2001) protein families; all but the last protein family are cofactorless systems. Specifically for the GNAT protein family, its counterpart in MadA is non-functional and its function remains unknown (Whaley et al., 2021). Despite the distinct protein scaffolds, the active site architectures of the cofactorless systems are strikingly similar, employing elements for substrate polarization, often with the amide backbones highlighting a common catabolic logic: the polarization of the terminal carboxylate and a resonance stabilized  $\beta$ -keto-enolate intermediate.

Unlike the first half of the malonyl decarboxylation, the final resolution of the enolate intermediate (IM2) remains elusive. It is unclear whether this process occurs spontaneously, through a proton exchange with a solvent species, or is assisted by an amino acid side chain. In the case of both methylmalonyl-CoA decarboxylase (MMCD, crotonase family) and LnmK (HDF family) a flexible acid catalyst was proposed facilitate this process; but this residue is not conserved or its equivalent is missing in other enzyme systems indicating an alternate means for the keto-enol tautomerization to proceed (Lohman and Shen, 2020; Stunkard et al., 2019, 2021). In MadB, this final acid catalyst role may be fulfilled by His80 as suggested by the predicted models of malonyl-derivatives in complex with MadB. Instead, His80 may possess additional role on an earlier stage of the reaction by attracting the electron density from the terminal carboxylate towards the  $\beta$ -keto carbonyl group. Nevertheless, other adjacent basic amino acids, such as His46 and Lys 47, may at least partially compensate for the substitution H80A, explaining the non-essential role of His80 for catalysis.

The analyses provided by the structural modeling and the DFT calculation corroborates the function of the residues identified via the *in vivo* alanine scanning; namely, the direct interaction with the reaction core (Asn45) or with distal regions of the substrate (Arg124), the stabilization of the dimer assembly of the substrate-binding site (Ile16, Leu18, and Asn43), or the formation of its underlying hydrophobic core (Trp64, Tyr90, and Tyr147). Other residues that are key for catalysis according to the proposed reaction mechanism (Gly52 and Ile83) were not identified as functional residues via mutagenesis because, according to our models, their interactions with the substrate involve backbone atoms. Indeed, many of the initial substitutions in the alanine scan did not result in a complete growth defect in the  $\Delta fabH$  *E. coli* background, likely due to the complex interplay between the mutant MadB and *EcMadA*. The result was only confounded as the amino acid residues involved in catalysis serve to polarize the substrate instead of acting as a direct proton shuttle. As such, the substitution effect would be

dampened for example through the effect of ordered water or chemical rescue. Additionally, other structural features such as helix II of ACP (“recognition helix”) may play an important role in stabilizing the substrate in a reactive conformation, which may contribute to the preferential activity of MadB toward malonyl-ACP instead of malonyl-CoA through the complementary hydrophobicity (Yadav et al., 2018).

In summary, our study describes three routes for FASII initiation in *P. putida* KT2440. FabH1 and FabH2 are responsible for the conventional Claisen condensation type reactions utilizing short and medium chain acyl-CoA donors, respectively. The third route utilizes the malonyl-ACP decarboxylase system, MadB. Understanding these enzymes opens an avenue for tunability of fatty acid metabolism in *P. putida* and provides insight toward the development of a more efficient biocatalyst such as in the production of fatty-acid derived bioproducts. For example, carbon flux may be redirected to produce non-native fatty acid products by employing FabH homologs with different substrate specificities (Qiu et al., 2005). Moreover, the prevalence of malonyl-ACP decarboxylase system in the domain Bacteria also opens an avenue for biomedical applications such as in the development of novel antibiotics.

## Author contributions

Conceptualization KJM, EK, GJB, CWJ, GTB.

Data curation KJM, EK, MZahn, ETP, HS, ARP, JNG, KVH, CJS, WEM, MSM.

Formal analysis KJM, EK, ETP, HS, ARP, BCP, DV, MZhou.

Funding acquisition DAJ, CWJ, JKM, KEB-J, KNH, JEM, GTB.

Investigation KJM, EK, MZahn, ETP, HS, ARP, BCP, CDN, DV, MZhou, NM.

Methodology KJM, EK, MZahn, ETP, HS, ARP, CL, DV, MZhou, JNS.

Supervision DAJ, JNS, KNH, CWJ, YK, NM, KEB-J, GTB.

Writing – original draft KJM, EK, ETP, HS.

Writing – review & editing KJM, EK, MZahn, ETP, HS, ARP, GJB, YK, KEB-J, MZhou, JEM, CWJ, GTB.

## Declaration of competing interest

KJM, CWJ, and GTB have filed a patent application on biotechnological applications of the enzyme discovered herein. The other authors declare no competing interests.

## Data availability

All data for this paper are available in the Supporting Information.

## Acknowledgments

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U. S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding was provided to KJM, GJB, JNG KVH, BCP, CL, WEM, CJS, CDN, DV, MZhou, NM, YK, JKM, KEB, CJW, and GTB by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Bioenergy Technologies Office (BETO) for the Agile BioFoundry. EK, ETP, DAJ, and GTB acknowledge funding from The Center for Bioenergy Innovation, a U.S. Department of Energy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. MZahn, ARP, and JEM acknowledge Research England for Expanding Excellence in England (E3) funding. This research was supported by the National Science Foundation (CHE-1764328) to KNH.

A portion of this research was performed on a project award (10.46936/reso.proj. 2020.51637/60000235) from the Environmental Molecular Sciences Laboratory, a DOE Office of Science User Facility sponsored by the Biological and Environmental Research program under Contract No. DE-AC05-76RL01830. This work used the Extreme Science

and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

We thank the Diamond Light Source (Didcot, UK) for beamtime (proposal MX-23269) and the beamline staff at I03 for support.

We thank Dr. Adam Guss for providing the temperature-sensitive plasmid (pGW26).

The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ymben.2023.02.006>.

## References

- Alberts, A.W., Majerus, P.W., Talamo, B., Vagelos, P.R., 1964. Acyl-carrier protein. II. Intermediary reactions of fatty acid synthesis. *Biochemistry* 3, 1563–1571.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- Benning, M.M., Haller, T., Gerlt, J.A., Holden, H.M., 2000. New reactions in the crotonase superfamily: structure of methylmalonyl CoA decarboxylase from *Escherichia coli*. *Biochemistry* 39, 4630–4639.
- Bergler, H., Wallner, P., Ebeling, A., Leitinger, B., Fuchsichler, S., Aschauer, H., Kollenz, G., Hogenauer, G., Turnowsky, F., 1994. Protein EnvM is the NADH-dependent enoyl-ACP reductase (FabI) of *Escherichia coli*. *J. Biol. Chem.* 269, 5493–5496.
- Buckel, W., 2001. Sodium ion-translocating decarboxylases. *Biochim. Biophys. Acta* 1505, 15–27.
- Choi, S.Y., Rhie, M.N., Kim, H.T., Joo, J.C., Cho, L.J., Son, J., Jo, S.Y., Sohn, Y.J., Baritugo, K.A., Pyo, J., Lee, Y., Lee, S.Y., Park, S.J., 2020. Metabolic engineering for the synthesis of polyesters: a 100-year journey from polyhydroxyalkanoates to non-natural microbial polyesters. *Metab. Eng.* 58, 47–81.
- Davies, C., Heath, R.J., White, S.W., Rock, C.O., 2000. The 1.8 Å crystal structure and active-site architecture of beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Escherichia coli*. *Structure* 8, 185–195.
- De Rosa, M.G., A Gliozzi, A., 1986. Structure, biosynthesis, and physicochemical properties of archaeobacterial lipids. *Microbiol. Rev.* 50, 70–80.
- Demling, P., Ankenbauer, A., Klein, B., Noack, S., Tiso, T., Takors, R., Blank, L.M., 2021. *Pseudomonas putida* KT2440 endures temporary oxygen limitations. *Biotechnol* 118, 4735–4750.
- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.
- Elmore, J.R., Dexter, G.N., Francis, R., Riley, L., Huenemann, J., Baldino, H., Guss, A.M., Egbert, R., 2020. The SAGE genetic toolkit enables highly efficient, iterative site-specific genome engineering in bacteria. *bioRxiv*, 2020.06.28.176339.
- Froese, D.S., Forouhar, F., Tran, T.H., Vollmar, M., Kim, Y.S., Lew, S., Neely, H., Seetharaman, J., Shen, Y., Xiao, R., Acton, T.B., Everett, J.K., Cannone, G., Puranik, S., Savitsky, P., Krojer, T., Pilka, E.S., Kiyani, W., Lee, W.H., Marsden, B.D., von Delft, F., Allerston, C.K., Spagnolo, L., Gileadi, O., Montelione, G.T., Oppermann, U., Yue, W.W., Tong, L., 2013. Crystal structures of malonyl-coenzyme A decarboxylase provide insights into its catalytic mechanism and disease-causing mutations. *Structure* 21, 1182–1192.
- Garwin, J.L., Klages, A.L., Cronan Jr., J.E., 1980. Structural, enzymatic, and genetic studies of beta-ketoacyl-acyl carrier protein synthases I and II of *Escherichia coli*. *J. Biol. Chem.* 255, 11949–11956.
- Heo, L., Arbour, C.F., Janson, G., Feig, M., 2021. Improved sampling strategies for protein model refinement based on molecular dynamics simulation. *J. Chem. Theor. Comput.* 17, 1931–1943.
- Holm, L., 2020. DALI and the persistence of protein shape. *Protein Sci.* 29, 128–140.
- Hunt, I.R., 2006. Decarboxylation of b-carbonyl esters. In: *Organic Chemistry Etext*, 2022. McGraw-Hill.
- Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., Yang, S., 2015. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* 81, 2506–2514.
- Johnson, L.S., Eddy, S.R., Portugaly, E., 2010. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinf.* 11, 431.
- Kass, L.R., Brock, D.J., Bloch, K., 1967. Beta-hydroxydecanoyl thioester dehydrase. I. Purification and properties. *J. Biol. Chem.* 242, 4418–4431.
- Kube, M., Chernikova, T.N., Al-Ramahi, Y., Belouqui, A., Lopez-Cortez, N., Guazzaroni, M. E., Heipieper, H.J., Klages, S., Kotsyurbenko, O.R., Langer, I., Nechitaylo, T.Y., Lünsdorf, H., Fernández, M., Juárez, S., Ciordia, S., Singer, A., Kagan, O., Egorova, O., Petit, P.A., Stogios, P., Kim, Y., Tchigvintsev, A., Flick, R., Denaro, R., Genovese, M., Albar, J.P., Reva, O.N., Martínez-Gomariz, M., Tran, H., Ferrer, M., Savchenko, A., Yakunin, A.F., Yakimov, M.M., Golyshina, O.V., Reinhardt, R., Golyshin, P.N., 2013. Genome sequence and functional genomic analysis of the oil-degrading bacterium *Oleispira antarctica*. *Nat. Commun.* 4, 2156.
- Lai, C.Y., Cronan, J.E., 2003. Beta-ketoacyl-acyl carrier protein synthase III (FabH) is essential for bacterial fatty acid synthesis. *J. Biol. Chem.* 278, 51494–51503.
- Lohman, J.R., Bingman, C.A., Phillips Jr., G.N., Shen, B., 2013. Structure of the bifunctional acyltransferase/decarboxylase LnmK from the leinamycin biosynthetic pathway revealing novel activity for a double-hot-dog fold. *Biochemistry* 52, 902–911.
- Lohman, J.R., Shen, B., 2020. The LnmK bifunctional acyltransferase/decarboxylase specifying (2R)-methylmalonyl-CoA and employing substrate-assisted catalysis for polyketide biosynthesis. *Biochemistry* 59, 4143–4147.
- Lu, Y.J., Zhang, Y.M., Rock, C.O., 2004. Product diversity and regulation of type II fatty acid synthases. *Biochem. Cell. Biol.* 82, 145–155.
- Maderbocus, R., Fields, B.L., Hamilton, K., Luo, S., Tran, T.H., Dietrich, L.E.P., Tong, L., 2017. Crystal structure of a *Pseudomonas* malonate decarboxylase holoenzyme hetero-tetramer. *Nat. Commun.* 8, 160.
- Marcella, A.M., Barb, A.W., 2017. The R117A variant of the *Escherichia coli* transacylase FabD synthesizes novel acyl-acyl carrier proteins. *Appl. Microbiol. Biotechnol.* 101, 8431–8441.
- Mezzina, M.P., Manoli, M.T., Prieto, M.A., Nikel, P.I., 2021. Engineering native and synthetic pathways in *Pseudomonas putida* for the production of tailored polyhydroxyalkanoates. *Biotechnol. J.* 16, e2000165.
- Mohan, S., Kelly, T.M., Eveland, S.S., Raetz, C.R., Anderson, M.S., 1994. An *Escherichia coli* gene (*fabZ*) encoding (3R)-hydroxymyristoyl acyl carrier protein dehydrase. *J. Biol. Chem.* 269, 32896–32903.
- My, L., Rekoske, B., Lemke, J.J., Viala, J.P., Gourse, R.L., Bouveret, E., 2013. Transcription of the *Escherichia coli* fatty acid synthesis operon *fabHDG* is directly activated by FadR and inhibited by ppGpp. *J. Bacteriol.* 195, 3784–3795.
- Pidugu, L.S., Maity, K., Ramaswamy, K., Surolia, N., Suguna, K., 2009. Analysis of proteins with the 'hot dog' fold: prediction of function and identification of catalytic residues of hypothetical proteins. *BMC Struct.* 10, 37.
- Prior, J.E., Lynch, M.D., Gill, R.T., 2010. Broad-host-range vectors for protein expression across gram negative hosts. *Biotechnol. Bioeng.* 106, 326–332.
- Qiu, X., Choudhry, A.E., Janson, C.A., Grooms, M., Daines, R.A., Lonsdale, J.T., Khandekar, S.S., 2005. Crystal structure and substrate specificity of the beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Staphylococcus aureus*. *Protein Sci.* 14, 2087–2094.
- Rock, C.O., Jackowski, S., 2002. Forty years of bacterial fatty acid synthesis. *Biochem. Biophys. Res. Commun.* 292, 1155–1166.
- Sanyal, R., Singh, V., Harinarayanan, R., 2019. A novel gene contributing to the initiation of fatty acid biosynthesis in *Escherichia coli*. *J. Bacteriol.* 201.
- Smith, S., 1994. The animal fatty acid synthase: one gene, one polypeptide, seven enzymes. *Faseb. J.* 8, 1248–1259.
- Smith, S., Witkowski, A., Joshi, A.K., 2003. Structural and functional organization of the animal fatty acid synthase. *Prog. Lipid Res.* 42, 289–317.
- Stunkard, L.M., Dixon, A.D., Huth, T.J., Lohman, J.R., 2019. Sulfonate/nitro bearing methylmalonyl-thioester isosteres applied to methylmalonyl-CoA decarboxylase structure-function studies. *J. Am. Chem. Soc.* 141, 5121–5124.
- Stunkard, L.M., Kick, B.J., Lohman, J.R., 2021. Structures of LnmK, a bifunctional acyltransferase/decarboxylase, with substrate analogues reveal the basis for selectivity and stereospecificity. *Biochemistry* 60, 365–372.
- Takamura, Y., Nomura, G., 1988. Changes in the intracellular concentration of acetyl-CoA and malonyl-CoA in relation to the carbon and energy metabolism of *Escherichia coli* K12. *J. Gen. Microbiol.* 134, 2249–2253.
- Thompson, M.G., Incha, M.R., Pearson, A.N., Schmidt, M., Sharpless, W.A., Eiben, C.B., Cruz-Morales, P., Blake-Hedges, J.M., Liu, Y., Adams, C.A., Haushalter, R.W., Krishna, R.N., Lichtner, P., Blank, L.M., Mukhopadhyay, A., Deutschbauer, A.M., Shih, P.M., Keasling, J.D., 2020. Fatty acid and alcohol metabolism in *Pseudomonas putida*: functional analysis using random barcode transposon sequencing. *Appl. Environ. Microbiol.* 86.
- Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461.
- Tsay, J.T., Oh, W., Larson, T.J., Jackowski, S., Rock, C.O., 1992. Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (*fabH*) from *Escherichia coli* K-12. *J. Biol. Chem.* 267, 6807–6814.
- Whaley, S.G., Radka, C.D., Subramanian, C., Frank, M.W., Rock, C.O., 2021. Malonyl-acyl carrier protein decarboxylase activity promotes fatty acid and cell envelope biosynthesis in Proteobacteria. *J. Biol. Chem.* 297, 101434.
- Wright, H.T., Reynolds, K.A., 2007. Antibacterial targets in fatty acid biosynthesis. *Curr. Opin. Microbiol.* 10, 447–453.
- Yadav, U., Arya, R., Kundu, S., Sundd, M., 2018. The "recognition helix" of the type II acyl carrier protein (ACP) utilizes a "ubiquitin interacting motif (UIM)"-like surface to bind its partners. *Biochemistry* 57, 3690–3701.

- Yao, Z., Davis, R.M., Kishony, R., Kahne, D., Ruiz, N., 2012. Regulation of cell size in response to nutrient availability by fatty acid biosynthesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2561–E2568.
- Yuan, Y., Leeds, J.A., Meredith, T.C., 2012a. *Pseudomonas aeruginosa* directly shunts beta-oxidation degradation intermediates into de novo fatty acid biosynthesis. *J. Bacteriol.* 194, 5185–5196.
- Yuan, Y., Sachdeva, M., Leeds, J.A., Meredith, T.C., 2012b. Fatty acid biosynthesis in *Pseudomonas aeruginosa* is initiated by the FabY class of  $\beta$ -ketoacyl acyl carrier protein synthases. *J. Bacteriol.* 194, 5171–5184.
- Zhu, X., Mitchell, J.C., 2011. KFC2: a knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features. *Proteins* 79, 2671–2683.