Biotin is an essential micro-nutrient across the three domains of life. The paradigm earlier step of biotin synthesis denotes “BioC-BioH” pathway in *Escherichia coli*. Here we report that BioZ bypasses the canonical route to begin biotin synthesis. In addition to its origin of Rhizobiales, protein phylogeny infers that BioZ is domesticated to gain an atypical role of β-ketoacyl-ACP synthase III. Genetic and biochemical characterization demonstrates that BioZ catalyzes the condensation of glutaryl-CoA (or ACP) with malonyl-ACP to give 5ʹ-keto-pimeloyl ACP. This intermediate proceeds via type II fatty acid synthesis (FAS II) pathway, to initiate the formation of pimeloyl-ACP, a precursor of biotin synthesis. To further explore molecular basis of BioZ activity, we determine the crystal structure of *Agrobacterium tumefaciens* BioZ at 1.99 Å, of which the catalytic triad and the substrate-loading tunnel are functionally defined. In particular, we localize that three residues (S84, R147, and S287) at the distant bottom of the tunnel might neutralize the charge of free C-carboxyl group of the primer glutaryl-CoA. Taken together, this study provides molecular insights into the BioZ biotin synthesis pathway.
Biotin is an essential enzyme cofactor for all living organisms across the three domains of life. The importance of biotin as a prosthetic group is attributed to its participation in the transfer of one carbon units (CO₂) in a number of reactions such as carboxylation, decarboxylation, and trans-carboxylation. Biotin-dependent enzymes include acetyl-CoA carboxylase subunit B (AcclB) of fatty acid synthesis, methyl-crotonyl-CoA carboxylase (MCC) and propionyl-CoA carboxylase (PCC) of amino acid metabolism, and pyruvate carboxylase (PC) of gluconeogenesis. Unlike plants and a number of microbes capable of de novo synthesis of biotin, mammals like humans rely on exogenous supply of biotin from either diet or the gut symbiotic microbiota. Dysfunction in biotin homeostasis is associated with a number of human pathologies, especially neurological disorders. In addition to its classical roles in biotin-dependent enzymes, certain noncanonical biotin utilization was also reported, which include histone biotinylilation and signal transduction in patients.

Structurally, biotin is a sultam-containing derivative of a C7 fatty acid, which consists of a bicyclic ring and a valeric acid side chain. Since its initial discovery in 1901, biochemical and physiological roles of biotin have been gradually elucidated. Our current knowledge on biotin metabolism mostly arisen from studies with the two model organisms, Escherichia coli and Bacillus subtilis. However, our understanding of the diversity of bacterial biotin biosynthetic pathways remains fragmentary. In general, biotin synthesis proceeds in two stages: (i) the generation of pimelate moiety, an atypical ω-dicarboxylic acid of seven-carbons; and (ii) the assembly of a fused heterocyclic rings of biotin. The latter steps are highly conserved in that it consistently proceed via four successive reactions catalyzed by four unique enzymes. Namely, they correspond to 8-amino-7-oxo-gononanate synthase (AONS) encoded by bioF, 7,8-diaminogononanate synthase (DANS), the product of bioA gene, and biotin synthase (DBTS) by bioD, and the biotin synthase BioB. In contrast, the initial steps of pimelate acid synthesis differ greatly amongst the varied bacterial lineages. In the paradigm organism E. coli, the BioC-BioH machinery exploits a disguising strategy of “methylation-to-demethylation” which allows methyl-malonate-ACP thioester as an unusual primer to trigger two cycles of type II fatty acid synthesis (FAS II), giving pimeloyl-ACP.

Not surprisingly, a large number of bioH-negative bacteria that retain bioC-BioH have evolved distinct nonhomologous isoenzymes to compensate for the loss of BioH. As expected, all the BioH functional equivalents belong to the family of α/β hydrolases, which include BioG of Haemophilus, the Synechococcus BioB, BioF in Francisella, and BioV from Helicobacter. This increases the diversity of demethylases in the context of biotin synthesis.

The second well-studied mechanism by which the pimeloyl moiety is generated, involves the “BioI-BioW” pathway of Bacillus subtilis and its relatives. Among them, the cytochrome P450 enzyme BioI in Bacillus, cleaves a specific carbon–carbon bond (C7–C8) of long chain acyl-ACP to liberate an equivalent of pimelinate. Complex structures of BioI liganded with different acyl-ACPs illustrate how fatty acyl chains are shaped into a “U-turn” above the heme iron cofactor within the enzyme active site. BioW of Bacillus acts as a pimeloyl-CoA synthetase, converting free pimelate acids into the thioester of pimeloyl-CoA via an ATP-dependent mechanism. Crystal structures of Biow enzymes from B. subtilis and Aquifex aeolicus provided mechanistic insights into the ligation of coenzyme A (CoA) with pimelic acid. 13C-labeling experiments by Manandhar and Cronan supplemented in vivo evidence that free pimelate also originates from the FAS II pathway in B. subtilis. In contrast to pimeloyl-ACP being preferentially used by the E. coli BioF as a substrate, pimeloyl-CoA is exclusively utilized by the B. subtilis BioF as a bona fide precursor of biotin synthesis. Therefore, the “BioI-BioW” pathway of B. subtilis represents a unique earlier route committed to pimelate production which is far distinct from the “BioC-BioH” route in E. coli.

The machinery of pimelate production seems strikingly different in the phylum of α-proteobacteria (like A. tumefaciens) because it has neither “bioC-bioH” nor “bioI-bioW” homologous pathways. In contrast, a fabH-like gene, bioZ, seems to be domesticated into a unique bioBFDAZ operon in certain α-proteobacterial species (Supplementary Fig. 1). This promoted the hypothesis that BioZ defines a third pathway for biotin precursor synthesis. However, the biochemical mechanism of BioZ remains an enigma as of this manuscript preparation. In this work, we aim at shedding some light on the role of BioZ in biotin precursor synthesis. Not only do we propose that BioZ arises from the FabH branch within the family of FAS enzymes, but also present structural basis for BioZ action in biotin synthesis. More importantly, we report integrative evidence that BioZ initiates the synthesis of 5-keto-pimeloyl-ACP from glutaryl-CoA along with malonyl-ACP. Therefore, this finding constitutes a functional proof that BioZ bypasses the early canonical steps of biotin synthesis, unveiling a previously unknown for pimelate production.

Results

Requirement of biotin for A. tumefaciens growth. The bioBFDA operon of A. tumefaciens encodes four enzymes responsible for the latter steps of biotin synthesis, assuring the physiological requirement of biotin. Presumably, the removal of this operon impairs its capabilities of biotin synthesis, giving the biotin-auxotrophic strain ΔbioBFDA as an indicator strain. As expected, the ΔbioBFDA mutant cannot appear on the M9 minimal medium without biotin in this biotin assay (Fig. 1b). The growth detection of this mutant was significantly restored by the exogenous addition of biotin (2–10 pmol, Fig. 1b), rather than its precursor DTB (Fig. 1c). The inability of DTB in supporting bacterial growth of ΔbioBFDA is due to the lack of Bioß, an essential biotin-converting enzyme from DTB (Fig. 1a). On the basis of DTB/biotin assay, cross-feeding experiments were also performed involving two different combinations (donor and recipient). Namely, (i) the wild-type strain of A. tumefaciens, NTL4 is a DTB/biotin producer, and in the ΔbioBFDA mutant acts as a recipient strain; (ii) cell-free bacterial supernatants of A. tumefaciens and Klebsiella pneumoniae Kp24 strain serve as the DTB/biotin donors, whereas the ER90 (ΔbioFCD) strain of E. coli functions as a DTB/biotin sensor. Not surprisingly, the ΔbioBFDA NTL4 strain (2 μl of log-phase culture) can cross-feed the biotin-auxotrophic strain of ΔbioBFDA, allowing its growth on the nonpermissive condition (Fig. 1d). This indicated that appreciable level of DTB/biotin is excreted by the A. tumefaciens NTL4 strain into its growth environment. In fact, we found that the cell-free supernatant of NTL4 strain (~5 μl) supports robust growth of ER90, as biotin (5 pmol) does (Fig. 1e). However, no obvious growth was detected despite of the culture medium (5 μl) of K. pneumoniae Kp24 strain. Evidently, this observation confirms that (i) biotin is an essential micronutrient for A. tumefaciens growth; and (ii) the NTL4 strain of A. tumefaciens can secret more DTB into growth medium than K. pneumoniae Kp24 does.

Functional replacement of BioC-BioH with BioZ. In the context of lipid metabolism, the phylum of α-proteobacterium such as A. tumefaciens has a higher frequency of gene duplication and
amplification when compared to E. coli (Supplementary Fig. 1). On the two chromosomes of A. tumefaciens, an array of redundant fab-related loci is dispersed (Supplementary Fig. 1), which includes two fabI genes [fabI1 (atu0149) and fabI2 (atu0752)], 4 fabF genes [fabF1 (atu1097) to fabF4 (atu4216)], and 4 acpP genes [acpP1 (atu1096) to acpP4 (atu8162)]. A similar distribution is seen in both Brucella melitensis and Rhizobium sp. IRBG74. While fabG is present as a single copy in A. tumefaciens, it appears as five copies in B. melitensis and seven copies in Rhizobium sp. IRBG74 (Supplementary Fig. 1). Intriguingly, a second copy of fabH gene that presumably encodes a β-ketoacyl-ACP synthase III (KAS III) of fatty acid synthesis, is consistently integrated into a single
bioBFAZ operon of biotin synthesis in certain α-proteobacterial species, like A. tumefaciens (Supplementary Figs. 1-2). Therefore, it was called bioZ (fabH2), in A. tumefaciens33, 35, B. melitensis34, and Rhizobium35-38 (Supplementary Fig. 1). An earlier genetic study by Sullivan and coworkers35, demonstrated that a transposon-based inactivation of bioZ impairs the growth of Mesorhizobium sp. strain R7A on media lacking biotin and the complementation of the bioZ gene into the bioH mutant of E. coli restored its viability on biotin-lacking media.

To functionally characterize BioZ in biotin metabolism, we cloned bioZ from three different species into the arabino-inducible plasmid pBADD4 (Supplementary Table 1). We then assayed the in vivo role of BioZ using E. coli biotin-auxotrophic strains (ΔbioH, ΔbioC, or the ΔbioH/ΔbioC double mutant) (Supplementary Figs. 2-5a). As predicted, the introduction of Brucella bioZ can rescue bacterial growth of both ΔbioH and ΔbioC on the biotin-free, nonpermissive growth condition (Supplementary Fig. 3). Furthermore, the expression of bioZ confers efficient growth of the ΔbioH/ΔbioC double mutant on biotin-deficient media (Supplementary and Fig. 3). Similar results were obtained with the A. tumefaciens and Rhizobium bioZ (Supplementary Fig. 4). Consequently, BioZ (FabH2) bypasses the roles of “BioC-BioH” in the canonical earlier steps of biotin precursor synthesis. From an evolutionary point of view, this result suggested that BioZ has been domesticated from a prototypical fatty acid synthetic enzyme to gain an additional role in biotin synthesis. Thus, BioZ could invoke a third pathway by which the biotin precursor pimelate is synthesized in α-proteobacteria, esp. A. tumefaciens.

Origin and phylogeny of BioZ. To probe the possible origin of BioZ, a database-wide BLAST search against GenBank was performed with manual collation. The majority of protein candidates obtained belong to the family of KAS enzymes: KAS I (FabF), KAS II (FabB), and KAS III (FabH1 and FabH2). BLAST search resulted in almost 150 hits, including 41 FabF, 15 FabB, 50 FabH, and 43 BioZ (Fig. 1a). Since these candidates were highly variable by homology searches, we thereby reevaluated their homologous domains via Pfam server (https://pfam.xfam.org/). Unlike the fab loci scattered on chromosomes, the fabH-like locus bioZ is frequently found to localize within the biotin synthesis gene cluster (Supplementary Fig. 1). These results, taken into the genetic context of neighboring genes, suggested that bioZ probably arises from recent integration events.

As shown in the unrooted tree, two distinct clades (FabF and FabH) are present (Fig. 2a). The FabH clade seem to be classified into three subgroups, one of which belongs to BioZ (Fig. 2a). Evidently, BioZ has likely evolved from FabH in certain species of Rhizobiales. During the domestication process with unknown selection pressure, such FabH ancestors might be accidently harnessed or hijacked to participate in biotin synthesis, despite its relative-low initial efficiency. Likewise, FabB is represented as a subclan within the FabF group, far different from the former FabH lineage (Fig. 2a). This suggests that FabB can be an ancestor of FabB. Noteworthy, A. tumefaciens C58 has evolved all the homologs across the aforementioned subclades (Fig. 2a), the specialized roles of which might benefit the evolutionary adaptability and metabolic fitness of Agrobacterium.

The phylogeny of BioZ homologs further inferred their potential evolutionary correlation across different species (Fig. 2b). Intriguingly, three highly similar bioZ genes (mll9094, mll5827, and mldr6070) were found in Mesorhizobium japonicum MAFF303099, indicating that the origination and evolution of BioZ is connected with gene duplication or horizontal transfer (Fig. 2b). In brief, the pMLA plasmid-borne mll9094 is 91.74% similar to the chromosomal mll5827 at amino acid sequence level. And mldr6070 is 86.52% identical to mll5827, of which sequence is 118 aa in size, significantly shorter than the BioZ prototype (>300aa). However, the core domains of this truncated chimera and mll5827 are both indicative of BioZ-type enzymes. In particular, we observed that the loci adjacent to mldr6070 (i.e., mll6064, mll6066-mll6068, mldr6069, and mll6075) encode a number of transposon elements40, 41. This underscored the possibility of on-going dissemination/variation of bioZ gene. The truncated version of mldr6070 is the remnant or degenerative in the domestication/evolution of bioZ, or alternatively, this homologous gene might provide fragment insertion sites. Therefore, these observations supported our proposal that bioZ originates from Rhizobiales (Fig. 2b). In summary, bioZ could be subjected to independent diversification rather than co-evolution with other biotin synthesis related genes.

Biochemical insights into BioZ action. To address its biochemical activity, we expressed BioZ from different organisms using a prokaryotic expression system (Supplementary Fig. 5a). Among the three bioZ-inserted constructs we examined, only AtBioZ could be overexpressed in the form of partially soluble protein, exhibiting an apparent mass of ~36 kDa (Supplementary Fig. 5b). The purity of the recombinant AtBioZ protein was determined with gradient PAGE (Supplementary Fig. 5b). Then, mass spectrometry validated its polypeptide fingerprint with a 44% sequence coverage (Supplementary Fig. 5c). Both gel filtration and chemical cross-linking experiment of AtBioZ revealed that its solution structure is of dimeric stoichiometry (Supplementary Fig. 5d, e). Despite of its capability to complement the E. coli ΔbioH mutant (Supplementary Figs. 3-4), AtBioZ cannot liberate the methyl group from the paradigm substrate of BioH, methyl-pimeloyl-ACP, to give pimeloyl-ACP, in the in vitro enzymatic system (Supplementary Fig. 6a-b). Using the in vitro reconstituted system of DTB/biotin synthesis, the resultant product by BioJ restored the growth of the biotin-auxotrophic ER90 strain on the biotin-lacking, nonpermissive condition27-42, whereas not for AtBioZ (Supplementary Fig. 6c-d). This underscored that BioZ represents a previously-unrecognized mechanism, and needed further experimental determination.

Given that AtBioZ originated from FabH, we then asked the two questions: (i) whether or not it retains the β-ketoacyl-ACP synthase III activity; (ii) how the primer substrate specificity differs between AtBioZ (AtFabH2) and AtFabH1. Thus, we purified the enzymes of fatty acid synthesis and reconstituted the in vitro FAS II system (Supplementary Figs. 7 and 8a). As for the initial reaction of fatty acid synthesis, a variety of FabH enzymes might recognize acyl-CoAs of different carbon chain lengths (Supplementary Fig. 8). Of note, the condensation of primer substrate by FabH with malonyl-ACP proceeds via the successive activities of FabG, FabZ (FabA) and FabI using NADH and NADPH as coenzymes. As expected, the E. coli FabH enzyme (EcFabH, control) was observed to ligate exclusively acetyl-CoA with malonyl-ACP to form aceto-acyl-ACP which enters the FAS II cycle, giving butyryl-ACP (C4-ACP) (Fig. 3a and Supplementary Fig. 8b). The AtFabH1 was detected active with at least two substrates, acetyl-CoA (C2-CoA) and butyryl-CoA (C4-CoA), suggesting its enzymatic promiscuity (Supplementary Fig. 8c-d). Unlike the scenario with AtFabH1, it seemed likely that AtBioZ catalyzes the condensation of primer substrate glutaryl-CoA with malonyl-ACP, producing a pool of C7-ACP (Supplementary Figs. 8e and 9). The resultant product mixture displayed a similar mobility of pimeloyl-ACP (positive control), slightly above holo-ACP and malonyl-ACP in the separation by conformationally sensitive, 0.5 M urea/PAGE (17.5%, pH9.5) (Fig. 3a). The pimeloyl-ACP positive control used in our assay
Fig. 2 Phylogeny of BioZ and its paralogs. a An unrooted tree of BioZ proteins and its putative homologs (FabH, FabB, and FabF). The phylogenetic tree was generated with the MEGA7 software using the NJ method (bootstrap: 1000 replicates). Two major clades of KAS enzymes are shown: FabF in red and FabH in orange. In contrast, FabB (shown in gray) is phylogenetically positioned as a sub-branch of FabF clade and BioZ was localized to a sub-branch of FabH clade. Therefore, we speculated that FabB originates from FabF (KAS II), and a recent ancestor of BioZ arises from FabH, the KAS III enzyme within the KAS pan-family. b Phylogenetic relationships of BioZ paralogs. Phylogeny was constructed using the MEGA7 software involving the NJ method (bootstrap: 1000 replicates, bootstrap values indicated by circle sizes). BioZ of Agrobacterium and Brucella melitensis are colored in red and blue, respectively.
was produced in our lab by the demethylation of pimeloyl-ACP methyl ester using the BioJ enzyme (Supplementary Fig. 7b). Not surprisingly, we also observed that BioZ possess the activity on glutaryl-ACP along with malonyl-ACP, giving a C7-ACP product (Supplementary Fig. 10a-b). In fact, our result of isothermal titration calorimetry (ITC) verified efficient binding of BioZ to glutaryl-ACP with the stoichiometry of N = 0.978 ± 0.028 and Kd = 6.167 ± 0.068 μM (Supplementary Fig. 11). This might be
Fig. 3 A role of BioZ in the FAS II-involving biotin biosynthesis. a In vitro biosynthesis of pimeloyl-ACP and/or its precursors. Using the FAS II system, BioZ catalyzes the synthesis reaction of pimeloyl-ACP from malonyl-ACP and glutaryl-CoA (ACP). The reaction mixture was separated with conformation-sensitive urea polyacrylamide gel electrophoresis (PAGE). A representative result is given from three trials. Of note: 17.5% PAGE (pH 9.5) containing 0.5 M urea was used here. The two controls (C4-ACP and C6-ACP) served as standards/markers for this conformation-sensitive urea gel. C7-ACP denotes four species of ACP attached with an acyl seven-carbon fatty acyl chain, namely, 5-keto-pimeloyl-ACP, 5-hydroxyl-pimeloyl-ACP, enoyl-pimeloyl-ACP, and pimeloyl-ACP. b MS/MS identification of 5-keto-pimeloyl-ACP, a primary product from the BioZ reaction coupled with FAS II, using glutaryl-CoA and malonyl-ACP as substrates. The use of MS/MS allowed us to detect the presence of four C7-ACP species in the above reaction system. As for an initial product of BioZ reaction, 5-keto-pimeloyl ACP, a 15-residue peptide fragment of interest is given. The C7-fatty acyl modification with high reliability is localized on the conserved Serine 36 of ACP. The two peaks of peptide fragments indicated with pink arrows were used to determine C7 acyl modification. The resultant mass was 498.1533, which is close to the theoretical mass (498.1430) of Ppan-linked keto-pimeloyl moiety. Of note, it might be not as stable as pimelic acid. c Cartoon illustration of the pimeloyl-ACP structure. This was generated from the complex structure of methyl-pimeloyl-ACP (ASU) (Supplementary Fig. 16a), of which the re...d A scheme for BioZ bypassing the canonical early steps of biotin synthesis. Unlike the paradigm “BioC-BioH” mechanism of biotin synthesis (above the dashed line), the BioZ reaction bypasses the earlier steps of “BioC-BioH” in biotin biosynthesis (below the dashed line). Recently, glutaryl-CoA is determined to physiologically arise from lysine catabolism in Agrobacterium species47. Designations: C4-ACP butanoyl-ACP, C6-ACP hexanoyl-ACP, Mal-ACP malonyl-ACP, Glu-ACP glutaryl-ACP, Pim-ACP pimeloyl-ACP;
triple decked sandwich that is formed by nine β-sheets and ten α-helices. In brief, the β-sheets of β1 and β4–β7 are sandwiched between the helices α3–α5 and α6, together with helices of α13–α14 to form the first layer, while the sheets of β8–β11 are sandwiched between the helices of α9–α12 and α6, together with the helix of α13–α14 to form the other layer of the “sandwich” (Supplementary Fig. 15a-b). Interestingly, a putative catalytic triad (C115, H255, and N285, Supplementary Figs. 15 and 19c-d) is centered in the core domain of AtBioZ (Figs. 4d, 5d). Consistent with the evolutionary placement seen in our phylogenetic studies (Fig. 2a), the structure of AtBioZ (PDB: 6KUE) is similar to FabH (PDB: 1HN9), whereas FabB (PDB: 1G5X) structurally is more related to that of FabF (PDB: 2GFW) (Fig. 4a, d). Even though the core domain and the active site are highly conserved (Figs. 4 and 5a, d), the lid domains differ dramatically across the four different members of the KAS family enzymes (Fig. 4a, d). In brief, the lid domain of FabB is mainly composed of five α-helices and two β-sheets, and the counterpart of FabF is formed by four α-helices and two β-sheets (Fig. 4a, b). However, the lid domain of FabH/BioZ is composed of two parts: one part is the long loop between β8 and α9 from the first layer of the “sandwich”, and the other part contains two regions (β1–α3 and β6–β7) on the second layer of the “sandwich” (Fig. 4c, d and Supplementary Fig. 15). Taken together, our results provide possible explanation for the diversity in substrate specificity/recognition amongst the KAS family enzymes.

Fig. 4 Structural comparison of four FAS-type enzymes. a Ribbon representation of the E. coli FabB structure (PDB: 1G5X). b Ribbon structure of the E. coli FabF (PDB: 2GFW). c Ribbon representation of the E. coli FabH structure (PDB: 1HN9). d Ribbon illustration of the A. tumefaciens BioZ structure (PDB: 6KUE). The lid domains in EcFabB (a), EcFabF (b), and EcFabH (c) are separately colored in green, cyan, and yellow, respectively. Whereas, the core domains are colored in silver-gray. As for AtBioZ, the lid domain is shown in red, and its core domain appears in blue (d). The three catalytic triad residues are indicated with pink letters.
Catalytic triad of BioZ. FAS enzymes possess an evolutionarily-conserved catalytic center. The catalytic triad (C163, H298, and H333) of EcFabB is almost structurally identical to that of EcFabF (C164, H304, and H341) (Fig. 5a, b). In contrast, the active site residues of AtBioZ correspond to C115, H255, and N285, which are highly similar to those of EcFabH (C112, H244, and N274) (Fig. 5c, d). To test the function of the putative catalytic center residues of AtBioZ, we performed site-directed mutagenesis (C115A, H255A, and N285A) followed by complementation assays on the nonpermissive growth condition without biotin. Experimentally, the ΔbioH biotin-auxotrophic strain was transformed with the wild-type or mutant bioZ plasmid and assayed for growth in liquid M9 minimal media lacking biotin (Fig. 5e). As predicted, the ΔbioH strain grew robustly when complemented with the wild-type bioZ plasmid and assayed for growth in liquid M9 minimal media lacking biotin (Fig. 5e). As predicted, the ΔbioH strain grew robustly when complemented with the wild-type bioZ plasmid and assayed for growth in liquid M9 minimal media lacking biotin (Fig. 5e).

Interaction between malonyl-ACP and BioZ. Since that BioZ was biochemically confirmed active on malonyl-ACP and glutaryl-CoA (ACP), we then asked the question how this BioZ enzyme interacts with the above two substrates. Although that our earlier efforts failed to co-crystallize AtBioZ complexed with malonyl-ACP or glutaryl-ACP (CoA), the recent availability of crystal structure of FabB crosslinked with ACP 16, allowed us to probe the interplay between malonyl-ACP (or its ACP moiety) and BioZ. Thus, we used the structure of FabB-ACP (2:1) complex to guide the building of a BioZ-ACP model (Fig. 6a, d).

Fig. 5 Structural and functional analyses of the BioZ catalytic triad. a Structural snapshot of the catalytic triad (C163, H298, and H333) of EcFabB. b Structural analysis of the catalytic triad (C164, H304, and H341) of EcFabF. c Structural presentation of the catalytic triad (C112, H244, and N274) of EcFabH. d Structural illustration of the catalytic triad (C115, H255, and N285) of AtBioZ. The figures of structures were generated using the PyMol software. e Functional assays of BioZ and its catalytic triad mutants using the ΔbioH biotin-auxotrophic strain. The tested strains were grown on nonpermissive M9 minimal media lacking biotin. Growth curves were generated from three independent experiments and displayed in an average ± standard deviation (SD).
Unlike that the FabB substrate-binding pocket whose entrance is close to the dimeric interface (Fig. 6a, c), the entrance of the AtBioZ substrate-binding pocket is relatively far away from the dimeric interface (Fig. 6b, d). In the case of FabB-ACP (Supplementary Fig. 19a-b), the ACP protein mainly interacts with one monomer while the acyl group on ACP inserts into the substrate-binding pocket of another monomer (Fig. 6a, c). In other words, both monomers of the FabB dimer cooperate for ACP protein and acyl group binding. The ClusPro-based molecular docking identified a putative interface between BioZ with its cognate ACP (Supplementary Figs. 18a and 19c, d). In contrast to its relatively weak affinity with ACP alone having the stoichiometry (N = 1.00 ± 0.17, Kd = 0.33 ± 0.25 mM, in Supplementary Fig. 18b), the mutant enzyme of BioZ(C115A) was found to bind modestly malonyl-ACP in our ITC experiments (N = 0.996 ± 0.096, Kd = 24.133 ± 0.602 μM, in Fig. 6e). Therefore, we proposed a “one-to-one” model for BioZ-ACP interaction in which an ACP protein can bind to only one BioZ monomer and the malonyl group of ACP interacts with the same BioZ monomer (Fig. 6b, d). Presumably, FabB-ACP interface mainly relies on an electrostatic interaction between negatively-charged residues on ACP and positively-charged residues on FabB (Fig. 7a). On the bottom of ACP helix II, Asp35, and Asp38 interact with Arg62, Lys63, and Arg66 on FabB, while Glu47 at the top of helix II interacts with Arg124 and Lys127 on FabB (Fig. 7b). Similarly, the surface around the entrance of BioZ pocket is enriched with...
Fig. 7 Structural insights into the recognition of malonyl-ACP by the BioZ enzyme. a Structural display of the ACP moiety of acyl-ACP substrate bound to the paradigm enzyme of FAS I, the E. coli FabB. b An enlarged view of the ACP-interacting interface of FabB that is constituted by six positively-charged residues. The image is generated through a counter-clockwise 90° rotation of the rectangle-dashed lined region (a). c Binding model of glutaryl-ACP to the positively-charged surface of AtBioZ. ACP is given in ribbon structure colored gray, which comprises four α-helices (helix-1 to helix-4). The interface of FabB (a) [and/or BioZ (c)] interacting with ACP group is illustrated in the surface electrostatic structure. The blue denotes positive charge, whereas the red refers to negative charge. d Visualization of the putative ACP-binding domain of BioZ. Presumably, it contains four basic residues, namely R39, R153, R221, and R260. This image is given through the counter-clockwise 90° rotation of the square-dashed lined region (c). e Growth curve-based assay to probe the role of the putative four basic residues-comprising, ACP-binding interface in the BioZ function. The strains used here were listed in Supplementary Table 1, and the growth curve was plotted as described in Fig. 5e. A representative result was given.
positively-charged amino acids, namely Arg39, Arg153, Arg221, and Arg260 (Fig. 7c, d). Among them, R39 and R153 was also predicted to interact with the CoA cargo. It is plausible that these residues play certain roles in the interaction of BioZ with the negatively-charged residues on ACP protein. 

To test this hypothesis, we generated a number of BioZ mutants (Supplementary Table 1) and examined them in vivo. These mutants of BioZ are divided into 4 groups: (i) four single mutants (R39A, R153A, R221A, and R260A), (ii) six sets of double mutants (R39A/R153A, R39A/R221A, R39A/R260A, R153A/R221A, R153A/R260A, and R221A/R260A), (iii) three triple mutants (R39A/R153A/R221A, R39A/R153A/R260A, R39A/R221A/R260A, and R153A/R221A/R260A), and (iv) a quadruple mutant (R39A/R153A/R221A/R260A). Next, genetic complementation of the ΔbioZ mutant was performed using pBAD24-borne bioZ mutants (Fig. 7e). As a result, (i) only one of four single bioZ mutants (i.e., R260A) lost full activity as illustrated in the phenotypic growth curves of ΔbioZ (ii) nearly all the six double mutants of bioZ exhibited poor activities; (iii) all the four triple mutants of bioZ were functionally impaired; and finally, (iv) the quadruple mutant of bioZ (R39A/R153A/R221A/R260A) was inactive (Fig. 7e). Together with our structural data, these in vivo results defined a functional interface between BioZ and malonyl-ACP, which is important in bypassing the canonical earlier steps of biotin synthesis.

Recognition of primer glutaryl-CoA by BioZ. The biochemical data indicated that BioZ exploits glutaryl-CoA and glutaryl-ACP both as the primer substrates in the in vitro reconstituted FAS reaction system (Fig. 3 and Supplementary Fig. 10). In contrast to those of malonyl-ACP (Fig. 6c) and holo-ACP alone (Supplementary Fig. 18), glutaryl-ACP exhibited more tight binding the AtBioZ(C115A) mutant enzyme, as determined by ITC experiments with the stoichiometry of N = 0.978 ± 0.028, and KD = 6.167 ± 0.068 μM (Supplementary Fig. 11). It seemed likely that BioZ enzyme can adjust (perhaps expand) the configuration of primer-loading tunnel to accommodate primer substrate of longer chain acyl-ACP (CoA). To address this issue, we thus applied AutoDock Vina-based molecular docking. As expected, the ACP-removed glutaryl moiety (i.e., glutarate) we selected was well docked into the AtBioZ enzyme (Supplementary Fig. 16a-b). Structural superposition revealed that this pocket occupied by glutarate is conserved amongst BioZ and its ancestor FabH enzymes (Supplementary Fig. 16c). This indicated that the primer-loading tunnel on BioZ is an evolutionary relic. Very recently, Hu and Cronan reported genetic evidence for glutaryl-CoA originating from bacterial lysine catabolism, and argued the feasibility of glutaryl-ACP as a physiological primer of BioZ. Because that both CoA and ACP are acidic molecules carrying a flexible Ppan moiety, it is not unusual that CoA and ACP mimic one another as noncognate substrates in the context of fatty acid metabolism. If so, this might explain why the above two C5-thioesters (C5-CoA and C5-ACP) can be recognized by BioZ in vitro and displayed comparable enzymatic affinity.

Since glutaryl-CoA is a most-likely intracellular primer/ligand, we then ask the question how the substrate-tunnel is tolerant with the free charged ω-carboxyl group. Subsequently, molecular docking of an intact glutaryl-CoA into AtBioZ was carried out (Fig. 8a). This allowed us to visualize a fine substrate-loading tunnel accessible to glutaryl-CoA (Fig. 8b). Not surprisingly, our ITC analyses demonstrated an efficient binding of glutaryl-CoA to BioZ in which the stoichiometry appears as follows: N = 0.80 ± 0.01, and KD = 2.64 ± 0.23 μM (Fig. 8c). This was quite similar to the scenario seen with glutaryl-ACP (Supplementary Fig. 11). Indeed, it in turn verified the reliability in molecular docking of BioZ/glutaryl-CoA. Structural analysis of the BioZ tunnel surface localized three residues of interest that might interact/neutralize the free carboxyl group of glutaryl-CoA substrate. They included Serine 84 (S84), Arginine 147 (R147), and Serine 287 (S287), respectively (Fig. 9a). Moreover, sequence alignments showed that all the three putative sites are extremely conserved across different BioZ homologs throughout α-proteobacteria (Fig. 9b). Thereafter, we applied alanine substitution to generate three single mutants (S84A, R147A, and S287A), and tested the in vivo roles in bypassing the need of BioH in demethylation of methyl-pimeloyl-ACP in E. coli. Although the R147A mutant of bioZ retained partial activity, the other two mutants (S84A and S287A) cannot allow the ΔbioH bioin-auxotrophic strain to well appear on the nonpermissive condition, bioin-deficient M9 minimal agar plates (Fig. 9c). A similar scenario was observed with growth curves of the aforementioned strains in the liquid M9 media lacking biotin (Fig. 9d). In addition, our structure-guided functional assays demonstrated that four additional residues (E152, T216, M217, and N258) play roles in the transient crosstalk/stabilization of BioZ with the CoA cargo. Taken together, these results constituted functional definition of the primer substrate-loading tunnel within the AtBioZ enzyme.

Discussion
Along with recent findings by other two research groups,20,22,23,34 the data we reported here contributes important biochemical and structural information of de novo biotin synthesis, which proceeds using three distinct pathways: (i) BioC-BioH pathway20-22, (ii) BioI-BioW pathway33,34 or (iii) noncanonical BioZ pathway (Fig. 3). Among them, the BioC-BioH pathway (and/or its equivalent, like BioJ27,42) is the prevalent route. In contrast, the BioZ pathway seems to be restricted to the phylum of α-proteobacteria (Supplementary Figs. 1-4). The exclusive occupation of BioZ in certain α-proteobacterial species, like A. tumefaciens, can be in part explained by its origin and evolution of FabH in the close relatives of Rhizobiales (Fig. 2). Probably, a long-term selection and gradual evolution domesticated this FabH-like enzyme, BioZ to do an “extra and/or redundant job”, providing a substrate for biotin synthesis (Fig. 3 and Supplementary Fig. 9). To less extents, the three residue-forming catalytic triad (C115, H255, and N285) of BioZ is conserved amongst BioZ and its ancestor FabH enzymes (Supplementary Fig. 16c). This indicated that the primer-loading tunnel on BioZ is an evolutionary relic. Very recently, Hu and Cronan reported genetic evidence for glutaryl-CoA originating from bacterial lysine catabolism, and argued the feasibility of glutaryl-ACP as a physiological primer of BioZ. Because that both CoA and ACP are acidic molecules carrying a flexible Ppan moiety, it is not unusual that CoA and ACP mimic one another as noncognate substrates in the context of fatty acid metabolism. If so, this might explain why the above two C5-thioesters (C5-CoA and C5-ACP) can be recognized by BioZ in vitro and displayed comparable enzymatic affinity.

Since glutaryl-CoA is a most-likely intracellular primer/ligand, we then ask the question how the substrate-tunnel is tolerant with the free charged ω-carboxyl group. Subsequently, molecular docking of an intact glutaryl-CoA into AtBioZ was carried out (Fig. 8a). This allowed us to visualize a fine substrate-loading tunnel accessible to glutaryl-CoA (Fig. 8b). Not surprisingly, our ITC analyses demonstrated an efficient binding of glutaryl-CoA to BioZ in which the stoichiometry appears as follows: N = 0.80 ± 0.01, and KD = 2.64 ± 0.23 μM (Fig. 8c). This was quite similar to the scenario seen with glutaryl-ACP (Supplementary Fig. 11). Indeed, it in turn verified the reliability in molecular docking of BioZ/glutaryl-CoA. Structural analysis of the BioZ tunnel surface localized three residues of interest that might interact/neutralize the free carboxyl group of glutaryl-CoA substrate. They included Serine 84 (S84), Arginine 147 (R147), and Serine 287 (S287), respectively (Fig. 9a). Moreover, sequence alignments showed that all the three putative sites are extremely conserved across different BioZ homologs throughout α-proteobacteria (Fig. 9b). Thereafter, we applied alanine substitution to generate three single mutants (S84A, R147A, and S287A), and tested the in vivo roles in bypassing the need of BioH in demethylation of methyl-pimeloyl-ACP in E. coli. Although the R147A mutant of bioZ retained partial activity, the other two mutants (S84A and S287A) cannot allow the ΔbioH bioin-auxotrophic strain to well appear on the nonpermissive condition, bioin-deficient M9 minimal agar plates (Fig. 9c). A similar scenario was observed with growth curves of the aforementioned strains in the liquid M9 media lacking biotin (Fig. 9d). In addition, our structure-guided functional assays demonstrated that four additional residues (E152, T216, M217, and N258) play roles in the transient crosstalk/stabilization of BioZ with the CoA cargo. Taken together, these results constituted functional definition of the primer substrate-loading tunnel within the AtBioZ enzyme.
Assume that dimeric BioZ enzyme successively binds to the aforementioned two substrates via surface electrostatic interaction at the same ACP-binding sites (Figs. 5, 6). Given that malonyl-ACP plays a critical role in the elongation steps catalyzed by β-ketoacyl-ACP synthases, we therefore formulate a working model for BioZ action, in which BioZ likely binds to glutaryl-CoA (ACP), then malonyl-ACP (Supplementary Fig. 20). In this model, a glutaryl-loaded CoA (ACP) first associates with BioZ, the active site of which, cysteine (Cys115) attacks the glutaryl moiety and releases the CoA (ACP) carrier. Then, the glutaryl-BioZ intermediate is ligated with a malonyl-ACP, leading to the addition of the acyl chain of glutaryl moiety with two carbon units (from C5 to C7). Finally, the 5-keto-pimeloyl ACP dissociates from the dimeric BioZ enzyme, which is then modified by FAS II system, prior to entry into the latter steps of biotin synthesis (Supplementary Fig. 9). That is the reason why the bioZ gene bypasses the genetic requirement of “bioC-bioH” in E. coli on the nonpermissive growth condition (Figs. 5e, f and 9c, d). It is unusual, but not without any precedent, that the activity of BioZ originates from its noncognate substrate preference/speciﬁcity in the context of FAS II fatty acid synthesis. Therefore, BioZ essentially serves as a noncanonical player within the family of β-ketoacyl-ACP synthase III, the enzymatic action of which satisfies the physiological demand of biotin in a number of bioC- and bioW-lacking bacteria (Supplementary Fig. 1). Structural comparison of different β-ketoacyl-ACP synthase III suggests that the gain of an atypical FabH-like activity of BioZ might be due to its distinct lid domain rather than the conserved catalytic domain (Fig. 3). This hints that an evolutionary advantage occurs in engineering a redundant (and/or duplicated) gene, but not creating an additional gene for a previously-unrecognized metabolic ability.

Because that biotin is recognized as a restriction (and/or nutritional) virulence factor42,48–50, the action of BioZ is presumably to participate into successful infection of the plant pathogen A. tumefaciens39 and the human pathogen B. melitensis36. Structure-guided functional dissection of both catalytic triad (Fig. 4) and substrate-loading tunnel (Fig. 6) might render BioZ as an antivirulence drug target, which can be utilized into the screen of lead compounds and design of inhibitors of small molecules. Given that A. tumefaciens, the bioZ-bearing microbe, exhibits the potential of profligate synthesis of biotin (perhaps its precursor DTB) in the excess of its physiological demand39, it is of much interest to ask the three questions bellowed: (i) if or not the excessive secretion of biotin/DTB is an ecologically selective outcome, benefiting those biotin-auxotrophic bacterial species inhabited in the same niche;
(ii) if *A. tumefaciens* can be developed into a ‘green chemistry’ tool for biotin production; and (iii) whether or not biotin production is greatly elevated in the BioZ-engineered *E. coli* strains? Along with the genetic & enzymatic study of Hu and Cronan\(^4\), our report extends biochemical and structural insights into molecular mechanism for bacterial BioZ biotin synthesis. To the best of our knowledge, the BioZ pathway represents a third machinery route for the formation of the biotin precursor pimeloyl-ACP, and poses extensively promising implications in the agricultural, pharmaceutical, and biomedical fields.
Phylogenetic analysis of BioZ. The genetic context of bioZ-like loci and genes coding for FabH, FabB, and FabF in bacterial genomes were comparatively investigated in NCBI RefSeq archived 334 a-proteobacteria genomes by using BLASTp-based searches. BioZ homologues were predicted by using BLASTp and BLASTn, respectively, against protein and nucleotide sequences of biotin synthesis gene clusters (bioBDFAZ + bioMN) from Agrobacterium tumefaciens str. C58 and Brucella melitensis bv. 1, str. 16 M75. Then, the functional relatedness of neighboring genes was carefully examined and correctly placed in the context of lipid metabolism37,42.54,55. In light of respective prototypes of FabH53, FabB53, and FabF54,55, the homology search identified 106 homologs of KAS enzymes, consisting of 50 FabH, 41 FabF, and 15 FabB. Finally, the phylogenetic relationships between BioZ and its homologs (FabH, FabB, and FabF) were calculated with the software of MEGAS7.

Genetic manipulations. To address the function of the BioZ enzyme, three bioZ homologs of a-proteobacteria (A. tumefaciens, B. melitensis, and R. rhizotorium) were ampliﬁed by PCR and/or synthesized in vitro. The genes were cloned into an arabinose-inducible pBAD24 expression vector, giving recombinant plasmids such as pBAD24: atbioZ (Supplementary Table 1). To probe whether BioZ bypasses the cognate earlier steps of biotin synthesis in E. coli, the recombinant plasmids were transformed into a single mutant ΔaioC (or ΔaioH) or into the double mutant ΔaioC/ΔaioH (Supplementary Table 1). Similarly, all the three versions of bioZ were cloned into an isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible expression pET28a vector, to produce a soluble form of BioZ proteins suitable for in vitro enzymatic assays and crystallization screens. Among them, only pET28: atbioZ from A. tumefaciens (Supplementary Table 1), was found to produce soluble BioZ protein. Additionally, the fabH1 of A. tumefaciens was also cloned into pET28 vector to prepare its soluble protein. To generate direct site-directed point mutations of bioZ, two different techniques were used accordingly: (i) primers-based overlapping PCR (Supplementary Table 2), and (ii) the Mut Express II fast mutagenesis kit v2 (Vazyme Biotech Co., Ltd.). All plasmids generated in this study are listed in Supplementary Table 2.

protein expression, purification, and identification. Isolation of recombinant protein from E. coli BL21 with pET28: atbioZ was used to produce the BioZ protein (Supplementary Table 1). Protein expression and purification (small batches) was ﬁrst optimized with GE buffer (25 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 2 mM DTT), followed by preparations with BJADN54, β-ketothiolase (ΔK172E) with little change57,42, the bacterial lysate containing six histidine-tagged AtBioZ protein was prepared with GF buffer, and then subjected to the puriﬁcation procedure. All the three proteins were puriﬁed with pepsin (rather than the routine trypsin). In brief, after 16 h of induction, at least 95% of the target protein was detected from the gel, and digested with pepsin (rather than the routine trypsin). In

proteolytic system of fatty acid synthesis in vitro was reconstituted as described earlier by Zhu and Cronan37,52. The recombinate E. coli FabH was prepared routinely58,63. Prior to the reaction of fatty acid synthesis in vitro, the recombinant FabH was expressed and puriﬁed with little change57,42, the bacterial lysate containing six histidine-tagged AtBioZ protein was prepared with GF buffer, and then subjected to the puriﬁcation procedure. All the three proteins were puriﬁed with pepsin (rather than the routine trypsin). In brief, after 16 h of induction, at least 95% of the target protein was detected from the gel, and digested with pepsin (rather than the routine trypsin). In

protein expression, purification, and identiﬁcation. Isolation of recombinant protein from E. coli BL21 with pET28: atbioZ was used to produce the BioZ protein (Supplementary Table 1). Protein expression and purification (small batches) was ﬁrst optimized with GE buffer (25 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 2 mM DTT), followed by preparations with BJADN54, β-ketothiolase (ΔK172E) with little change57,42, the bacterial lysate containing six histidine-tagged AtBioZ protein was prepared with GF buffer, and then subjected to the puriﬁcation procedure. All the three proteins were puriﬁed with pepsin (rather than the routine trypsin). In brief, after 16 h of induction, at least 95% of the target protein was detected from the gel, and digested with pepsin (rather than the routine trypsin). In
Molecular docking between BioZ and substrates. The protein three-dimensional structures were obtained from RCSB PDB database49. Structure superposition was performed using MatchMaker tool of UCSF Chimera software50. The structure images were generated using UCSF Chimera and PyMOL software (https://pymol.org). BioZ possesses similar fold and active site structure as FabB. We built a schematic model of BioZ-ACP complex based on CluPro server-aided molecular docking52. Given that (i) three negatively charged amino acids (Asp55, Asp58, and Glu47) have roles in the interaction between FabB and ACP, and (ii) FabB accesses its substrate by interacting with ACP protein46, four arginines (Arg9, Arg53, Arg221, and Arg260) of BioZ were located in the proximity of its substrate-loading entrance.

To further dissect binding model of the primer glutaryl-CoA (or glutarate alone) and BioZ, molecular docking was applied using AutoDock Vina (version 1.1.2)53. Apart from the high-resolution (1.992 Å) 3D structure of BioZ (PDB ID: 6KUE), the glutaryl-CoA structure was extracted from a of glutaryl-CoA (PDB ID: 6KUE) as the searching model66,67. Model building and crystallographic refinement were conducted with COOT and PHENIX38. The final structure of AtBioZ was solved at 1.99 Å (Table1) and deposited into the Protein Data Bank (PDB) with the accession entry: 6KUE. The images for interpreting the NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-021-22360-4 www.nature.com/naturecommunications

Received: 22 November 2019; Accepted: 12 March 2021; Published online: 06 April 2021

References
1. Beckett, D. Biotin sensing: universal influence of biotin status on transcription. Dev. Biol. 41, 463–464 (2007).
2. Rodionov, D. A., Mironov, A. A. & Gelfand, M. S. Conservation of the biotin regulon and the BirA regulatory signal in Eubacteria and Archaea. Genome Res. 12, 1507–1516 (2002).
3. Beckett, D. Biotin sensing at the molecular level. J. Nutr. 139, 167–170 (2009).
4. Cronan, J. E. Jr The biotinyl domain of Euarchonia coli acetyl-CoA carboxylase. Evidence that the “thumb” structure is essential and that the domain functions as a dimer. J. Biol. Chem. 276, 37355–37364 (2001).
5. Huang, C. S., Ge, P., Zhou, Z. H. & Tong, L. An unanticipated architecture of the 750-kDa α6β6 holoenzyme of 3-methylcrotonyl-CoA carboxylase. Nature 481, 219–223 (2011).
6. Huang, C. S. et al. Crystal structure of the α6β6 holoenzyme of propionyl-coenzyme A carboxylase. Nature 466, 1001–1005 (2010).
7. Lazar, N. et al. Control of biotin biosynthesis in mycobacteria by a pyruvate carboxylase dependent metabolic signal. Mol. Microbiol. 106, 1018–1031 (2017).
8. Pacheco-Alvarez, D., Solozzano-Vargas, R. S. & Del Rio, A. L. Biotin in metabolism and its relationship to human disease. Arch. Med. Res. 33, 439–447 (2002).
9. Cronan, J. E. Biotin and lipoid acid: synthesis, attachment, and regulation. EcoSal Plus, 6 https://doi.org/10.1128/ecosalplus.EPS-0001-2012 (2014).
10. Leon-Del-Rio, A. Biotin in metabolism, gene expression, and human disease. J. Inherit. Metab. Dis. 42, 1811–1821 (2019).
11. du Vigneaud, V. The structure of biotin. Science 96, 455–461 (1942).
12. du Vigneaud, D. U. V., Melville, D. B., Gyorgy, P. & Rose, C. S. On the identity of vitamin H with biotin. Science 92, 62–63 (1940).
13. Lin, S. & Cronan, J. E. Closing in on complete pathways of biotin biosynthesis. Mol. BioSyst. 7, 1811–1821 (2011).
14. Chapman Smith, A. & Cronan, J. E. Jr Molecular biology of biotin attachment to proteins. J. Nutr. 129, 4775–4845 (1999).
15. Tong, L. Structure and function of biotin-dependent carboxylases. Cell Mol. Life Sci. 70, 863–891 (2013).
16. Cronan, J. E. Advances in synthesis of biotin and assembly of lipoyl acid. Curr. Opin. Chem. Biol. 47, 60–66 (2018).
17. Manandhar, M. & Cronan, J. E. Panellic acid, the first precursor of the Bacillus subtilis biotin synthesis pathway, exists as the free acid and is assembled by fatty acid synthesis. Mol. Microbiol. 104, 595–607 (2017).
18. Manandhar, M. & Cronan, J. E. A canonical biotin synthesis enzyme, 8-amino-7-oxononanoate synthase (BioF), utilizes different acyl chain donors in Bacillus subtilis and Escherichia coli. J. Bacteriol. 190, e02084-17 (2018).
19. Rodionov, D. A., Dubchak, I., Arkpin, A., Alm, E. & Gelfand, M. S. Reconstruction of regulatory and metabolic pathways in metal-reducing delta-proteobacteria. Genome Biol. 5, R90 (2004).
20. Lin, S., Hanson, R. E. & Cronan, J. E. Biotin synthesis begins by hijacking the fatty acid synthetic pathway. Nat. Chem. Biol. 6, 682–688 (2010).
21. Lin, S. & Cronan, J. E. The C2-acyl synthase catalyzes methyl esterification of malonyl-acyl carrier protein, an essential step in biotin synthesis. J. Biol. Chem. 287, 37010–37020 (2012).
22. Agarwal, V., Lin, S., Løkk, T., Nair, S. K. & Cronan, J. E. Structure of the enzyme-acyl carrier protein (ACP) substrate gatekeeper complex required for biotin synthesis. Proc. Natl Acad. Sci. USA 109, 17406–17411 (2012).
23. Pshapire, M. M., Chakravorty, V. & Cronan, J. E. Remarkable diversity in the enzymes catalyzing the last step in synthesis of the pimelate moiety of biotin. PLoS ONE 7, e49440 (2012).
24. Bi, H., Zha, L., Jia, J. & Cronan, J. E. A biotin biosynthesis gene restricted to Helicobacter. Sci. Rep. 6, 21162 (2016).
25. Feng, Y. et al. A Francisella virulence factor catalyzes an essential reaction of biotin synthesis. Mol. Microbiol. 91, 300–314 (2014).
26. Shi, J., Cao, X., Chen, Y., Cronan, J. E. & Guo, Z. An atypical αβ-hydrolase fold revealed in the crystal structure of pimeloyl-acyl carrier protein methyl esterase BioG from Haemophilus influenzae. Biochemistry 55, 6705–6717 (2016).
27. Wei, W. et al. Molecular basis of BioG, a unique gatekeeper in bacterial biotin synthesis. Science 19, 790 (2019).
28. Manandhar, M. & Cronan, J. E. Proofofreading of noncognate acyl adenylates by an acyl-coenzyme a ligase. Chem. Biol. 20, 1441–1446 (2013).
29. Stok, J. E. & De Voss, J. Expression, purification, and characterization of Biol-a carbon-carbon bond cleaving cycytomere P450 involved in biotin biosynthesis in Bacillus subtilis. Arch. Biochem. Biophys. 384, 351–360 (2000).
30. Cryle, M. J. & Schlichting, L. Structural insights from a P450 carboxylic acid methyl complex reveal how specificity is achieved in the P450(BioI) ACP complex. Proc. Natl Acad. Sci. USA 105, 15696–15701 (2008).
31. Cryle, M. J. & De Voss, J. J. Carbon-carbon bond cleavage by cycytomere P450(BioI)(CYP107H1). Chem. Commun. 1, 86–87 (2004).
32. Ploux, O., Soulauré, P., Marquet, A., Gleöcker, R. & Lemoine, Y. Investigation of the first step of biotin biosynthesis in Bacillus subtilis. Purification and
characterization of the pimeloyl-CoA synthase, and uptake of pimelate. Biochem. J. 287, 685–690 (1992).

33. Wang, M. et al. Using the pimeloyl-CoA synthetase adenylation fold to synthesize fatty acid thioesters. Nat. Chem. Biol. 13, 660–667 (2017).

34. Estrada, P. et al. The pimeloy-CoA synthetase BioW defines a new fold for adenylate-forming enzymes. Nat. Chem. Biol. 13, 668–674 (2017).

35. Croman, J. E. & Lin, S., Synthesis of the alpha,-omega-dicarboxylic acid precursor of biotin by the canonical fatty acid biosynthetic pathway. Curr. Opin. Chem. Biol. 15, 407–413 (2011).

36. Feng, Y., Xu, J., Zhang, H., Chen, Z. & Srinivas, S. Brucella BioA regulator defines a complex regulatory mechanism for bacterial biotin metabolism. J. Bacteriol. 195, 3451–3467 (2013).

37. Entcheva, P., Phillips, D. A. & Streit, W. R. Functional analysis of Serratia marcescens fab genes involved in biotin synthesis and transport. Appl. Environ. Microbiol. 68, 2843–2848 (2002).

38. Sullivan, J. T., Brown, S. D., Yocum, R. R. & Ronson, C. W. The bio operon on the acquired symbiosis island of Mesorhizobium sp. strain R7A includes a novel gene involved in pimeloy-CoA synthesis. Microbiology 147, 1315–1322 (2001).

39. Feng, Y., Zhang, H. & Croman, J. E. Folligiotin biotin synthesis in alpha-proteobacteria—a developing or degenerating regulatory system? Mol. Microbiol. 88, 77–92 (2013).

40. Kaneko, T. et al. Complete genome structure of the nitrogen-fixing bacterial bacterium Mesorhizobium loti. DNA Res. 7, 331–338 (2000).

41. Kaneko, T. et al. Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti (supplement). DNA Res. 7, 381–406 (2000).

42. Feng, Y. et al. A Francisella virulence factor catalyses an essential reaction of biotin synthesis. Mol. Microbiol. 91, 300–314 (2014).

43. Qiu, X. et al. Crystal structure of beta-ketoacyl-ACP carrier protein synthase III. A key condensing enzyme in bacterial fatty acid biosynthesis. J. Biol. Chem. 274, 36465–36471 (1999).

44. Zhang, Y. M. et al. Identification and analysis of the acyl carrier protein (ACP) docking site on beta-ketoacyl-ACP synthase III. J. Biol. Chem. 276, 8231–8238 (2001).

45. Wang, J. et al. Platsensimycin is a selective FabF inhibitor with potent antibiotic properties. Nature 441, 358–361 (2006).

46. Milligan, J. C. et al. Molecular basis for interactions between an acyl carrier protein and a ketosynthase. Nat. Chem. Biol. 15, 669–671 (2019).

47. Hu, Y. & Croman, J. E. Alpha-proteobacteria biosynthesize biotin precursor pimeloyl-ACP using BioZ 3-ketoacyl-ACP synthase and yiscine catabolism. Nat. Commun. 11, 5598 (2020).

48. Feng, Y. et al. The atypical occurrence of two biotin protein ligases in Francisella novicida is due to distinct roles in virulence and biotin metabolism. mBio 6, e00591 (2015).

49. Napier, B. A. et al. Link between intraphagosomal bacterial and rapid phagosomal escape in Francisella. Proc. Natl Acad. Sci. USA 109, 18084–18089 (2012).

50. Woong Park, S. et al. Evaluating the sensitivity of Mycobacterium tuberculosis to biotin deprivation using regulated gene expression. PLoS Pathog. 7, e1002264 (2011).

51. Feng, Y., Kumar, R., Ravcheev, D. A. & Zhang, H. Paracoccus denitrificans possesses two BioB homologs having a role in regulation of biotin metabolism. Microbiol. Open 4, 644–659 (2015).

52. Wei, W. et al. Crystal structure and acetylation of BioQ suggests a novel regulatory switch for biotin biosynthesis in Mycobacterium smegmatis. Mol. Microbiol. 109, e62–e62 (2018).

53. Lai, C. Y. & Croman, J. E. Beta-ketoacyl-ACP carrier protein synthase III (FabH) is essential for bacterial fatty acid synthesis. J. Biol. Chem. 278, 51494–51503 (2003).

54. Wang, H. & Croman, J. E., Functional replacement of the FabA and FabB proteins of Escherichia coli fatty acid synthase by Enterococcus faecalis FabZ and FabB homologues. J. Biol. Chem. 279, 34489–34495 (2004).

55. Feng, Y. & Croman, J. E. Escherichia coli unsaturated fatty acid synthesis: complex transcription of the fabA gene and in vivo identification of the essential reaction catalyzed by FabB. J. Biol. Chem. 284, 29526–29535 (2009).

56. Kumar, S., Stecher, G. & Tamura, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874 (2016).

57. Zhu, L. et al. The two functional enoyl-ACP carrier protein reductases of Enterococcus faecalis do not mediate triclosan resistance. mBio 4, e00613–e00613 (2013).

58. Flugel, R. S., Wangbo, Y., Lamblot, R. H., Croman, J. E. & Walsh, C. T. Holo-(acyl carrier protein) synthase and phosphopantetheinyl transfer in Escherichia coli. J. Biol. Chem. 275, 959–968 (2000).

59. Jiang, Y., Chan, C. H. & Croman, J. E. The soluble acyl-ACP carrier protein synthetase of Vibrio harveyi B392 is a member of the medium chain acyl-CoA synthetase family. Biochemistry 45, 10008–10019 (2006).

60. Subrahmanyan, S. & Croman, J. E., Overproduction of a functional fatty acid biosynthetic enzyme blocks fatty acid synthesis in Escherichia coli. J. Bacteriol. 180, 4359–4362 (1998).

61. Lai, C. Y. & Croman, J. E. Isolation and characterization of beta-ketoacyl-ACP carrier protein reductase (FabG) mutants of Escherichia coli and Salmonella enterica serovar Typhimurium. J. Bacteriol. 186, 1689–1687 (2004).

62. Bergler, H. et al. Protein EnvM is the NADH-dependent enoyl-ACP reductase (FabM) of Escherichia coli. J. Biol. Chem. 269, 5493–5496 (1994).

63. Rock, C. O. & Croman, J. E., Jr Acyl carrier protein from Escherichia coli. Methods Enzymol. 71, 341–351 (1981). Pt C.

64. Ren, J. et al. Acetylation of lysine 201 inhibits the DNA-binding activity of PhoP to regulate Salmonella virulence. PLoS Pathog. 12, e1005458 (2016).

65. Otwinowski, Z. & Minor, W. (20) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326 (1997).

66. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr. D. Biol. Crystallogr. 63, 32–41 (2007).

67. Qiu, X. et al. Crystal structure and substrate specificity of the beta-ketoacyl-ACP carrier protein synthase III (FabH) from Staphylococcus aureus. Protein Sci. 14, 2087–2094 (2005).

68. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D. Biol. Crystallogr. 66, 486–501 (2010).

69. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D. Biol. Crystallogr. 66, 213–220 (2010).

70. Burley, S. K. et al. RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. Nucleic Acids Res. 47, D464–D474 (2019).

71. Pettersen, E. F. et al. UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612 (2004).

72. Kozakov, D. et al. The ClusPro web server for protein-protein docking. Nat. Protoc. 12, 255–278 (2017).

73. Trotj, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455–461 (2010).

74. Wischgoll, S. et al. Structural basis for promoting and preventing decarboxylation in glutaryl-coenzyme A dehydrogenases. Biochemistry 49, 5350–5357 (2010).

75. Morris, G. M. et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J. Comput. Chem. 30, 2785–2789 (2009).
