CellPress

Article

Molecular Basis of BioJ, a Unique Gatekeeper in Bacterial Biotin Synthesis

Catalytic triad of BioJ

H278

D248

S151

Growth of biotin auxtropic strain

+DTB

+BioJ

Chemical reaction by BioJ

H_2O

CH_3OH

H_3CO

Me-pimeloyl-ACP

S—ACP

BioJ

Pimeloyl-ACP

S—ACP

Wenhui Wei,
Hongxin Guan,
Tong Zhu, Sitao
Zhang,
Chengpeng Fan,
Songying Ouyang,
Youjun Feng

ouyangsy@fjnu.edu.cn (S.O.)
fengyj@zju.edu.cn (Y.F.)

HIGHLIGHTS

BioJ is a distinct
gatekeeper of biotin
synthesis

We report structural and
functional definition of
BioJ

We propose a working
model for BioJ-substrate
binding in addition to
catalytic triad

It furthers our
understanding how the
substrate is gatekept by
BioJ in biotin synthesis

DATA AND CODE
AVAILABILITY
6K1T

iScience

CellPress

Article

Molecular Basis of BioJ, a Unique Gatekeeper in Bacterial Biotin Synthesis

Wenhui Wei,[1,3,7] Hongxin Guan,[2,4,7] Tong Zhu,[5,7] Sitao Zhang,[2,4,7] Chengpeng Fan,[6] Songying Ouyang,[2,4,*] and Youjun Feng[1,3,8,*]

SUMMARY

Biotin is an indispensable cofactor in the three domains of life. The unusual virulence factor BioJ of *Francisella* catalyzes the formation of pimeloyl-ACP, an intermediate in biotin synthesis. Here, we report the 1.58 Å crystal structure of BioJ, the enzymatic activity of which is determined with the *in vitro* reconstituted reaction and biotin bioassay *in vivo*. Unlike the paradigm BioH, BioJ displays an atypical α/β-hydrolase fold. A structurally conserved catalytic triad (S151, D248, and H278) of BioJ is functionally defined. A proposed model for BioJ catalysis involves two basic residues-rich cavities, of which cavity-1, rather than cavity-2, binds to the ACP moiety of its physiological substrate, pimeloyl-ACP methyl ester. In summary, this finding provides molecular insights into the BioJ gatekeeper of biotin synthesis.

INTRODUCTION

Biotin (vitamin B7 or H), is an indispensable micronutrient required in three domains of life (Beckett, 2007, 2009; Rodionov et al., 2002). As the carboxyl group carrier, this prosthetic cofactor biotin participates into the enzymatic reactions of carboxylation, decarboxylation, and trans-carboxylation in the context of central metabolism, such as type II fatty acid synthesis (Beckett, 2007, 2009). Plants and certain microorganisms possess the ability of *de novo* biotin synthesis, whereas mammals and birds do not (Cronan, 2014). Therefore, it is reasonable that a scavenging/uptake pathway of biotin from food is present in animals or from the inhabiting niche in the biotin auxotrophic microorganisms (Hebbeln et al., 2007) (like *Lactococcus* [Zhang et al., 2016] and *Streptococcus* [Ye et al., 2016]). The most of knowledge on biotin synthesis is from studies with the model bacterium *Escherichia coli* (Cronan, 2014). In general, the *de novo* pathway involves two critical steps: (1) the synthesis of pimelate, a seven-carbon α, ω-dicarboxylate intermediate (Cronan and Lin, 2010; Lin and Cronan, 2010; Lin et al., 2010); and (2) the assembly of the fused heterocyclic rings of biotin (Rodionov et al., 2002). The latter step of the biotin synthesis route is well known for years (Lin and Cronan, 2010), which is extremely conserved and successively catalyzed by four enzymes, namely, BioF, BioA, BioD, and BioB (Figure 1A). In contrast, the earlier steps by which the precursor, pimeloyl moiety (pimeloyl-CoA or pimeloyl-ACP), is synthesized remained a mystery for around 70 years (Cronan, 2014), until a recent discovery by Lin et al. (Lin and Cronan, 2010, 2012; Lin et al., 2010) that biotin synthesis begins by hijacking a modified type II fatty acid synthesis pathway (FAS II). Unlike acetyl-CoA, a normal primer of FAS II synthesis, the molecule of methyl malonyl-CoA is recruited and elongated by two FAS II cycles, giving the product of methyl pimeloyl-ACP (Lin et al., 2010). The unusual primer, methyl malonyl-CoA (ACP) is produced in the first-committed reaction, i.e., BioC-catalyzed SAM-dependent methylation (Lin and Cronan, 2012). BioH, a prototypical member of pimeloyl-acyl carrier protein (ACP) methyl esterase, removes an extra methyl moiety from pimeloyl-ACP methyl ester (Me-pimeloyl-ACP) to release an intermediate product of pimeloyl-ACP (Agarwal et al., 2012; Sanishvili et al., 2003) (Figure 1A), which directly enters as a dedicated substrate of BioF, into the latter steps of biotin heterocyclic ring formation (Agarwal et al., 2012; Rodionov et al., 2002). The cleavage of Me-pimeloyl-ACP by BioH efficiently prevents its further elongation and functions as a gatekeeper of connecting FAS II with biotin biosynthesis (Lin and Cronan, 2010; Lin et al., 2010). Unlike the representative "BioC-BioH" pathway, many *bioC*-containing microorganisms lack *bioH* homologues, raising the possibility that non-homologous isoenzymes are present (Shapiro et al., 2012). As expected, no less than four additional enzymes have been discovered, which consistently belong to the superfamily of α/β hydrolase (Figure 1B) (Bi et al., 2016; Feng et al., 2014; Shapiro et al., 2012). These paradigmatic members separately refer to BioK in *Synechococcus* (Shapiro et al., 2012), BioG of *Haemophilus influenzae* (Shapiro et al., 2012; Shi et al., 2016), BioJ exclusively in *Francisella* (Feng et al., 2014), and BioV restricted to *Helicobacter* (Bi et al., 2016), respectively.

[1]Department of Pathogen Biology & Microbiology and General Intensive Care Unit of Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310058, China

[2]The Key Laboratory of Innate Immune Biology of Fujian Province, Provincial University Key Laboratory of Cellular Stress Response and Metabolic Regulation, Biomedical Research Center of South China, Key Laboratory of OptoElectronic Science and Technology for Medicine of the Ministry of Education, College of Life Sciences, Fujian Normal University, Fuzhou 350117, China

[3]College of Animal Science, Zhejiang University, Hangzhou 310058, China

[4]Laboratory for Marine Biology and Biotechnology, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao 266237, China

[5]CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China

[6]Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Wuhan University, Wuhan 430071, China

[7]These authors contribute equally

[8]Lead Contact

*Correspondence: ouyangsy@fjnu.edu.cn (S.O.), fengyj@zju.edu.cn (Y.F.)

https://doi.org/10.1016/j.isci.2019.08.028

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

A
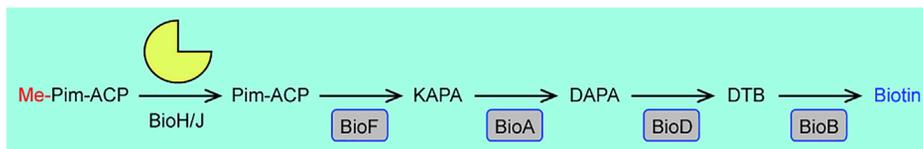
B
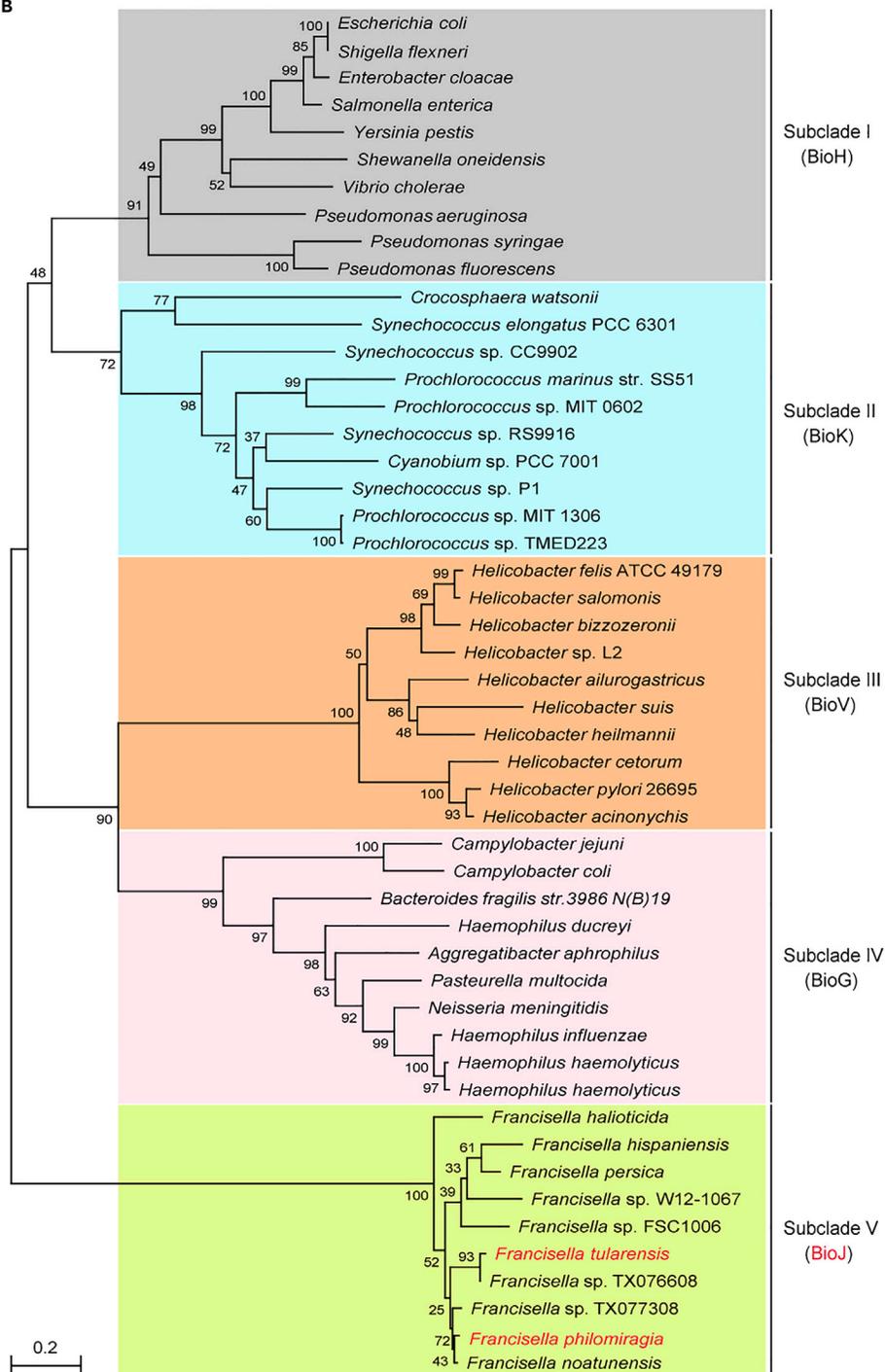
Figure 1. A Role of BioJ in Biotin Synthesis and its Phylogeny

(A) Scheme for physiological role of BioJ in biotin biosynthesis pathway. ACP, Acyl carrier protein; Me-Pim-ACP, Methyl pimeloyl-ACP ester; Pim-ACP, pimeloyl-ACP; KAPA, 7-keto-8-aminopelargonic acid; DAPA, 7,8-diaminopelargonic acid; DTB, dethiobiotin; BioF, 7-keto-8-aminopelargonic acid (KAPA) synthase; BioA, 7,8-diaminopelargonic acid aminotransferase; BioD, Dethiobiotin synthase; BioB, Biotin synthase.

(B) Phylogeny of a family of BioJ-containing α/β-hydrolases. The protein sequences of pimeloyl-ACP methyl esterase family were collected from NCBI database. The maximum likelihood-based phylogenetic analysis was conducted using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo), and the final output is given with MEGA (https://www.megasoftware.net/mega4).

Tularemia (i.e., rabbit fever) is a highly infectious zoonosis (Oyston et al., 2004), whose spread/transmission is mainly dependent on arthropod vectors, such as ticks (Santic et al., 2010). The causative agent of this disease refers to *Francisella tularensis*, a facultatively intracellular Gram-negative bacterium, which possesses an unusual ability to primarily infect macrophage cells within the host (Celli and Zahrt, 2013). As a category A bioterrorism agent in the United States, *F. tularensis* seems to be a most virulent bacterium because an inhalation of as few as 10 bacteria is sufficient to result in severe and even fatal disease (Jones et al., 2012). On bacterial uptake by and/or entry into macrophage, *F. tularensis* exploits multiple strategies to rapidly respond the limited nutrition (e.g., cysteine [Alkhuder et al., 2009] and biotin [Napier et al., 2012]) within the harsh micro-environment of phagocytic cells accompanied with the burst of reactive oxygen species (Celli and Zahrt, 2013). It is a prerequisite for *F. tularensis* as a cytosolic pathogen to escape into host cytosols for its successful proliferation and survival during the life cycle of infection within hosts (Celli and Zahrt, 2013; Ray et al., 2009). A genome-wide *in vivo* negative screen originally suggested that *bioJ* (formerly designated as FTN_0818) of *F. tularensis* is a genetic determinant essential for both intracellular replication and bacterial infection in mice (Weiss et al., 2007). Then, Napier et al. (2012) reported that the *bioJ* protein product links biotin biosynthesis to efficient escape from the *Francisella*-containing phagosome, implying biotin as a nutritional limitation factor during infection. Subsequently, we found that BioJ is a non-homologous isoenzyme of *E. coli* BioH, the best-studied Me-pimeloyl-ACP carboxyl-esterase (Feng et al., 2014). Consistent with that of BioH (Lin et al., 2010), our data *in vitro* and *in vivo* demonstrated that BioJ also functions as a gatekeeper in *Francisella* biotin synthesis and determines the chain length of the biotin valeryl side chain (Feng et al., 2014). A similar scenario was also seen in the other cytosolic pathogen, *Mycobacterium tuberculosis*, because removal of *bioA*, an essential gene of *de novo* biotin synthesis impairs the establishment and maintenance of its chronic tuberculosis infections (Woong Park et al., 2011). This finding represents metabolic evidence for the link of biotin synthesis to bacterial virulence. However, structural and mechanistic aspects of BioJ remains largely elusive.

In this study, we aimed to close this knowledge gap. Here we report a high-resolution X-ray crystal structure of BioJ at 1.58 Å, illustrating a distinct architecture from the paradigm BioH gatekeeper of biotin synthesis. In addition to the biochemical role played by BioJ in *Francisella* biotin synthesis, structure-guided functional analyses define a catalytic triad (S151, D248, and H278) and a cavity for binding of its physiological substrate Me-pimeloyl-ACP. In summary, this finding extends our understanding of the biotin synthesis pathway and provides the structural basis for BioJ virulence factor, a potential drug target against the deadly infections with *Francisella*.

RESULTS

BioJ Is a Gatekeeper in Biotin Synthesis

In total, five types of α/β-hydrolases have been assigned to demethylase of pimeloyl-ACP methyl ester. The phylogeny of these non-homologous isoenzymes suggested that they are evolutionarily distinct (Figure 1B). Unlike the well-studied BioH, which is distributed in γ-proteobacteria, BioJ is restricted to the zoonotic pathogen *Francisella* and constitutes a unique sub-lineage, Subclade V (Figure 1B). It seems very true that BioJ terminates an alternative route of pimeloyl moiety to be elongated, assuring its entry of this C7 chain into the latter steps of biotin synthesis pathway (Figures 1A and 2A). *In vitro* enzymatic assays elucidated that BioJ (but not BioZ) cleaves its physiological substrate Me-pimeloyl-ACP into Pimeloyl-ACP (Figure 2B). Similar to BioH (Lin et al., 2010), BioJ is also a promiscuous enzyme in that it removes the methyl moiety from the non-physiological substrates, acyl-ACP methyl esters (like C6 and C8, Figure 2C).

To address the role of BioJ *in vivo*, an indicator strain-based biotin bioassay was used, in which a visible deposition of a red formazan suggests bacterial viability (Figures 2D and 2E) (Feng et al., 2014). This strain ER90 (Δ*bioF* Δ*bioC* Δ*bioD*) is biotin auxotrophic, whose growth depends on the availability of biotin or its precursor
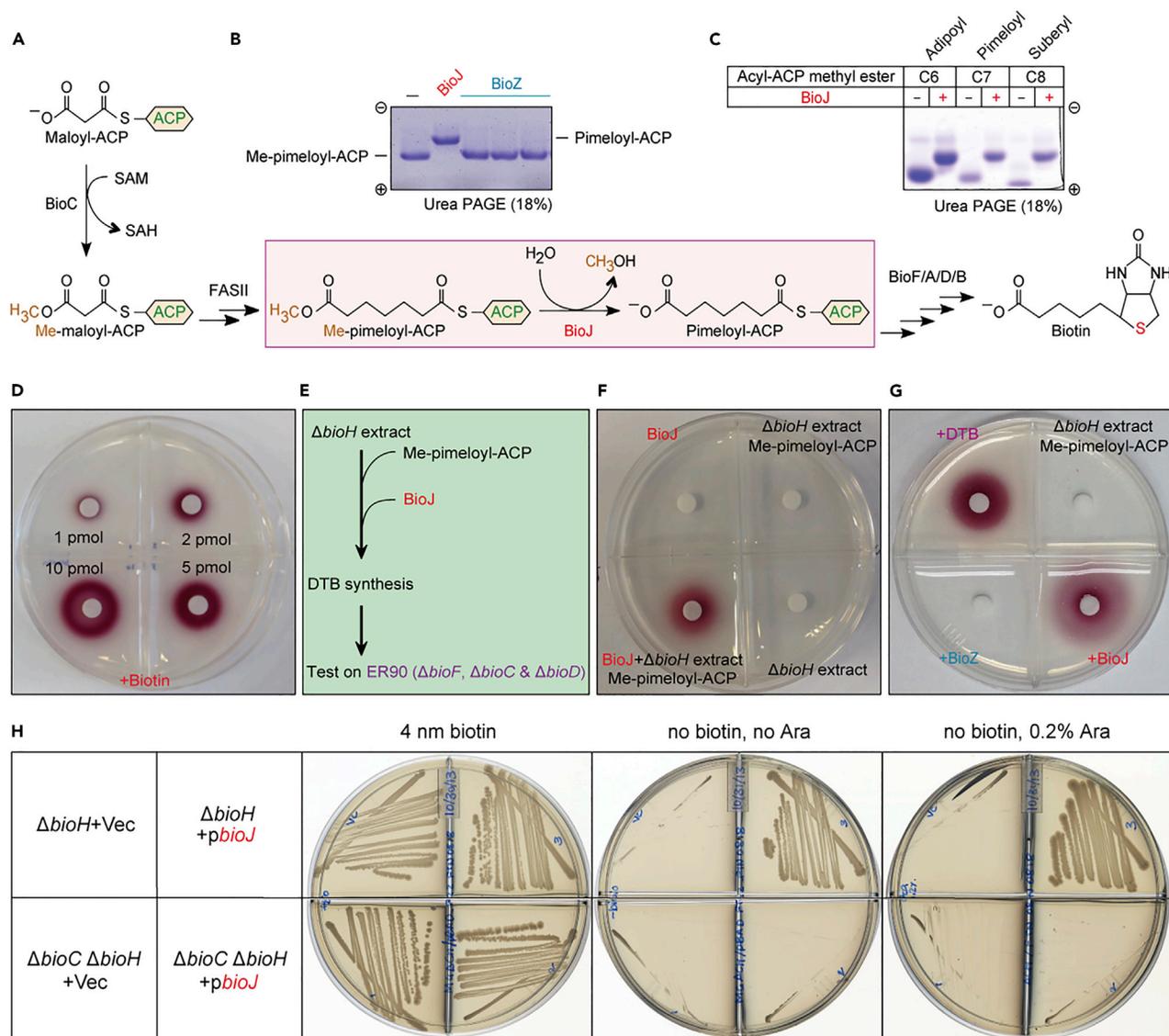
Figure 2. The *Francisella* BioJ Catalyzes an Essential Reaction of Biotin Synthesis

(A) Schematic diagram for the biotin synthesis pathway in *Francisella*. The enzymatic reaction catalyzed by BioJ is underlined by a rectangle with pink background. SAM, *S*-adenosylmethionine; SAH, *S*-adenosylhomocysteine; BioC, Malonyl-ACP O-methyltransferase; Me-malonyl ACP, Malonyl-ACP methyl ester; Me-pimeloyl-ACP, Pimeloyl-ACP methyl ester; BioJ, the iso-enzyme of BioH, a pimeloyl-ACP methyl ester carboxylesterase; FASII, type II fatty acid synthesis pathway.

(B) *In vitro* enzymatic assays reveal that BioJ can hydrolyze pimeloyl-ACP methyl ester to pimeloyl-ACP, rather than BioZ. The BioZ of *Agrobacterium* is used as a negative control.

(C) Eighteen percent urea PAGE-aided analyses of substrate specificity of BioJ. The plus sign denotes addition of BioJ (or BioZ) enzyme, whereas the minus sign refers to no addition of the protein.

(D) Use of bioassay to visualize bacterial growth of the biotin auxotrophic strain ER90 on minimal media at varied level of biotin.

(E) Scheme of the system of *in vitro* biotin synthesis.

(F) Biotin bioassay suggests that BioJ enzyme works together with the ΔbioH extract to synthesize biotin (and/or DTB) in the presence of its physiological substrate pimeloyl-ACP methyl ester.

(G) BioJ (rather than BioZ) can restore the ability of ΔbioH in biotin (DTB) synthesis. The bioassay was performed using the reporter strain ER90 (ΔbioF bioC bioD). The red formazan deposit is due to the reduction of the indicator tetrazolium, and this suggests the presence of biotin supply (or synthesized in the *in vitro* system) conferring bacterial growth of the biotin auxotrophic strain ER90 on the non-permissive condition.

(H) Functional expression of *Francisella bioJ* rescues its growth of the *E. coli bioH* mutant on the non-permissive condition of biotin-free medium. *E. coli* strains used here included the single ΔbioH mutant and a double mutant (ΔbioC/ΔbioH). To test the function of *bioJ* in vivo, the strains expressing *bioJ* (or empty vector) were stripped on minimal media with/without 4 nM biotin and maintained at 30°C for 36h. Vec, pBAD322 vector; Ara, arabinose; DTB, Dethiobiotin.

dethiobiotin (DTB) (Figure 2D). The addition of BioJ into the *in vitro* system of biotin synthesis (Figure 2E) conferred bacterial growth of the indicator strain on the biotin-free, non-permissive condition (Figure 2F), which is consistent with our earlier observation (Feng et al., 2014). Evidently, the growth potency by BioJ is quite similar to scenarios with 5 pmol of biotin (Figure 2D) or its precursor DTB (Figure 2G). However, BioZ cannot work as BioJ does in our assays (Figures 2B and 2G), suggesting a different mechanism. Finally, the plasmid-borne *bioJ* expression restored the growth of the *E. coli* Δ*bioH* mutant but not the double mutant of Δ*bioC*/Δ*bioH* (Figure 2H), indicating that BioJ is functionally exchangeable with BioH. Given the fact that BioJ displays BioH-like activity, we therefore believe it is a unique gatekeeper of biotin synthesis (Feng et al., 2014).

Overall Structure of BioJ

The crystals of BioJ diffracted the X-ray to 1.58 Å with the space group of $P12_11$ (Table 1) and contained just one copy of BioJ per asymmetric unit (ASU). We solved the structure of BioJ by molecule replacement using the structure of esterase 2 (EST2) from *Alicyclobacillus acidocaldarius* (PDB:1QZ3) as the search model (De Simone et al., 2004). In the finally refined model, we located 299 amino acids of BioJ except the first methionine and the loop of E213 to S217 (relevant data collection and refinement statistics are presented in Table 1). As a member of the α/β-hydrolase superfamily enzymes, BioJ, BioH, and BioG come from different bacteria and lack significant sequence identity (19% identity shared by BioJ and BioH; 15.7% identity between BioJ and BioG), but all of them contain two domains, a core domain and an α-helical lid domain. The structural comparison gave the root-mean-square deviation (RMSD) of 2.2 Å (165 Cα atoms between BioJ and BioH) and 2.7 Å (125 Cα atoms between BioJ and BioG), respectively. BioH and BioG have a core domain with just a seven-member central β-sheet flanked on either side by α-helices (or loop) and covered by an auxiliary domain (Figures 3A–3D). In contrast, the core domain of BioJ contains eight central β-sheets (Figures 3E and 3F), which is sandwiched in the middle by three helices on each side (h4, h12, and h13 on one side; h5, h6, and h11 on the other side, Figure 3E).

BioH is architecturally similar to that of BioG (Figures 3A–3D), whose RMSD value is 1.078 Å. Meanwhile, the RMSD value of BioJ is 1.110 Å for BioG, and 1.175 Å for BioH, respectively. The major difference is described as follows: four (BioH, Figures 3A and 3B) or three (BioJ, Figures 3E and 3F) of the six α-helices flanking both sides of the core β-sheet are replaced with long loops in BioG (Figures 3C and 3D), thus forming an unusual α/β-hydrolase fold in BioG (Shi et al., 2016). The lid domain of BioJ caps on its core domain by three helices (α1-α3) at the N terminus and four helices (α7-α10) from the C terminus (Figures 3E and 3F). Topological comparison further illustrates obvious difference among the lid domains of BioH (Figure 3B), BioG (Figure 3D), and BioJ (Figure 3F). In relation to the counterpart of BioH that consists of four continuous helices (spanning from K121 to T185, Figure 3B) (Agarwal et al., 2012), BioJ exhibits a bigger lid domain comprising two separate parts (1–42 and185–231, Figure 3F). In fact, the lid domain of BioG is the smallest in that it is of only 57 residues (Y101–Q157, Figure 3D) (Shi et al., 2016).

Functional Validation of Catalytic Triad in BioJ

The prototypical catalytic triad (S82, H235, and D207, Figure 4A) of BioH lies at α10, a loop between β6 and α8 and loop between β7 and α9 (Figure 3B), respectively. Although it is an atypical α/β hydrolase, BioG also has evolved into a similar catalytic triad (S65, D175, and H200, Figure 4B) (Shi et al., 2016). Not surprisingly, the structure of BioJ defines a canonical catalytic triad (S151, D248, and H278) at the interface of the core domain and the auxiliary domain (Figures 3F and 4C), which is analogous to the counterparts in BioH (Figure 4A) (Agarwal et al., 2012) and BioG (Figure 4B) (Shi et al., 2016). Multiple sequence alignment of BioJ homologs showed that they exhibit nearly 80% amino acids identity and consistently possess an identical catalytic triad candidate across different species of *Francisella* (Figure S1).

The conserved steric configuration of catalytic pockets across BioH (Figure 4A), BioG (Figure 4B), and BioJ (Figure 4C) verified our earlier prediction for BioJ active sites guided by structural modeling (Feng et al., 2014). Indeed, the alanine substitution revealed that the point-mutants of BioJ (e.g., S151A and H278A) inefficiently hydrolyze its substrate of Me-pimeloyl-ACP (Figure S2). In the assays of genetic complementation, the wild-type version of *bioJ* allowed the Δ*bioH* mutant to grow well on the minimal agar plates lacking biotin. However, none of three point mutants of *bioJ* (S151A, D248A, and H278A) supported bacterial growth of the Δ*bioH* mutant on the biotin-free condition, which is almost identical to scenarios with only empty vector introduced (Figure 4D). In addition, the supplementation of exogenous biotin (4 nM) reversed/bypassed functional defection of catalytic triad (Figure 4D). Together, it constitutes a first structure-function proof of catalytic triad in BioJ.

Dataset	BioJ
PDB ID	6K1T
Data Collection	
Beamline	BL-17U1, SSRF
Wavelength	0.979
Resolution range[a]	51.32–1.58 (1.67–1.58)
Space group	$P12_11$
Cell dimensions	
a, b, c (Å)	43.94, 67.31, 55.57
α, β, γ (°)	90, 112.55, 90
Total reflections	1,12,765 (17,680)
Unique reflections	37,190 (5,609)
Multiplicity	3.0
Completeness (%)	91.5 (95.6)
Mean I/sigma(I)	6.7 (3.6)
Wilson B-factor	17.46
R-merge	0.129 (0.745)
R-meas	0.179 (1.022)
R-pim	0.123 (0.695)
CC1/2	0.972 (0.620)
Refinement	
Reflections used in refinement	36,992 (3,847)
Reflections used for R-free	1,808 (172)
R-work	0.2067 (0.2591)
R-free	0.2350 (0.2999)
Number of non-hydrogen atoms	2,659
Macromolecules	2,443
Solvent	216
Protein residues	299
RMS (bonds)	0.006
RMS (angles)	0.81
Ramachandran favored (%)	96.61
Ramachandran allowed (%)	3.05
Ramachandran outliers (%)	0.34
Rotamer outliers (%)	0.00
Clash score	4.32
Average B-factor	20.62

Table 1. The X-Ray Crystallographic Data Collection and Refinement Statistics

(Continued on next page)

Dataset	BioJ
Macromolecules	19.97
Solvent	27.99

Table 1. Continued

[a]Highest resolution shell is shown in parentheses.

Structural Insight into BioJ-ACP Interplay

Our long-term exploration had no success in obtaining the structure of BioJ complexed with Me-pimeloyl-ACP. However, the availability of the complex structure of the substrate pimeloyl-ACP methyl ester with the point mutant (S82A) of BioH (Figures S3, 5A, and 5B) rendered it possible to gain a putative glimpse of the interplay between the substrate gatekeeper BioJ and the ACP group of its substrate (Figures 5E and 5F). First, four basic residues (namely, R138, R142, R155, and R159) at the capping domain of BioH close to its cavity (Figures 5A and 5B) are mapped to form ionic interactions with the ACP group (Figure S3A). This is due to the formation of salt bridges between the side chain of basic amino acids with positively electronic density area (R138 and R142 in Figure S3B; R155 and R159 in Figure S3C) and the negatively charged ACP-α2 (Q13, D34, and D37 in Figure S3B; E46, I53, and D55 in Figure S3C). Using these criteria, a similar pattern of BioG contacting ACP is also defined by Shi et al. (Shi et al., 2016), which includes the three basic residues, namely, K118, K127, and R132 (Figures 5C and 5D). As expected, we also illustrated that BioJ has a similar positively charged interface with the potential to interact with the ACP moiety (Figures 5E and 5F). This interface is mainly constituted by the three key basic amino acids (K184, K221, and R223) in Cavity-1 at the lid domain close to cavity (Figures 5E and 5F). More importantly, the synergistic roles of the three residues are demonstrated in the assays for growth curves (Figure 5G) and bacterial viability on the agar plates carrying 4 nM biotin (Figure S4A) or lacking biotin (Figure S4B), following the genetic manipulation of *bioJ* with a series of combined alanine substitution (single, double, and triple). The alignment of BioJ with esterase 2 (EST2) from the thermophilic bacterium *Alicyclobacillus acidocaldarius* exhibits 30.95% identity (Figure S5). Although their catalytic triads are almost identical (i.e., S155, D252, and H282 in EST2; S151, D248, and H278 in BioJ), the substrate-interacting pocket between BioJ and EST2 varies greatly (Figure S5). In brief, the long hydrophobic tunnel in EST2 for the entry of the hexadecane moiety comprises ten critical residues (N15, S26, S35, L36, G82, V87, S211, L215, F284, and F287, Figure S5) (De Simone et al., 2004). By contrast, the three distinct residues (K184, K221, and R223) are implied to participate in the interplay between BioJ and its substrate pimeloyl-ACP methyl ester (Figure 5). This scenario is also seen with those of BioH (Agarwal et al., 2012) and BioG (Shi et al., 2016) (Figure 5).

In particular, the analyses of cavities and molecular docking were applied to infer the residues participating in the BioJ-ACP Interaction, which allowed us to observe the presence of a unique Cavity-2 in BioJ (Figure S6A), but not in canonical BioH or atypical BioG. It indicated that two channels went through the catalytic triad of BioJ, the two ends of which are Cavity-1 and a unique Cavity-2 (Figure S6A). Interactions with residues in Cavity-1 was not observed from the docking result, because the space of channel-1 formed by Cavity-1 is insufficient to accept the substrate pimeloyl-ACP methyl ester. This indicates that the configuration of BioJ may alter during the catalytic process. Unlike those of canonical BioH and atypical BioG, Cavity-2 consists of three continuous α-helices (α1, α2, and α3) from the N terminus of BioJ (Figures 3E and 3F). The channel generated by Cavity-2 is opposite to the other one formed by Cavity-1 (Figure S6A). The analysis of cavities suggests that BioJ creates a bigger channel-2 connected with Cavity-2 than the channel-1(Figure S6A). Of note, both of them are interconnected and go through the catalytic triad. Following the docking process, the analysis of electrostatic distribution further unveiled the occurrence of three basic acids (K29, K40, and R41) in the vicinity of Cavity-2 (Figures S6A and S6B). However, bacterial growth assays along with the alanine substitutions confirm that they are not implicated in the BioJ function (Figure S6C). This suggested that Cavity-2 is not exploited for interacting with the substrate of pimeloyl-ACP methyl ester but might be for expelling product of methyl.

DISCUSSION

In relation to the *bio* operon, the genomic context of these gatekeeper-encoding genes is divergent, although they are consistently grouped into the superfamily of α/β-hydrolase (Shapiro et al., 2012). Unlike the paradigm *E. coli bioH* that is not integrated into *bio* operon, the counterpart of *Pseudomonas* appears within the *bioBFHCD* operon (Shapiro et al., 2012). Quite different from *bioGC* organization in
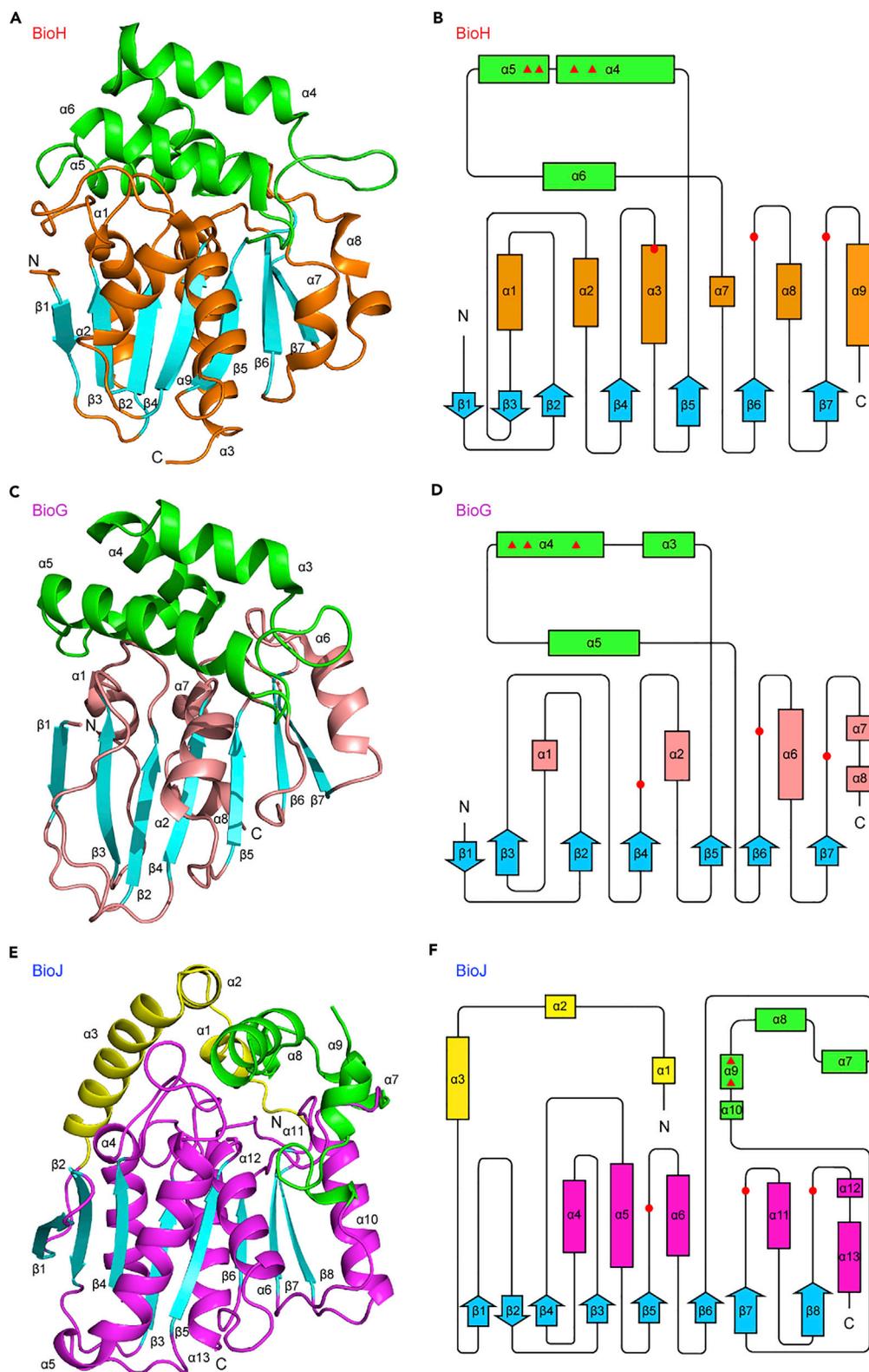
Figure 3. Comparison of Overall Structures of BioH and BioJ, Two pimeloyl-ACP Methyl Ester Carboxylesterases
Ribbon structure (A) and topological diagram (B) of the *E. coli* BioH. Overall structure of BioH (A) consists of two domains, the capping domain and the α/β core domain. The helices 4, 5, and 6 of the capping domain are colored green, whereas

Figure 3. *Continued*

the α/β core domains are labeled orange in α-helices and cyan in β-sheets. In the topological diagram of BioH (B), an arrow, rectangle and line refer to β-sheet, α-helix, and loop, respectively. The four residues that make salt bridge with ACP are underscored with red triangle in capping domain helix 4 and 5. The catalytic triad (S82, D207, and H235) located in the core domain is featuring with red dots. Overall architecture (C) and topological scheme (D) for the *Haemophilus influenzae* BioG. The lip domain of BioG is composed of four α-helices (α3–α6) and highlighted by green. In the α/β core domain, there are seven β-sheets (cyan) circled by three α-helices (grayish peach). In the topological diagram of BioG (D), typical catalytic triads (S65, D175, and H200) are labeled with red dots. Overall architecture (E) and topological scheme (F) for the *Francisella* BioJ. In the lip domain of BioJ, the α-helices 1, 2, and 3 are highlighted by yellow, whereas the α-helices 7, 8, 9, and 10 interacting with ACP are colored green. In the α/β core domain, the eight β-sheets (cyan) are circled with six α-helices (pink). Three basic amino acids are marked by red triangle in the helix 6 plus loop between sheet 6 and helix 7. Typical catalytic triads (S151, D248, and H278) are labeled with red dots. PDB entry is 1M33 for BioH, 6K1T for BioJ, and 5GNG for BioG.

Campylobacter and *Haemophilus* (Shapiro et al., 2012), the *Francisella bioJ* acts as a neighbor, but not within its *bioBFCD* operon (Feng et al., 2014). Thus, it seems likely that these gatekeeper enzymes (synthesizing pimeloyl moiety of biotin) are "wild cards" during the ongoing domestication of the *bio* operon. Probably, *bioJ* is a partially domesticated gene in *Francisella* (Feng et al., 2014). Intriguingly, together with the observation by Napier and coworkers (Napier et al., 2012), we speculated that this gatekeeper BioJ is associated with *Francisella* virulence (Feng et al., 2014). In addition, we found that efficient utilization of biotin by BplA (FTN_0568), a biotin protein ligase lacking a DNA-binding motif, is essential for the intracellular pathogen *F. novicida* to survive within macrophages and mice (Feng et al., 2015). Because the disruption of *de novo* biotin synthesis impairs the initiation and maintenance of chronic infection with *Mycobacterium tuberculosis* (Ren et al., 2016), it is reasonable to propose biotin as a nutritional (or restricted) virulence factor.

Intriguingly, the physiological requirement of biotin in *Francisella* (Feng et al., 2014) is much less than that of *M. smegmatis* (Wei et al., 2018) and *Agrobacterium tumefaciens* (Feng et al., 2013). That is because a single biotinylated AccB protein is predicted, whereas multiple biotinylated enzymes of *M. smegmatis* (Wei et al., 2018) and *A. tumefaciens* (Feng et al., 2013) are experimentally verified. Clearly, non-homologous isoenzymes of BioJ gatekeeper of biotin synthesis are distributed within a family of diversified pathogens (Figure 1), which is partially featuring with BioH of *Salmonella* and *Vibrio*, BioG of *Campylobacter* and *Neisseria*, and BioV of *Helicobacter*. Subsequently, the question to ask is if these isoenzymes participate in bacterial virulence. Probably it might consolidate the aforementioned hypothesis in view of a common role played by biotin metabolism in bacterial colonization, competitive infection, and survival within hosts.

The biochemical and structural data reported here represent the molecular basis for this atypical BioJ gatekeeper in biotin synthesis. A working model for BioJ action is proposed here, which is described with three putative steps as follows: (1) the recognition of BioJ to Me-pimeloyl-ACP; (2) the removal of methyl group from Me-pimeloyl-ACP by catalytic triad; and (3) configuration change-dependent exclusion of pimeloyl-ACP product from Cavity-1 and release of methyl group via Cavity-2 (Figure S6A). As for the final step, the core size of Cavity-1 presumably becomes smaller to extrude pimeloyl-ACP, whereas the core size of Cavity-2 is enlarged for the release of the methyl product. Because its overall architecture is quite different from those of BioH (Agarwal et al., 2012) and BioG (Shi et al., 2016), the high-resolution structure of BioJ we solved (Figure 3) extends our mechanistic understanding of the functional unification across BioH-like esterase within the distinct evolutionary placement (Figure 1). The fact that the conserved catalytic triad is shared among BioJ (Figure 4C), BioH (Figure 4A) (Agarwal et al., 2012), and BioG (Figure 4B) (Shi et al., 2016) suggests a promising anti-virulence drug target. Similarly, the predictive substrate-binding mechanism of BioJ offered us an avenue for the discovery of leading anti-bacterial drugs. Taken together, large-scale screen and/or computational design of small molecule inhibitors targeting these two motifs of the *bio* gatekeeper represent a new approach for combating the deadly infections with multidrug-resistant superbugs.

Limitations of the Study

Although a complex structure of BioH with Me-Pim-ACP provides direct evidence for its binding to physiological substrate (Agarwal et al., 2012), we have no success in securing a complex structure of BioJ with Me-Pim-ACP. Therefore, it is probable that cavity analyses (and molecular docking)-based insights are
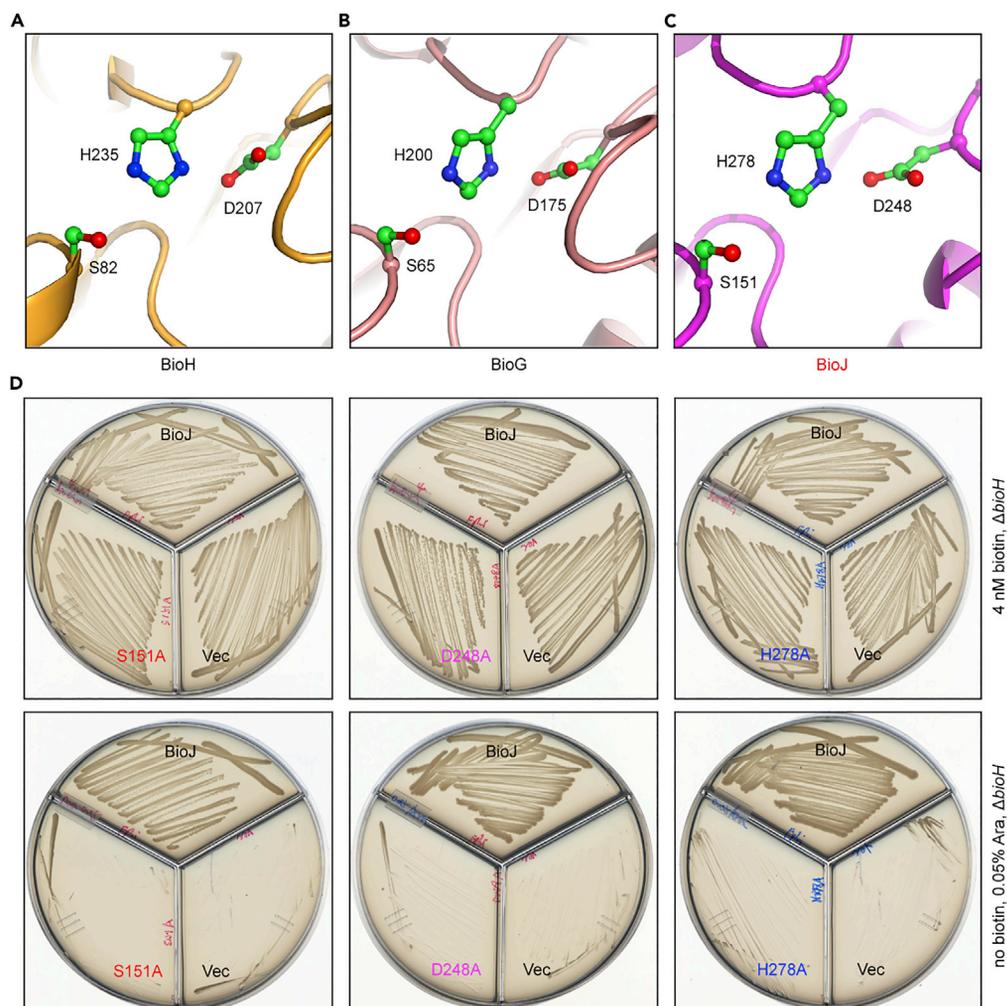
Figure 4. Structural and functional definition of catalytic triad motif in BioJ

(A–C) An enlarged view of the catalytic triad motif in BioH (A), BioG (B), and BioJ (C). Structural snapshots in (A) and (B) were separately generated via 180° rotation of the enlarged views illustrated in Figures 3A, 3C, and 3E.

(D) Genetic complementation reveals that the catalytic triad (S151, D248, and H278) is required for BioJ function. Site-directed mutagenesis was performed to give three *bioJ* derivatives, namely, S151A, D248A, and H278A. The Δ*bioH* strains of *E. coli* expressing *bioJ* (or its point mutants) were plated on minimal media with/without 4 nM biotin and incubated at 30°C for 36h. Ara, arabinose; Vec, an arabinose-inducible vector of pBAD322.

incomplete, with respect to an interplay between BioJ and its substrate. Also, *in vitro* enzymatic actions of all the mutant protein with defection in substrate binding are needed to consolidate the scenario seen with the assays of genetic complementation. In the near future, the picture of gatekeeper in biotin synthesis is relatively complete upon the availability of the BioV structure.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

DATA AND CODE AVAILABILITY

The accession number for the atomic coordinates of BioJ protein reported in this paper is PDB: 6K1T.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2019.08.028.
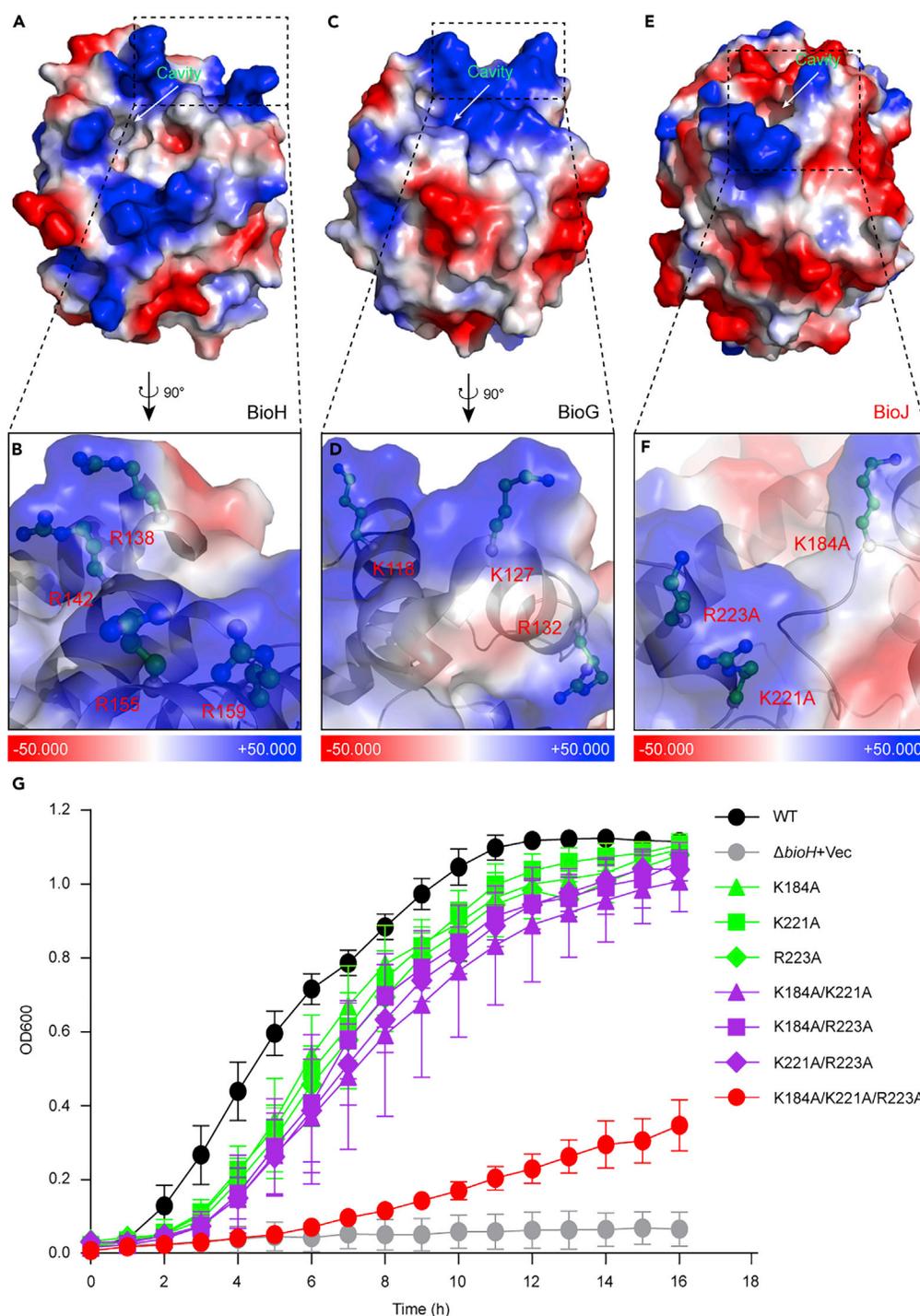
Figure 5. Structure-based search for binding sites of BioJ to its physiological substrate, methyl pimeloyl-ACP

(A–D) Electrostatic surfaces of the BioH (A) and BioG (C). Black dashed frame is used to highlight the positively charged surface of BioH forming salt bridges with the ACP domain of methyl pimeloyl-ACP near the substrate cavity. The PDB entries of BioH and BioG are 4ETW and 5GNG, respectively. An enlarged view of the positively charged basic residues of BioH (B) and BioG (D) that interact with the acidic ACP moiety of methyl pimeloyl-ACP. The side chains of basic amino acids (A and C) are given in a clockwise rotation of 90°.

(E) Surface structure of BioJ. A similar cavity surrounded with positively charged residues is labeled by black frame.

Figure 5. *Continued*

(F) Magnified view of positively charged, side-chain residues in BioJ. It is derived from (E) in the clockwise rotation of 90°. The three putative basic amino acids refer to K184, K221, and R223, respectively. The gradual changing polarity of the protein surface from negative to positive is marked by color from red to blue.

(G) Use of growth curves to evaluate a role played by the three basic residues-containing substrate cavity in the BioJ action. The recipient strain for *bioJ* and its mutants is STL24 (Δ*bioH*, Table S1), and plasmid vector is pBAD322 (Table S1). The addition of Arabinose (0.20%) induces the expression of *bioJ* (and/or its mutants) in STL24 that grows into M9 minimal media with glycerol as sole carbon source. It is given in means ±SD. Three independent experiments were conducted. The single, double, and triple mutants are colored green, magenta, and red.

AUTHOR CONTRIBUTIONS

Y.F. and S.O. designed and supervised this project; Y.F., W.W., S.Z., T.Z., and H.G. performed experiments; Y.F., S.O., C.F., W.W., and T.Z. analyzed the data and prepared figures; Y.F., T.Z., and S.O. drafted this manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Agarwal, V., Lin, S., Lukk, T., Nair, S.K., and Cronan, J.E. (2012). Structure of the enzyme-acyl carrier protein (ACP) substrate gatekeeper complex required for biotin synthesis. Proc. Natl. Acad. Sci. U S A *109*, 17406–17411.

Alkhuder, K., Meibom, K.L., Dubail, I., Dupuis, M., and Charbit, A. (2009). Glutathione provides a source of cysteine essential for intracellular multiplication of *Francisella tularensis*. PLoS Pathog. *5*, e1000284.

Beckett, D. (2007). Biotin sensing: universal influence of biotin status on transcription. Annu. Rev. Genet. *41*, 443–464.

Beckett, D. (2009). Biotin sensing at the molecular level. J. Nutr. *139*, 167–170.

Bi, H., Zhu, L., Jia, J., and Cronan, J.E. (2016). A biotin biosynthesis gene restricted to *Helicobacter*. Sci. Rep. *6*, 21162.

Celli, J., and Zahrt, T.C. (2013). Mechanisms of *Francisella tularensis* intracellular pathogenesis. Cold Spring Harb. Perspect. Med. *3*, a010314.

Cronan, J.E. (2014). Biotin and lipoic acid: synthesis, attachment, and regulation. EcoSal Plus *6*, https://doi.org/10.1128/ecosalplus.3.6.3.5.

Cronan, J.E., and Lin, S. (2010). Synthesis of the alpha, omega-dicarboxylic acid precursor of biotin by the canonical fatty acid biosynthetic pathway. Curr. Opin. Chem. Biol. *15*, 407–413.

De Simone, G., Mandrich, L., Menchise, V., Giordano, V., Febbraio, F., Rossi, M., Pedone, C., and Manco, G. (2004). A substrate-induced switch in the reaction mechanism of a thermophilic esterase: kinetic evidences and structural basis. J. Biol. Chem. *279*, 6815–6823.

Feng, Y., Chin, C.Y., Chakravartty, V., Gao, R., Crispell, E.K., Weiss, D.S., and Cronan, J.E. (2015). The atypical occurrence of two biotin protein ligases in *Francisella novicida* is due to distinct roles in virulence and biotin metabolism. MBio *6*, e00591.

Feng, Y., Napier, B.A., Manandhar, M., Henke, S.K., Weiss, D.S., and Cronan, J.E. (2014). A *Francisella* virulence factor catalyzes an essential reaction of biotin synthesis. Mol. Microbiol. *91*, 300–314.

Feng, Y., Zhang, H., and Cronan, J.E. (2013). Profligate biotin synthesis in alpha-proteobacteria - a developing or degenerating regulatory system? Mol. Microbiol. *88*, 77–92.

Hebbeln, P., Rodionov, D.A., Alfandega, A., and Eitinger, T. (2007). Biotin uptake in prokaryotes by solute transporters with an optional ATP-binding cassette-containing module. Proc. Natl. Acad. Sci. U S A *104*, 2909–2914.

Jones, C.L., Napier, B.A., Sampson, T.R., Llewellyn, A.C., Schroeder, M.R., and Weiss, D.S. (2012). Subversion of host recognition and defense systems by *Francisella* spp. Microbiol. Mol. Biol. Rev. *76*, 383–404.

Lin, S., and Cronan, J.E. (2010). Closing in on complete pathways of biotin biosynthesis. Mol. Biosyst. *7*, 1811–1821.

Lin, S., and Cronan, J.E. (2012). The BioC O-methyltransferase catalyzes methyl esterification of malonyl-acyl carrier protein, an essential step in biotin synthesis. J. Biol. Chem. *287*, 37010–37020.

Lin, S., Hanson, R.E., and Cronan, J.E. (2010). Biotin synthesis begins by hijacking the fatty acid synthetic pathway. Nat. Chem. Biol. *6*, 682–688.

Napier, B.A., Meyer, L., Bina, J.E., Miller, M.A., Sjostedt, A., and Weiss, D.S. (2012). Link between intraphagosomal biotin and rapid phagosomal escape in *Francisella*. Proc. Natl. Acad. Sci. U S A *109*, 18084–18089.

Oyston, P.C., Sjostedt, A., and Titball, R.W. (2004). Tularaemia: bioterrorism defence renews interest in *Francisella tularensis*. Nat. Rev. Microbiol. *2*, 967–978.

Ray, K., Marteyn, B., Sansonetti, P.J., and Tang, C.M. (2009). Life on the inside: the intracellular lifestyle of cytosolic bacteria. Nat. Rev. Microbiol. *7*, 333–340.

Ren, J., Sang, Y., Tan, Y., Tao, J., Ni, J., Liu, S., Fan, X., Zhao, W., Lu, J., Wu, W., et al. (2016). Acetylation of lysine 201 inhibits the DNA-binding ability of PhoP to regulate *Salmonella* virulence. PLoS Pathog. *12*, e1005458.

Rodionov, D.A., Mironov, A.A., and Gelfand, M.S. (2002). Conservation of the biotin regulon and the BirA regulatory signal in Eubacteria and Archaea. Genome Res. *12*, 1507–1516.

Sanishvili, R., Yakunin, A.F., Laskowski, R.A., Skarina, T., Evdokimova, E., Doherty-Kirby, A., Lajoie, G.A., Thornton, J.M., Arrowsmith, C.H., Savchenko, A., et al. (2003). Integrating structure, bioinformatics, and enzymology to discover function: BioH, a new carboxylesterase from *Escherichia coli*. J. Biol. Chem. *278*, 26039–26045.

Santic, M., Al-Khodor, S., and Abu Kwaik, Y. (2010). Cell biology and molecular ecology of *Francisella tularensis*. Cell Microbiol. *12*, 129–139.

Shapiro, M.M., Chakravartty, V., and Cronan, J.E. (2012). Remarkable diversity in the enzymes catalyzing the last step in synthesis of the pimelate moiety of biotin. PLoS One *7*, e49440.

Shi, J., Cao, X., Chen, Y., Cronan, J.E., and Guo, Z. (2016). An atypical α/β-hydrolase fold revealed in the crystal structure of pimeloyl-acyl carrier protein methyl esterase BioG from *Haemophilus influenzae*. Biochemistry *55*, 6705–6717.

Wei, W., Zhang, Y., Gao, R., Li, J., Xu, Y., Wang, S., Ji, Q., and Feng, Y. (2018). Crystal structure and acetylation of BioQ suggests a novel regulatory switch for biotin biosynthesis in *Mycobacterium smegmatis*. Mol. Microbiol. *109*, 642–662.

Weiss, D.S., Brotcke, A., Henry, T., Margolis, J.J., Chan, K., and Monack, D.M. (2007). *In vivo* negative selection screen identifies genes required for *Francisella* virulence. Proc. Natl. Acad. Sci. U S A *104*, 6037–6042.

Woong Park, S., Klotzsche, M., Wilson, D.J., Boshoff, H.I., Eoh, H., Manjunatha, U., Blumenthal, A., Rhee, K., Barry, C.E., 3rd, Aldrich, C.C., et al. (2011). Evaluating the sensitivity of *Mycobacterium tuberculosis* to biotin deprivation using regulated gene expression. PLoS Pathog. *7*, e1002264.

Ye, H., Cai, M., Zhang, H., Li, Z., Wen, R., and Feng, Y. (2016). Functional definition of BirA suggests a biotin utilization pathway in the zoonotic pathogen *Streptococcus suis*. Sci. Rep. *6*, 26479.

Zhang, H., Wang, Q., Fisher, D.J., Cai, M., Chakravartty, V., Ye, H., Li, P., Solbiati, J.O., and Feng, Y. (2016). Deciphering a unique biotin scavenging pathway with redundant genes in the probiotic bacterium *Lactococcus lactis*. Sci. Rep. *6*, 25680.

Supplemental Information

Molecular Basis of BioJ, a Unique Gatekeeper

in Bacterial Biotin Synthesis

Wenhui Wei, Hongxin Guan, Tong Zhu, Sitao Zhang, Chengpeng Fan, Songying Ouyang, and Youjun Feng

Supplemental Information

Molecular Basis of BioJ, a Unique Gatekeeper in Bacterial Biotin Synthesis

Wenhui Wei, Hongxin Guan, Tong Zhu, Sitao Zhang, Chengpeng Fan, Songying Ouyang, Youjun Feng*
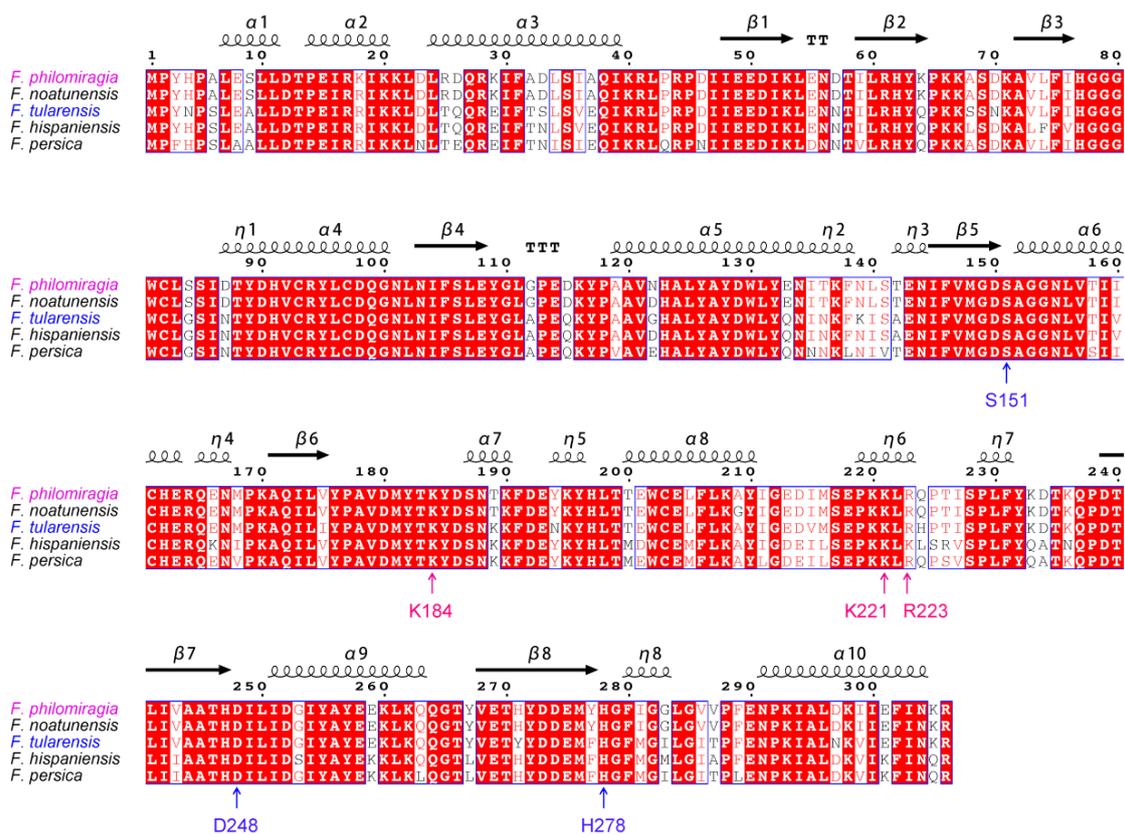
Figure S1 Sequence analyses of BioJ homologs from five *Francisella* species, related to **Figure 1**

Multiple sequence alignment of BioJ homologs proceeded with Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/), giving the final output in the form of ESPript (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi). Given the crystal structure of BioJ (PDB: 6K1T), protein secondary structure is illustrated in carton on the top. Three critical residues of catalytic triad (namely S151, D248, and H278) are consistently localized on the end of a certain β-sheet (namely β5, β7, and β8). Three putative ACP-interacting residues include K184, K221, and R223, respectively.

The accession numbers of five BioJ homologs separately refer to WP_004287811 for *F. philomiragia*; WP_088821195, *F. noatunensis*;

WP_003039032 for *F. tularensis*; WP_066046921 for *F. hispaniensis*; and WP_068594369 for *F. persica*.

Abbreviations: *F. philomiragia*, *Francisella philomiragia*; *F. noatunensis*, *Francisella noatunensis*; *F. tularensis*, *Francisella tularensis*; *F. hispaniensis*, *Francisella hispaniensis*; *F. persica*, *Francisella persica*; α, α-helices; β, β-sheet; η, coil.
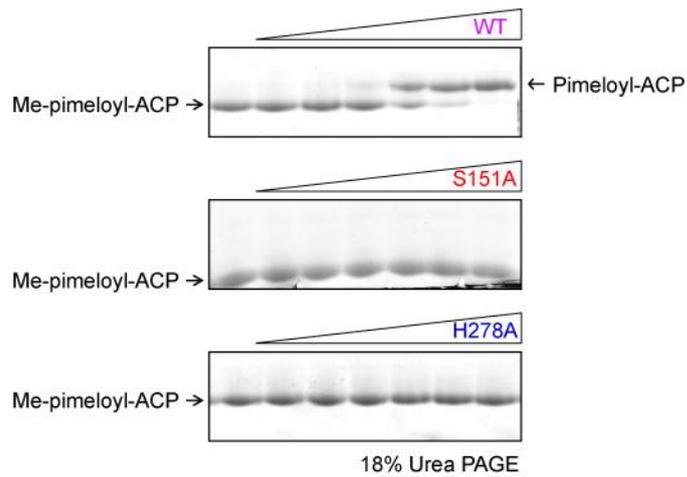
Figure S2 *In vitro* enzymatic assays for the point-mutants of BioJ with an inactivation in its catalytic triad, related to **Figure 2**

The reaction of BioJ-catalyzed hydrolysis is conducted as we earlier described with little change (Feng et al., 2014). The reaction products are separated with 18% urea PGAE identical to that of **Figs 2B-C**.

Designations: Me-pimeloyl-ACP, pimeloyl-ACP methyl ester.
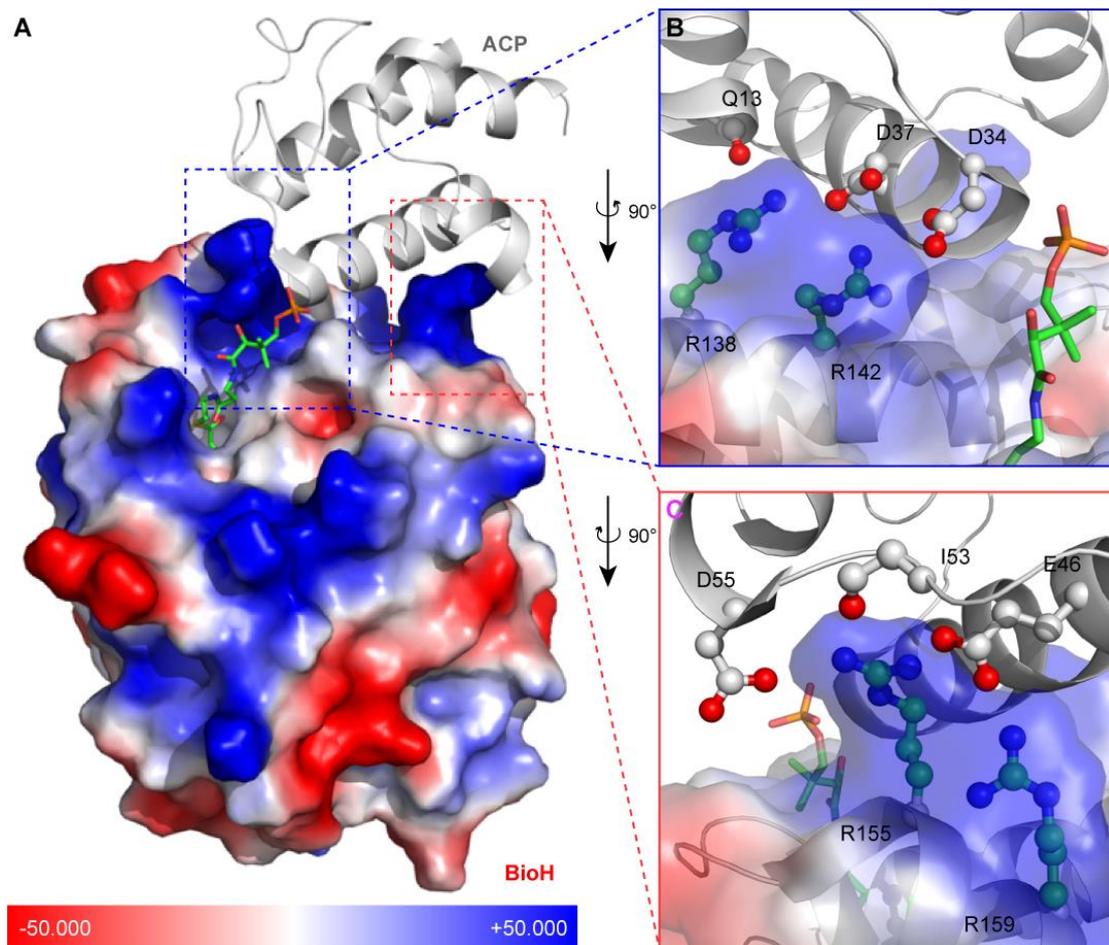
Figure S3 Interplay between BioH capping positive charge residues and ACP-α2 negative charge residues, related to **Figure 5**

A. Overall structure of BioH with its substrate Me-pimeloyl-ACP

The overall structure of BioH is shown in electrostatic potential surface, whereas the ACP (colored white) is in ribbon and methyl pimeloyl ester is highlighted in stick-balls. The left and right-sided regions marked respectively with blue and green frame represent ionic interactions between BioH and ACP.

B. An enlarged view of ionic interaction in left region

It is given through anti-clockwise 90° rotation of blue frame position (**Panel A**).

The two residues (R138 and R142) of BioH and three residues (Q13, D34 and D37) in ACP are implicated into ionic interactions.

C. Enlarged picture of ionic interaction in right region

It is generated through the clockwise 90° rotation of green frame position (**Panel A**). Ionic interactions are formed between ACP and BioH. ACP donates the side-chain carboxylate oxygen atoms from two acidic residues (D55 and E46) and main-chain amide carbonyl oxygen (I53), whereas BioH provides the side-chain nitrogen atoms of basic residues R155 and R159.
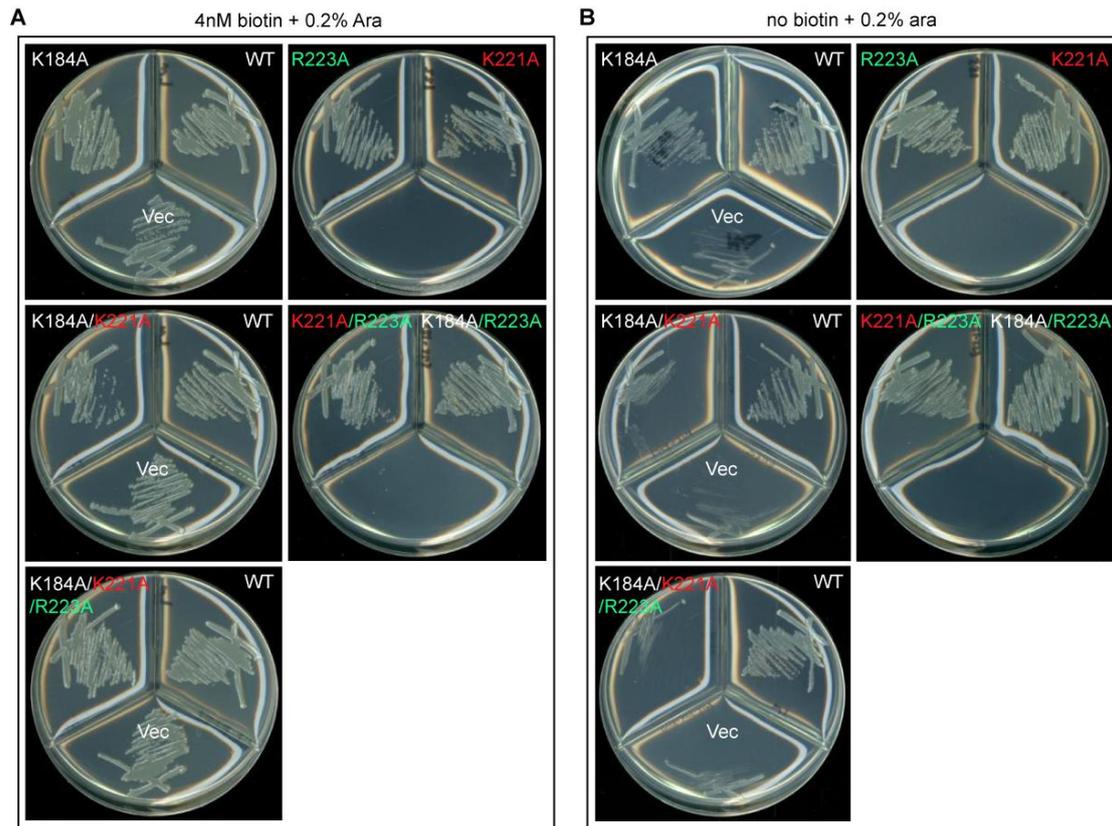
Figure S4 Activity of BioJ and its various mutants on M9 minimal agar plates, related to **Figure 5**

A. Biotin recovers growth phenotype of auxotrophic strains

The strains streaked in a minimal medium adding 4nM biotin can grow vigorous similar with wide type.

B. The activity of BioJ triple mutant reduced obviously on free-biotin medium

Alanine substitutions of residues (K184, K221 and R223) affect the enzymatic activity of BioH in the assay of colony formation on M9 minimal agar plates. The strain streaked in a sector is identified by the number. The single, double and triple mutants are labeled with brown, purple and red, respectively. Of note, the triple mutant significantly reduces bacterial growth on the biotin free condition.
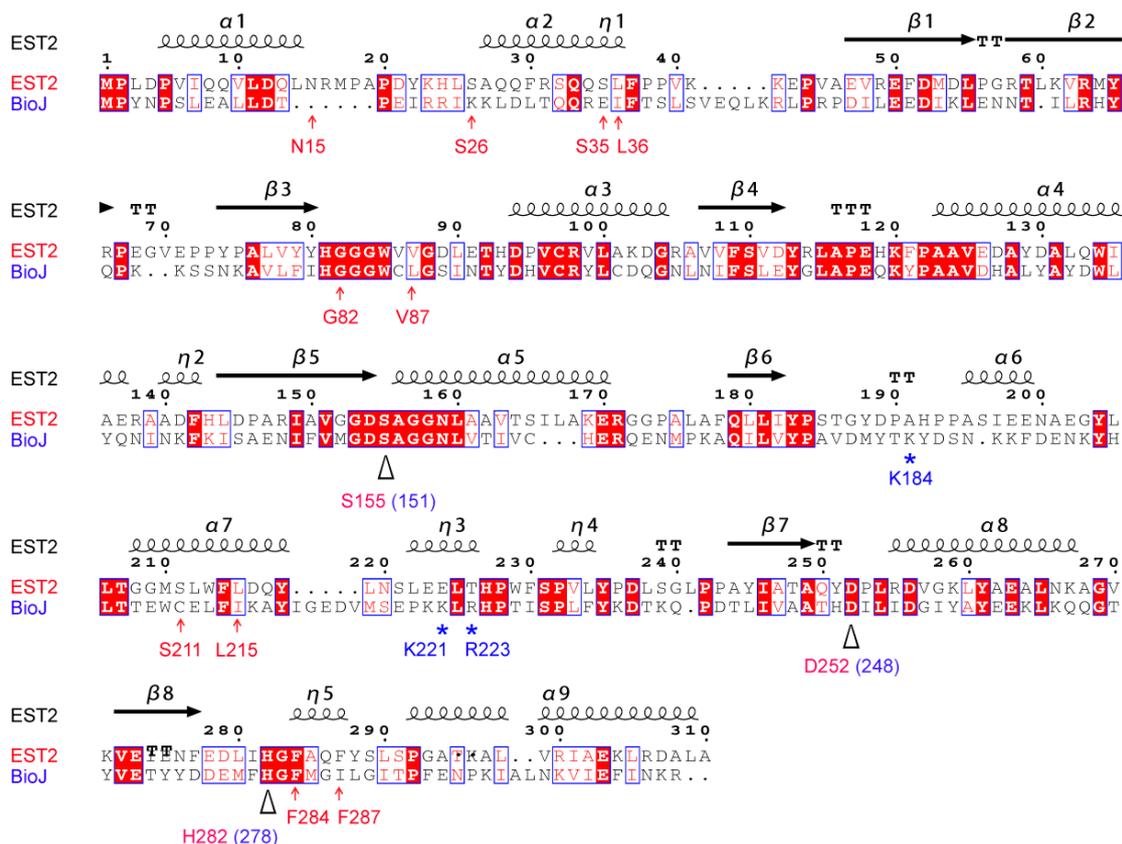
Figure S5 Sequence alignment of BioJ with EST2 of *Alicyclobacillus acidocaldarius*, related to **Figure 4**

EST2 (310aa) of *Alicyclobacillus acidocaldarius* (PDB: 1QZ3) (De Simone et al., 2004) was aligned with BioJ (306aa) using ClustalOmega (https://www.ebi.ac.uk/Tools/msa/clustalo/). The catalytic triad is indicated with dark triangles (S155, D252 & H282 in EST2; S151, D248 & H278 in BioJ). The long hydrophobic tunnel in EST2 for the entry of hexadecane moiety involves ten distinct critical residues (N15, S26, S35, L36, G82, V87, S211, L215, F284, and F287) are highlighted with red arrows. However, the three important residues (K184, K221 & R223) participating into the interaction between BioJ and its substrate pimeloyl-ACP methyl ester are underscored with asterisks. Though that the catalytic triad is conserved, the substrate-interacting pocket of BioJ seems to be far distinct from that of EST2.
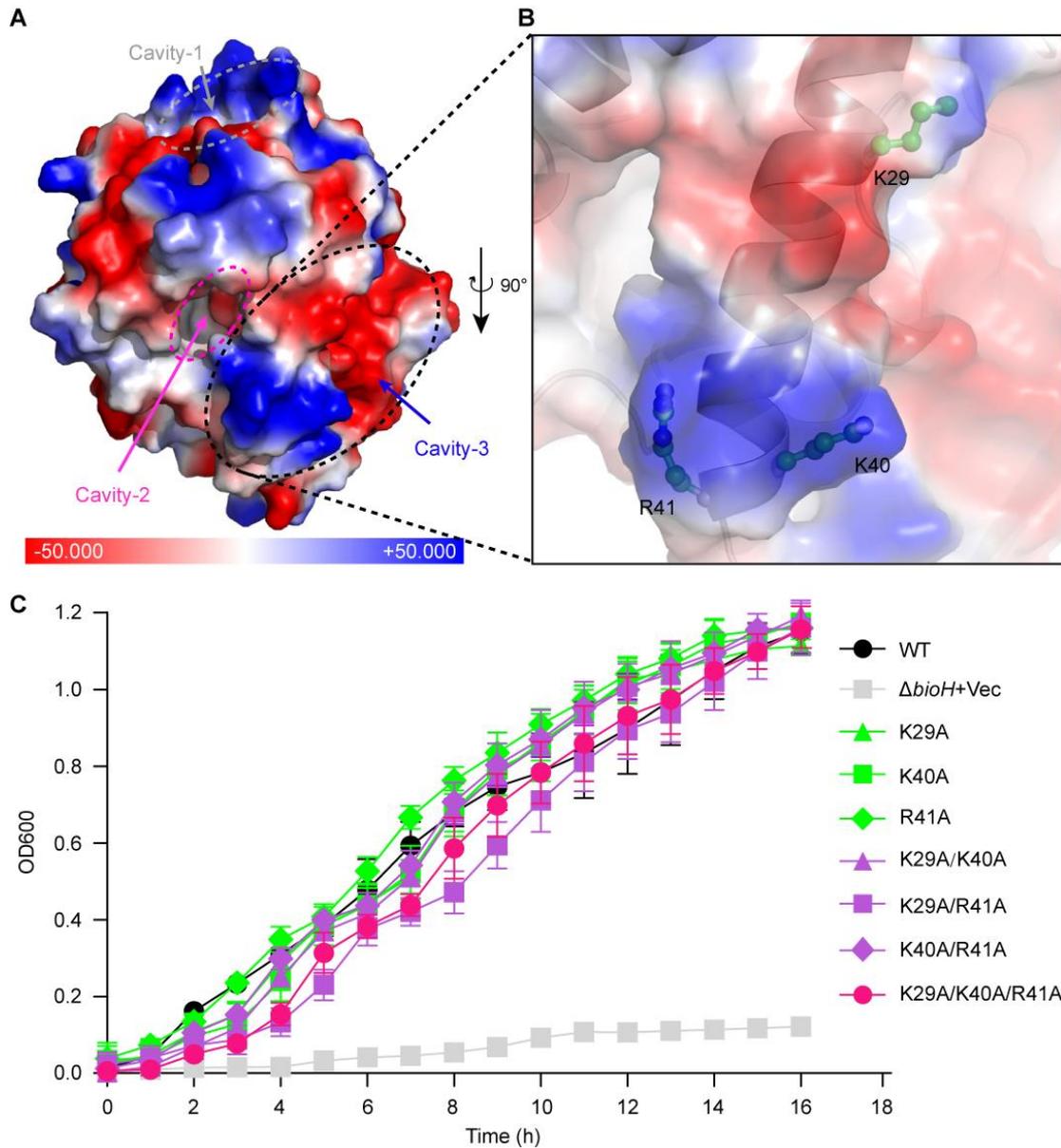
Figure S6 Discovery of a unique channel, Cavity-2, connected with the catalytic triad in BioJ, related to **Figure 5**

A. Electrostatic surface structure of BioJ harboring three unique cavities

It is given with PyMol. Indicated by the results of analysis of cavities with CAVER Analyst, Cavity-2 marked by pink circle has bigger channel (approximately 1.4-fold volume of channel-1) than the ACP-interacting Cavity-1 (circled with grey) does. These two cavities are connected with the catalytic triad. A small positively-charged area is located near Cavity-2.

B. An enlarged view of basic residues-forming Cavity 3 that is in the vicinity of Cavity 2

It is shown through the clockwise 90° rotation of **Panel A**. As we determined in Cavity-1, three basic residues (K29, K40, and R41) forecasted by the results of docking are chosen from Cavity 3. The side-chains are shown with stick-ball.

C. Site-directed mutagenesis of *bioJ* (K29A, K40A and R41A) suggests that this Cavity-3 merely affects BioJ function

Table S1 Bacteria and plasmids used in this study, related to **Figure 4**

Strain or plasmids	Relevant characteristics	Origins
Strains		
DH5α	A cloning host of *E. coli*	Lab stock
STL24	MG1655, Δ*bioH*	(Lin et al., 2010)
FYJ391	STL24 carrying pBAD322	(Feng et al., 2014)
FYJ392	STL24 carrying pBAD322::*bioJ*	(Feng et al., 2014)
FYJ394	BL21 (DE3) carrying pET28a::*bioJ*	(Feng et al., 2014)
FYJ398	STL24 carrying pBAD322::*bioJ* (S151A)	(Feng et al., 2014)
FYJ405	STL24 carrying pBAD322::*bioJ* (D248A)	(Feng et al., 2014)
FYJ406	STL24 carrying pBAD322::*bioJ* (H278A)	(Feng et al., 2014)
FYJ1513	STL24 carrying pBAD322::*bioJ* (K184A)	Lab stock
FYJ1514	STL24 carrying pBAD322::*bioJ* (K221A)	Lab stock
FYJ1515	STL24 carrying pBAD322::*bioJ* (R223A)	This work
FYJ1516	STL24 carrying pBAD322::*bioJ* (K181A/K221A)	This work
FYJ1517	STL24 carrying pBAD322::*bioJ* (K181A/R223A)	This work
FYJ1518	STL24 carrying pBAD322::*bioJ* (K221A/R223A)	This work
FYJ1519	STL24 carrying pBAD322::*bioJ* (K184A/K221A/R223A)	This work
FYJ1520	STL24 carrying pBAD322::*bioJ* (K29A)	This work
FYJ1521	STL24 carrying pBAD322::*bioJ* (K40A)	This work
FYJ1522	STL24 carrying pBAD322::*bioJ* (R41A)	This work
FYJ1523	STL24 carrying pBAD322::*bioJ* (K29A/K40A)	This work
FYJ1524	STL24 carrying pBAD322::*bioJ* (K29A/R41A)	This work
FYJ1525	STL24 carrying pBAD322::*bioJ* (K40A/R41A)	This work
FYJ1526	STL24 carrying pBAD322::*bioJ* (K29A/K40A/R41A)	This work
Plasmids		
pBAD322	An arabinose-inducible expression vector, Amp^R	(Cronan, 2006)
pET28a	An IPTG-inducible T7-driven	Novagen

	expression vector, Km^R	
pBAD322::*bioJ*	pBAD322 encoding *bioJ*, Amp^R	(Feng et al., 2014)
pET28a::*bioJ*	pET28a encoding *bioJ*, Km^R	(Feng et al., 2014)
pBAD322::*bioJ* (S151A)	pBAD322 encoding the mutant version of *bioJ* (S151A), Amp^R	(Feng et al., 2014)
pBAD322::*bioJ* (D248A)	pBAD322 encoding the mutant version of *bioJ* (D248A), Amp^R	(Feng et al., 2014)
pBAD322::*bioJ* (H278A)	pBAD322 encoding the mutant version of *bioJ* (H278A), Amp^R	(Feng et al., 2014)
pBAD322::*bioJ* (K184A)	pBAD322 encoding the mutant version of *bioJ* (K184A), Amp^R	This work
pBAD322::*bioJ* (K221A)	pBAD322 encoding the mutant version of *bioJ* (K221A), Amp^R	This work
pBAD322::*bioJ* (R223A)	pBAD322 encoding the mutant version of *bioJ* (R223A), Amp^R	This work
pBAD322::*bioJ* (K184A/K221A)	pBAD322 encoding the mutant version of *bioJ* (K184A/K221A), Amp^R	This work
pBAD322::*bioJ* (K184A/R223A)	pBAD322 encoding the mutant version of *bioJ* (K184A/R223A), Amp^R	This work
pBAD322::*bioJ* (K221A/R223A)	pBAD322 encoding the *bioJ* mutant (K221/R223A), Amp^R	This work
pBAD322::*bioJ* (K183A/K221A/R 223A)	pBAD322 encoding the *bioJ* mutant (K184A/K221A/R223A), Amp^R	This work
pBAD322::*bioJ* (K29A)	pBAD322 encoding the *bioJ* mutant (K29A), Amp^R	This work
pBAD322::*bioJ* (K40A)	pBAD322 encoding the *bioJ* mutant (K40A), Amp^R	This work
pBAD322::*bioJ* (R41A)	pBAD322 encoding the *bioJ* mutant (R41A), Amp^R	This work
pBAD322::*bioJ* (K29A-K40A)	pBAD322 encoding the *bioJ* mutant (K29A/K40A), Amp^R	This work
pBAD322::*bioJ* (K29A/R41A)	pBAD322 encoding the *bioJ* mutant (K29A/R41A), Amp^R	This work
pBAD322::*bioJ* (K40A/R41A)	pBAD322 encoding the *bioJ* mutant (K40A/R41A), Amp^R	This work
pBAD322::*bioJ* (K29/K40A/R41A)	pBAD322 encoding the *bioJ* mutant (K29A/K40A/R41A), Amp^R	This work

Table S2 Primers used in this study, related to **Figures 4** and **5**

Primers	Sequences
bioJ-K184A-F	5'-CGA TAT GTA TAC TGC ATA TGA TAG TAA TAC TA-3'
bioJ-K184A-R	5'-TAG TAT TAC TAT CAT ATG CAG TAT ACA TAT CG-3'
bioJ-K221A-F	5'-GCC TAA AGC ACT TAG ACA ACC TAC AAT ATC TCC ATT ATT TT-3'
bioJ-K221A-R	5'-GTC TAA GTG CTT TAG GCT CTG ACA TGA TAT CTT CG-3'
bioJ-K233-F	5'-GCA GAT ACA AAT CAA CCT GAT ACT CTA ATT GTA G-3'
bioJ-K233-R	5'-GGT TGA TTT GTA TCT GCA TAA AAT AAT GGA GAT ATT GTA GGT TGT CTA A-3'
bioJ-K221A/K223A-F	5'-AAA GCA CTT GCA CAA CCT ACA ATA TCT CCA TTA TTT TAT AAA G-3'
bioJ-K221A/K223A-R	5'-GGT TGT GCA AGT GCT TTA GGC TCT GAC ATG ATA TCT TCG-3'
bioJ-K29A-F	5'-CAG CGT GCA ATA TTT GCA GAT CTT TCA ATA GCA CA-3'
bioJ-K29A-R	5'-GCA AAT ATT GCA CGC TGA TCT CTT AAA TCT AAT TTT TTT-3'
bioJ-K40A-F	5'-ACA AAT CGC ACG CCT ACC TCG CCC TGA TAT TA-3'
bioJ-K40A-R	5'-GTA GGC GTG CGA TTT GTG CTA TTG AAA GAT CTG CA-3'
bioJ-R41A-F	5'-AAT CAA GGC ACT ACC TCG CCC TGA TAT TAT CGA-3'
bioJ-R41A-R	5'-GAG GTA GTG CCT TGA TTT GTG CTA TTG AAA GAT CTG-3'
bioJ-K41A/R41A-F	5'-AAT CGC AGC ACT ACC TCG CCC TGA TAT TAT CGA-3'
bioJ-K41A/R41A-R	5'-GAG GTA GTG CTG CGA TTT GTG CTA TTG AAA GAT C-3'

Transparent Methods

Bacterial Strains, Plasmids and Growth Condition

The bacterial strains used in this study were *E. coli* K-12 derivatives (**Table S1**), and maintained in Luria-Bertani (LB) liquid broth and/or on solid LB agar (LBA) plates at 37°C. As for biotin-sensitive assays, the biotin-free M9 minimal media contained 0.1% vitamin-free Casamino Acids and 0.2% (wt/vol) glycerol as sole carbon source (Feng et al., 2014). If required, the antibiotics (like ampicillin, 100μg/ml & kanamycin, 50μg/ml) were supplemented. The *bioJ* of *Francisella* was re-cloned into the pET28a-SUMO vector with an N-terminal 6×His-SUMO-tag. Using the Mut Express II Fast Mutagenesis Kit V2 (Vazyme Biotech Co., Ltd), site-directed mutagenesis of *bioJ* was routinely conducted, in which the plasmid pBAD322*::bioJ* (**Table S1**) acted as a template, together with appropriate primer pairs (**Table S2**). The resultant recombinant plasmids like pBAD322*::bioJ* (K29/K40A/R41A), were validated with both PCR detection and direct DNA sequencing. In the pET28-based protein production system, 0.2mM of Isopropyl β-D-thiogalactoside (IPTG) was utilized. In general, 0.2%(vol/vol) arabinose was added to trigger expression of pBAD24-borne *bioJ* and its mutants (**Table S1**).

Experiments of Genetic Complementation

The STL24 strain, an *E. coli* Δ*bioH* mutant (**Table S1**), functioned as a recipient for assessing the physiological role of pBAD24 plasmid-borne *bioJ* and its variants (**Table S1**). Totally, point-mutants of *bioJ* covered both amino acids critical for either catalytic triad (S151, D248 & H278, in **Fig. S1**) and putative ACP-interacting residues (namely K184, K221 & K223, in **Fig. S1**). The strains of interest were stripped on biotin-free M9 minimal plates and incubated at 37°C overnight for photographing. To plot growth curves, all these engineered strains were sub-cultured (1:200) into liquid M9 minimal media lacking biotin. In this case, the 96-well cultivation plates were kept at 37°C, each well of which contains 200μl of M9 media. The bacterial optical density at

wave-length of 600nm (OD600) was monitored in real-time using BMG SPECTROstar Nano (LRBTECH). A regular interval was set as 1h within the whole monitoring period of 16h (Feng et al., 2014).

Protein Expression and Purification

The *E. coli* BL21 (DE3) cells carrying pET28a-SUMO*::bioJ* were induced with the addition of 0.2 mM IPTG when the optical density (OD600) reached 0.8, and then kept at 18°C for 18 hrs. Bacterial cultures were harvested through the centrifugation at 5,000 rpm for 15 min at 4°C, and subsequently resuspended in the lysis buffer [50 mM Tris-HCl (pH 8.0) and 500 mM NaCl]. Following the lysis by ultrasonication, the cell lysate was centrifuged at 17,000 rpm for 30 min at 4°C. Then, the BioJ protein was purified as we recently described (Jiang et al., 2017) with minor adjustment. Briefly, the supernatant was loaded onto a. After the removal of all the protein contaminants from Ni-NTA column (Qiagen) with extensive wash, the resin-bound BioJ was incubated with Ulp1 peptidase overnight to remove the SUMO tag. The target BioJ protein without tag was eluted, concentrated, and then purified to homogeneity by size-exclusion chromatography. The gel filtration was conducted with Superdex 200 increase column (GE Healthcare) equilibrated with the running buffer (25 mM Tris-HCl (pH 8.0), 100 mM NaCl and 2 mM DTT).

Enzymatic Assay for BioJ

The in vitro enzymatic activities of BioJ and its derivatives (e.g., S151A) were determined as we recently described (Feng et al., 2014). The mixture of reaction (20μl) consists of 50 mM HEPES buffer (pH 7.0), 5% glycerol, 150 μM acyl-ACP methyl esters (C6, C7 & C8), and 5 nM BioJ. The reaction was kept at 37°C for 1 h. Conformation-sensitive urea (2.5M)-PAGE (20%) was used to separate the products from the reactants (130 V, 2.5 h).

Biotin Bioassay

As earlier described by Lin *et al.* (Lin et al., 2010), the system for DTB biosynthesis was reconstituted *in vitro.* The use of dialysis allowed us to prepare the cell-free Δ*bioH* crude extract, which was applied in the catalysis of DTB synthesis. The purified BioJ protein was added into the crude extract of strain STL24 (*E. coli* Δ*bioH*) (Feng et al., 2014). An indicator strain of this DTB synthesis system referred to the biotin auxotrophic strain ER90 (Δ*bioF* Δ*bioC* Δ*bioD*) (Lin et al., 2010). The supply of biotin/DTB provided by the *in vitro* system restored bacterial growth of the auxotrophic strain ER90. It was illustrated by the formation of insoluble red formazan in the reduction of 2,3,5-triphenyl tetrazolium chloride (TTC).

In brief, the system of enzymatic reaction (100 µl) comprised 1 mg of Δ*bioH* cell-free crude extract protein, 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 0.1 mM pyridoxal-5′-phosphate (PLP), 1 mM L-alanine, 1 mM $KHPO_4$, 1 mM NADPH, 1 mM ATP, 1 mM glucose 6-phosphate (G6P), 1 mM S-adenosy-L-methionine, 60 µM pimeloyl-ACP methyl ester (Me-Pim-ACP) and 0.5 µM (~15 µg/ml) BioJ protein. The mixture of reaction was initiated at 37°C for ~3hrs and then quenched by the immersion in boiling water for 15 min.

Crystallization, Data Collection and Structural Determination

The purified BioJ (~15 mg/ml) was subjected to crystallization screens with the hanging drop vapor diffusion method. Each drop contained 0.5 µl of the protein solution together with 0.5 µl of reservoir solution, and kept at 16°C. As a result, diffraction-quality crystals were obtained in 20% (v/v) PEG 6000 and 100 mM BICINE/sodium hydroxide (pH 9.0). Crystals were harvested and stored in the abovementioned buffer supplemented with 15% (v/v) glycerol as a cryoprotectant before flash freezing. X-ray diffraction data were collected on beamline BL-17U1 at Shanghai Synchrotron Radiation Facility (SSRF).

Diffraction images were processed with the HKL-2000 program (Otwinowski and Minor, 1997). Further data processing was performed in the CCP4 suite (Winn et al., 2011) (**Table 1**). The phase was identified using principle of molecular replacement (MR). The first 46 amino acids from the N-terminal of EST2 were removed from the structure and the remainder structure was used as the search model for MR (PDB code: 1QZ3). The LLG score and Z-score of the solution are 107 and 10.1, respectively. Model building and crystallographic refinement were conducted with COOT (Emsley et al., 2010) and PHENIX (Adams et al., 2010). The interactions were analyzed with PyMol (http://www.pymol.org/) and PDBsum (Laskowski, 2007). The RMSD values were calculated with Chimera (Pettersen et al., 2004).

Molecular Docking

Initially, the crystal structure was analyzed with CAVER Analyst 2.0 (Jurcik et al., 2018). The analyses of cavities were performed using default parameters, and the volume of each cavity was calculated. The largest two cavities were selected as the sites for the subsequent docking process. Unlike the structural architectures of BioH (PDB: 1M33 & 4NMW) and BioG (PDB: 5H3B & 5GNG) alone (Sanishvili et al., 2003; Shi et al., 2016), the crystal structure of *E. coli* BioH is complexed with pimeloyl-ACP methyl ester (PDB: 4ETW) (Agarwal et al., 2012). The latter (PDB: 4ETW) was selected as the initial template for structural modelling and molecular docking. The structures of Me-pimeloyl-ACP and BioJ were initially energy-minimized using YASARA (Krieger and Vriend, 2014, 2015) with Yamber3 force field (Krieger et al., 2004), including 2,000 steps of steepest decent minimization and 3,000 steps of conjugate gradient minimization. Then Me-pimeloyl-ACP was docked into BioJ structure using YASARA local docking program that adopted the AutoDock VINA algorithm (Trott and Olson, 2010). In light of the binding mode of BioH to Me-pimeloyl-ACP, the simulation cell was fixed around the hydrophobic pocket and the lid domain of BioJ, with box size 60 Å in all three dimensions. For each

of the two cavities, 900 docking runs were performed in total, and the results were clustered based on the root-mean-square deviations (RMSD) against 4ETW. Obtained conformations were subsequently ordered by binding energy, and the conformation of the complex with the lowest free energy of binding was chosen to present the structural information of docking results. Other parameters involved in the process above were set as default parameters if they are not mentioned.

Supplemental References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

Agarwal, V., Lin, S., Lukk, T., Nair, S.K., and Cronan, J.E. (2012). Structure of the enzyme-acyl carrier protein (ACP) substrate gatekeeper complex required for biotin synthesis. Proc Natl Acad Sci USA *109*, 17406-17411.

Cronan, J.E. (2006). A family of arabinose-inducible *Escherichia coli* expression vectors having pBR322 copy control. Plasmid *55*, 152-157.

De Simone, G., Mandrich, L., Menchise, V., Giordano, V., Febbraio, F., Rossi, M., Pedone, C., and Manco, G. (2004). A substrate-induced switch in the reaction mechanism of a thermophilic esterase: kinetic evidences and structural basis. The Journal of biological chemistry *279*, 6815-6823.

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

Feng, Y., Napier, B.A., Manandhar, M., Henke, S.K., Weiss, D.S., and Cronan, J.E. (2014). A *Francisella* virulence factor catalyzes an essential reaction of biotin synthesis. Molecular microbiology *91*, 300-314.

Jiang, Y., Zhu, Y., Qiu, W., Liu, Y.J., Cheng, G., Liu, Z.J., and Ouyang, S. (2017). Structural and functional analyses of human DDX41 DEAD domain. Protein Cell *8*, 72-76.

Jurcik, A., Bednar, D., Byska, J., Marques, S.M., Furmanova, K., Daniel, L., Kokkonen, P., Brezovsky, J., Strnad, O., Stourac, J.*, et al.* (2018). CAVER Analyst 2.0: analysis and visualization of channels and tunnels in protein structures and molecular dynamics trajectories. Bioinformatics *34*, 3586-3588.

Krieger, E., Darden, T., Nabuurs, S.B., Finkelstein, A., and Vriend, G. (2004). Making optimal use of empirical energy functions: force-field parameterization in crystal space. Proteins *57*, 678-683.

Krieger, E., and Vriend, G. (2014). YASARA View - molecular graphics for all devices - from smartphones to workstations. Bioinformatics *30*, 2981-2982.

Krieger, E., and Vriend, G. (2015). New ways to boost molecular dynamics simulations. Journal of computational chemistry *36*, 996-1007.

Laskowski, R.A. (2007). Enhancing the functional annotation of PDB structures in PDBsum using key figures extracted from the literature. Bioinformatics *23*, 1824-1827.

Lin, S., Hanson, R.E., and Cronan, J.E. (2010). Biotin synthesis begins by hijacking the fatty acid synthetic pathway. Nat Chem Biol *6*, 682-688.

Otwinowski, Z., and Minor, W. (1997). [20] Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol *276*, 307-326.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. Journal of computational chemistry *25*, 1605-1612.

Sanishvili, R., Yakunin, A.F., Laskowski, R.A., Skarina, T., Evdokimova, E., Doherty-Kirby, A., Lajoie, G.A., Thornton, J.M., Arrowsmith, C.H., Savchenko, A.*, et al.* (2003). Integrating structure, bioinformatics, and enzymology to discover function: BioH, a new carboxylesterase from *Escherichia coli*. The Journal of biological chemistry *278*, 26039-26045.

Shi, J., Cao, X., Chen, Y., Cronan, J.E., and Guo, Z. (2016). An atypical α/β-hydrolase fold revealed in the crystal structure of pimeloyl-acyl carrier protein methyl esterase BioG from *Haemophilus influenzae*. Biochemistry *55*, 6705-6717.

Trott, O., and Olson, A.J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry *31*, 455-461.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A*., et al.* (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr *67*, 235-242.