Published in final edited form as:

ACS Chem Biol. 2015 December 18; 10(12): 2785–2793. doi:10.1021/acscchembio.5b00641.

Bacterial genome mining of enzymatic tools for alkyne biosynthesis

Xuejun Zhu¹, Michael Su¹, Kadirvel Manickam², and Wenjun Zhang^{1,3,*}

¹Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA

²Department of Bioengineering, University of California, Berkeley, CA 94720, USA

³Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Abstract

The alkyne is an important functionality widely used in material science, pharmaceutical science, and chemical biology, but the importance of this functionality is contrasted by the very limited number of enzymes known to be involved in alkyne biosynthesis. We recently reported the first known carrier protein-dependent pathway for terminal alkyne formation, and *in silico* analysis suggested that this mechanism could be widespread in bacteria. In this paper, we screened additional homologous gene cassettes presumed to be involved in alkyne biosynthesis using both *in vitro* biochemical study and an *E. coli*-polyketide synthase (PKS) reporting system for *in vivo* analysis. We discovered and characterized a new terminal alkyne biosynthetic pathway comprised of TtuA, B, and C from *Teredinibacter turnerae* T7901. While the acyl-CoA ligase homolog (TtuA) demonstrated promiscuity in the activation and loading of medium-chain fatty acids onto the carrier protein (TtuC), the desaturase homolog (TtuB) showed stringent substrate specificity towards C₁₀ fatty acyl moieties. In addition, TtuB was demonstrated to be a bifunctional desaturase/acetylenase that efficiently catalyzed two sequential O₂-dependent dehydrogenation reactions. A novel terminal-alkyne bearing polyketide was further produced upon co-expression of *ttuABC* and a PKS gene in *E. coli*. The discovery and characterization of TtuA, B, and C provides us with a new bifunctional desaturase/acetylenase for mechanistic and structural study and expands the scarce enzyme inventory for the biosynthesis of the alkyne functionality, which has important applications in synthetic and chemical biology.

Introduction

The alkyne is a readily derivatized functionality found in molecules that are widely used in chemical synthesis, pharmaceuticals, and materials^{1, 2}. It is a particularly useful moiety in azide-alkyne bioorthogonal chemistry, which has emerged as one of the most powerful tools

*wjzhang@berkeley.edu.

Supporting Information

Primers, plasmids, strains, gene sequences, SDS-PAGE analysis of purified proteins, TtuB topology, selected mass chromatograms and compound characterizations. This material is available free of charge via the Internet at <http://pubs.acs.org>.

in chemical biology^{3–5}. For example, tagging biomolecules with an alkyne functionality and coupling them with bioorthogonal reactions have allowed the imaging and mode of action study of biologically important molecules^{4–7}. Despite the importance of the alkyne functionality and its prevalence in synthetic molecules, the biological routes to alkynes are poorly understood. Acetylenases are the enzymes responsible for alkyne biosynthesis, and they are categorized as a special family of membrane-bound desaturases that catalyze the abstraction of hydrogen atoms from C-C bonds with the use of a di-iron active site^{8, 9}. However, the notorious challenges in working with membrane-bound desaturases have impeded the discovery and mechanistic understanding of this family of enzymes. To date, only a limited number of acetylenases have been identified, resulting in a shortage of alkyne biosynthetic tools^{10–16}.

We recently reported the study of JamB, a desaturase/acetylenase that is essential for the formation of a terminal alkyne functionality found in the cyanobacterial natural product jamaicamide B (Figure 1A)^{16, 17}. *JamB* is embedded in a tri-gene cassette *jamABC* that encodes an acyl-acyl carrier protein (ACP) synthetase, a membrane-bound desaturase/acetylenase, and an ACP, respectively. These three proteins employ an ACP-dependent pathway to generate the terminal alkyne functionality: JamA activates and loads 5-hexenoic acid onto JamC¹⁸, and the resulting 5-hexenoyl-JamC is modified by JamB to yield 5-hexenoyl-JamC as a starter unit for the downstream polyketide synthase/non-ribosomal peptide synthetase (PKS/NRPS) assembly line. The specific recognition of the ACP-bound substrate by JamB explains the necessity of the co-localization of *jamA* and *jamC* with *jamB* in the same operon. In addition, more than 80 gene operons homologous to *jamABC*, including the ones from the carmabin and caryoynencin biosynthetic gene clusters (Figure 1A), have been identified across diverse bacterial genera through genome-mining^{14, 16, 19}, indicating that this ACP-dependent pathway is likely adopted in the synthesis of a variety of alkyne-bearing natural products. Further *in silico* analysis revealed that a number of *jamABC* homologs are clustered with genes encoding PKSs/NRPSs, suggesting their possible involvement in generating alkynes residing in polyketide/non-ribosomal peptide (PK/NRP) molecular scaffolds; and some other operons encode multiple desaturases and are probably responsible for polyene biosynthesis (Figure 1B)^{14, 15}. Yet the majority of these gene clusters have no known associated metabolites. The study of these gene operons homologous to *jamABC* will thus facilitate the discovery of a variety of alkyne-bearing natural products and lead to the expansion of the alkyne biosynthetic toolbox capable of producing acetylenic groups with altered substrate specificities and improved efficiency.

In this paper, we report the screening of the activities of selected JamA, B, and C homologs using both *in vitro* biochemical assays and *in vivo* heterologous reconstitution (Figure 2). We previously demonstrated that a carrier protein-bound fatty acyl moiety could efficiently serve as the starter unit for a promiscuous type III PKS, and as a result, a terminal alkyne-tagged polyketide was produced in *E. coli* during the co-expression of *jamABC* and *hspksI*¹⁶. These results suggested that the *E. coli*-PKS platform could be a useful *in vivo* reporting system for probing the activities of JamA, B, and C homologs. In this work, we first determined the substrate preference of JamA homologs towards fatty acids of different chain lengths. We then extended the *E. coli* platform to include a variety of type III PKSs

with the corresponding chain length specificity of the starter unit to reconstitute the activities of JamA, B, and C homologs *in vivo*. Our screening efforts resulted in the discovery of new terminal alkyne biosynthetic machinery from *Teredinibacter turnerae* T7901 that specifically recognized C₁₀ fatty acids and led to the production of a novel terminal alkyne-bearing polyketide in *E. coli* with high efficiency.

Results and Discussion

Eleven gene operons homologous to *jamABC* were targeted for screening

A phylogenetic analysis was performed on JamB and its homologs encoded by the genes from finished or permanent draft of sequenced bacterial genomes (Figure 3). The results revealed that JamB homologs are widespread in bacteria, including cyanobacteria, proteobacteria, actinobacteria, and planctomycetes. Among these JamB homologs, we selected 11 representative desaturases from different clades for functional investigation. The *jamABC* homologous operon from *Pseudomonas fluorescens* Pf-5 was cloned from its genomic DNA while the other 10 operons were obtained through gene synthesis.

Biochemical analysis of JamA and C homologs for activation and loading of fatty acids of different chain lengths

To probe the abilities of acyl-ACP synthetases to activate and load various fatty acids onto ACPs, we individually cloned *jamA* and *jamC* homologs into an expression vector encoding a C-terminal His₆-tag. The recombinant proteins were overproduced in *E. coli* and purified using Ni-NTA affinity chromatography. The *E. coli* BAP1 strain containing a chromosomal copy of the phosphopantetheinyl transferase Sfp was used for the expression of *jamC* homologs to ensure their posttranslational modification to the pantetheinylated forms²⁰. For the JamA homologs that were not soluble in *E. coli*, a maltose-binding protein (MBP) tag was fused onto the N-terminus to possibly promote the folding process²¹. Seven out of eleven pairs of JamA and JamC homologs were successfully purified from *E. coli* as soluble proteins (Supplementary Figure 1). The substrate selectivity of JamA homologs was then tested against fatty acids of different chain lengths ranging from C₆ to C₁₈ using the ACP loading assay, followed by liquid chromatography-high-resolution mass spectrometry (LC-HRMS) analysis (Figure 4A and Supplementary Figure 2). Based on the fatty acid chain length preferences, JamA homologs were categorized into two groups. Group A showed high specificity towards fatty acids of mediumchain length (C₆–C₁₂) and included the ones from *Burkholderia pseudomallei* K96243, *Teredinibacter turnerae* T7901, *Schlesneria paludicola* DSM 18645, and *Cyanothece* sp. ATCC 51142. Group B was composed of the homologs from *Pseudomonas entomophila* L48, *Pseudomonas fluorescens* Pf-5, and *Inquilinus limosus* DSM 16000, which were specific for both medium- and long-chain fatty acids (C₈–C₁₈). Bioinformatics analysis of these JamA homologs showed a close phylogenetic relationship in each group, consistent with our *in vitro* biochemical results (Figure 4B).

***In vivo* reconstitution of the activities of JamA, B, and C homologs using an *E. coli*-PKS reporting system**

We previously reported that a promiscuous type III PKS can utilize a carrier proteinbound fatty acyl starter unit to yield a terminal alkyne-tagged polyketide in *E. coli* during the coexpression of *jamABC* and *hspks1*¹⁶. This *E. coli*-PKS platform could thus be feasible as an *in vivo* reporting system for probing the activities of JamB homologs with unknown substrate preferences. We first extended the *E. coli*-PKS platform by including additional promiscuous type III PKSs that can accept acyl-ACPs of altered fatty acyl chain lengths as starter units. HsPKS1 is known to accept medium-chain fatty acyl-ACPs (C₆–C₁₁)^{16, 22}, and it was thus used for the study of the desaturases that pair with JamA homologs that activate medium-chain fatty acids. To study JamB homologs that may recognize long-chain fatty acyl-ACP substrates, we searched for type III PKSs with activities towards long-chain fatty acyl-ACP starter units. We tested the *in vivo* activities of three type III PKSs that have been reported to take long-chain fatty acyl-CoAs *in vitro*, including ORAS from *Neurospora crassa*, PKS18 from *Mycobacterium tuberculosis* and CsyA from *Aspergillus oryzae*^{23–26}. Upon overexpression of the individual PKSs with the feeding of long-chain fatty acids, all three of the type III PKSs enabled the efficient production of polyketides in *E. coli* with fatty acyl starter units ranging in chain length from 10 to 16 carbons, while PKS18 and CsyA also showed the incorporation of C₁₈ fatty acyl starter units (Supplementary Figure 3). The substrate tolerance of these type III PKSs towards ACPs was also investigated *in vitro*. Several lauryl-ACPs were prepared *in situ* using enzymatic reactions containing lauric acid, acyl-ACP synthetases and their corresponding ACPs, and were tested as substrates for the type III PKSs. While CsyA poorly recognized these acyl-ACPs, ORAS and PKS18 could efficiently utilize all of the lauryl-ACPs to yield the expected product. This product identity was further confirmed by comparison to the compound synthesized using enzymatic reactions containing lauryl-CoA, malonyl-CoA and the type III PKSs as positive controls (Supplementary Figure 4). Consequently, ORAS and PKS18 were selected to replace the role of HsPKS1 in the *in vivo* reporting system to study the activities of JamB homologs that may recognize long-chain fatty acyl-ACP substrates.

We next built the *E. coli* platform for the reconstitution of the activities of JamA, B, and C homologs using the promiscuous type III PKSs HsPKS1, ORAS, or PKS18. *E. coli* BAP1 was transformed with two plasmids encoding *hspks1/oras/pks18* and *jamABC* homologs, with the expression of each gene under a T7 promoter. In parallel, we omitted the *jamB* homolog to yield a control strain. The resulting *E. coli* strains were induced with IPTG and grown in shake flasks at 20 °C for two days with the supplementation of selected fatty acids (Supplementary Table 4). LC-UV-MS and LC-HRMS analyses of these culture extracts revealed the production of a putative alkyne-bearing polyketide **1** by the strains XZ2, XZ3, XZ4 that contained *ttuABC* from *T. turnerae* T7901 and *hspks1/oras/pks18*, respectively. (Figure 5, Supplementary Table 4, and Supplementary Figures 5–7 and 11). Compound **1** was produced with a titer of ~0.5 mg L⁻¹ when the cultures of XZ2 were fed with 9-decenoic acid, and omitting *ttuB* completely abolished the production of **1** (Figure 5B, Supplementary Table 4, and Supplementary Figures 6–7). The current titer of compound **1** was comparable to the typical polyketide titer in *E. coli* under similar culture conditions²⁷. Product **1** was further purified from large-scale cultures of XZ2 grown in F1 minimal

medium, and the molecular structure of **1** was confirmed by NMR analysis (Supplementary Table 5 and Supplementary Figures 14–18). In particular, the presence of the terminal alkyne functionality was demonstrated by ^1H ($\delta_{\text{H-14}} = 2.73$) and ^{13}C NMR ($\delta_{\text{C-13}} = 84.5$, $\delta_{\text{C-14}} = 71.2$), together with strong ^1H , ^{13}C heteronuclear single quantum correlations (^1H , ^{13}C HSQC) of H-14/C-14 and strong ^1H , ^{13}C heteronuclear multiple-bond correlations (^1H , ^{13}C HMBC) of H-14/C-13, H-14/C-12, H-12/C-13, H-12/C-14, and H-11/C-13. Compound **1** was presumably assembled by the type III PKSs which catalyze the condensation of the 9-decynoyl moiety generated by TtuA, B, and C with two malonyl-CoAs, followed by a spontaneous intramolecular cyclization (Figure 5A). In addition to **1**, we also observed a major byproduct **2** that retains a terminal alkene functionality from the 9-decenoyl precursor fed to the cultures of XZ2–4 (Figure 5, Supplementary Table 4, and Supplementary Figure 12). Unfortunately, no production of alkyne-bearing polyketides was observed from the culture extracts of the strains that overexpressed *hspks1/oras/pks18* and *jamABC* homologs from *B. pseudomallei* K96243, *S. paludicola* DSM 18645, *C. sp.* ATCC 51142, *P. entomophila* L48, *P. fluorescens* Pf-5 and *I. limosus* DSM 16000, though these strains still produced alkyl or terminal alkene-bearing polyketides in a significant amount (Supplementary Table 4). SDS-PAGE analyses revealed that the JamB homologs from *S. paludicola* DSM 18645, *B. pseudomallei* K96243, and *I. limosus* DSM 16000 were poorly or not expressed, while the others from *C. sp.* ATCC 51142, *P. entomophila* L48, and *P. fluorescens* Pf-5 were expressed at a level similar to TtuB (Supplementary Figure 8). Although inactive desaturases might be a possible explanation for the failure of terminal alkyne formation in *E. coli*, we cannot rule out the possibility of alternative substrates being required for the JamB homologs. It is likely that the fatty acyl moiety has to first be modified by other enzymes in the cluster before acetylenation catalyzed by a JamB homolog can further take place.

Biochemical analysis of TtuA, B, and C

Our screening efforts led to the discovery of new terminal alkyne biosynthetic machinery consisting of TtuA, B, and C. In addition to the reconstitution of its activity *in vivo*, *in vitro* biochemical characterization was further performed to confirm the functions of these three proteins. TtuA shows moderate sequence similarity to JamA (Identity/Similarity = 34%/51%), and was demonstrated to ligate medium-chain fatty acids to TtuC (Figure 4). To determine the substrate preference of TtuA, we set out to measure the kinetic parameters of TtuA towards fatty acids of different chain lengths. Due to the strong intrinsic ATPase activity of TtuA, an ACP loading assay followed by LC-HRMS analysis was used to obtain the kinetic data of TtuA by monitoring the formation of fatty acyl-TtuC. Comparison of the substrate specificity constants of TtuA showed that TtuA prefers the C_{10} fatty acid: the $k_{\text{cat}}/K_{\text{m}}$ value for decanoic acid ($27.73 \pm 2.36 \text{ min}^{-1}\text{mM}^{-1}$) was higher than those for other fatty acids of different chain lengths (Table 1 and Figure 6A), which is consistent with the observed starter unit for compound **1** biosynthesis.

We then tried to reconstitute the activity of TtuB *in vitro* for terminal alkyne biosynthesis. We used the extracted TtuB-containing membrane fraction in enzymatic assays and the membrane fraction from *E. coli* transformed with an empty vector as a negative control. LC-HRMS analysis demonstrated that 9-decynoyl-TtuC (**4**) was successfully generated after the

incubation of TtuA, C, and the TtuB-containing membrane fraction with 9-decenoic acid and ATP, but it was not detectable in the controls without TtuB or TtuC (Figure 6B and C). These results unequivocally confirmed the role of TtuB as a terminal acetylenase that transforms a terminal alkenoyl-TtuC (**5**) to a terminal alkynoyl-TtuC (**4**). However, few turnovers were observed, probably due to the lack of an efficient electron transport system and reducing equivalents *in vitro*, although the semi-purified TtuB-containing membrane fraction could provide these essential components to some extent^{16, 28}.

***In vivo* characterization of TtuB using the *E. coli*-PKS reporting system**

Because of the low activity of TtuB *in vitro*, we decided to take advantage of the *in vivo* reporting system for a detailed investigation of this newly identified acetylenase. We prepared two additional *E. coli* strains by omitting *ttuA* or *ttuC* from the strain XZ2. 9-Decenoic acid was fed to all of the cultures. Deletion of *ttuA* decreased the titer of **1** by ~100-fold, as determined by selected MS ion monitoring, similar to the behavior of the *jamABC* system (Supplementary Table 4 and Supplementary Figure 6)¹⁶. To our surprise, omitting *ttuC* did not completely abolish the production of **1**; instead, the titer of **1** was decreased by ~100-fold (Supplementary Table 4 and Supplementary Figure 6). In addition, the omission of both *ttuA* and *ttuC* still resulted in the production of a trace amount of **1** (Supplementary Table 4 and Supplementary Figure 6). These results show that unlike JamB, the reaction catalyzed by TtuB does not have a stringent requirement for the presence of TtuC. TtuB could presumably recognize 9-decenoyl-CoA or 9-decenoyl-ACP formed by an acyl-CoA ligase or an acyl-ACP synthetase endogenous to *E. coli*, albeit with a much lower efficiency.

Since acetylenases are often reported to be able to catalyze two sequential O₂-dependent dehydrogenations of a C-C bond, the first introducing a double bond and then a triple bond^{8, 13, 16}, we investigated the activity of TtuB on decanoyl-TtuC by feeding decanoic acid into the cultures of XZ2. LC-UV-MS analyses demonstrated that XZ2 was capable of producing compounds **1** and **2**, while the production of **1** and **2** was completely abolished in the control strain without *ttuB* (Figure 7 and Supplementary Table 4). These results show that TtuB is essential for the production of both the terminal alkene and alkyne functionalities. The production of **3**, a byproduct with a saturated fatty acyl starter unit was also observed (Figure 7, Supplementary Table 4, and Supplementary Figure 13). Since HsPKS1 has no preference for starter units with different terminal functionalities, the efficiency in the production of **1**, **2** and **3** is most likely determined by the activity of TtuB. Products **1**, **2** and **3** were produced in a ~3:1:3 ratio in *E. coli*, indicating that TtuB was able to transform decanoyl-TtuC to 9-decynoyl-TtuC through a four-electron oxidation in a very efficient way (Figure 5A). The TtuB-catalyzed conversion of decanoyl-TtuC to 9-decenoyl-TtuC and of 9-decenoyl-TtuC to 9-decynoyl-TtuC was estimated to be ~86% and ~75%, respectively (Figure 5A). Compared to the JamB studies in which the feeding of hexanoic acid only led to the production of a trace amount of the terminal alkyne-bearing polyketide, our results from this decanoic acid feeding experiment demonstrate that TtuB is a much more efficient bifunctional desaturase/acetylenase than JamB in *E. coli*.

We next scrutinized the acyl chain length specificity of TtuB by feeding nonanoic acid or 10-undecenoic acid to the cultures of XZ2 (Supplementary Table 4). Kinetic data suggested that similar to decanoic acid, 10-undecenoic acid could be effectively activated and loaded onto TtuC catalyzed by TtuA, but nonanoic acid was only weakly activated (Table 1 and Figure 6A). No production of the corresponding alkyne-bearing products was observed by selected MS monitoring, demonstrating that TtuB has a strict chain length specificity towards C₁₀. Sequence alignment revealed that TtuB shares 44% identity and 60% similarity to JamB and has three conserved histidine cluster motifs (HX₄H, HX₂HH, and QX₂HH) (Supplementary Figure 9). The conserved regions are proposed to be catalytically essential as ligands for the di-iron cluster of the enzyme. It is notable that the third histidine box domain in JamB aligns to the consensus sequence of ⁵-/⁶-desaturases (QX₂HH), consistent with the regiospecificity of JamB^{9, 11, 16}. However, this His box is also conserved in TtuB, which raises an interesting question about the key residues of terminal acetylenases in determining the regiospecificity. TtuB, along with our previously characterized JamB, may thus serve as model enzymes for further mechanistic and structural studies of desaturases/acetylenases.

Conclusions

Following our recent efforts on characterizing the first ACP-dependent terminal alkyne biosynthetic machinery from the marine cyanobacteria *Moorea producens*¹⁶, we here have screened additional homologous gene cassettes presumably involved in alkyne biosynthesis using both *in vitro* biochemical assays and the *E. coli*-PKS *in vivo* reporting system. We discovered and characterized a new terminal alkyne biosynthetic pathway from *T. turnerae* T7901 with altered substrate specificity and improved efficiency in *E. coli*. Through both *in vitro* and *in vivo* analyses, we demonstrated that decanoic acid was activated and loaded onto TtuC by TtuA, and it was then modified by TtuB to form a terminal alkyne moiety. TtuB is thus a bifunctional desaturase/acetylenase that is specific for the C₁₀ fatty acyl substrate, and it is able to efficiently catalyze two sequential O₂-dependent dehydrogenation reactions in *E. coli*. Further bioinformatic analysis showed that *ttuABC* are in proximity with genes that encode a membranebound desaturase, a dehydrogenase, and hypothetical proteins (Supplementary Figure 10), suggesting that they are likely involved in the biosynthesis of a polyne metabolite in *T. turnerae* T7901. The elucidation of the functions of TtuA, B, and C in terminal alkyne biosynthesis thus also facilitates the discovery of the corresponding alkyne-bearing natural products from *T. turnerae* by trapping metabolites from crude mixtures using the azide-alkyne bioorthogonal reaction¹⁴. Furthermore, this work expands the terminal alkyne biosynthetic toolbox that has wide applications in synthetic and chemical biology.

Methods

Materials

Phusion High-Fidelity PCR Master Mix (NEB) was used for PCR reactions. Restriction enzymes were purchased from Thermo Scientific. 9-Decenoic acid was purchased from Oakwood Chemical. *d*6-DMSO was purchased from Cambridge Isotope Laboratories, Inc.

LB medium was purchased from EMD Chemicals. Other chemicals were obtained from Alfa Aesar or Sigma Aldrich.

Phylogenetic analysis

We searched for homologs of JamABC within the IMG/JGI database (<http://img.jgi.doe.gov/>) containing finished or permanent draft of sequenced genomes using default parameters. A cladogram of JamB homologs was made using default parameters in IMG/JGI. MEGA 6.06 was applied for sequence alignment and molecular evolutionary analysis of the JamA homologs. The consensus phylogenetic tree was constructed using the Neighbor- Joining method tested with Bootstrap with 1000 replications.

Bacterial strains, plasmids and DNA manipulations

Selected genes from *Pseudomonas entomophila* L48, *Burkholderia pseudomallei* K96243, *Burkholderia vietnamiensis* G4, *Stigmatella aurantiaca* DW4/3-1, *Cyanotheca* sp. ATCC 51142, *Teredinibacter turnerae* T7901, *Nocardia brasiliensis* ATCC 700358, *Cyanotheca* sp. PCC 7425, *Schlesneria paludicola* DSM 18645 and *Inquilinus limosus* DSM 16000 were synthesized with or without codon optimization as gBlocks (IDT). *JamABC* homologs from *Pseudomonas fluorescens* Pf-5 were PCR amplified from genomic DNA extracted from the bacteria. *Pks18* was PCR amplified from genomic DNA extracted from *Mycobacterium tuberculosis* H37Rv. *Hspks1*, *csyA*, and *oras* were provided by I. Abe (the University of Tokyo), J. Zhan (Utah State University) and H. Zhao (the University of Illinois, Urbana-Champaign), respectively. Plasmid constructions were performed using standard protocols. Plasmids were purified from *E. coli* XL1-Blue with a QIAprep Spin Miniprep Kit and confirmed by DNA sequencing.

Protein expression and purification

Soluble proteins purified in this work contained C-terminus hexahistidine tags or were fused with N-terminally His-tagged MBP. The expression plasmids were transformed into *E. coli* BL21-Gold (DE3) or BAP1 for protein expression. The cells were grown at 37 °C in 750 mL of LB medium with appropriate concentrations of antibiotics to an OD₆₀₀ of 0.4–0.6, followed by the induction with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 16 h at 16 °C. The cells were harvested by centrifugation (4424 × g, 15 min, 4 °C), resuspended in 30 mL of lysis buffer (50 mM HEPES, pH 8.0, 300 mM NaCl, 10 mM imidazole) and lysed by homogenization on ice. Cell debris was removed by centrifugation (15,000 × g, 1 h, 4 °C). Ni-NTA agarose resin was added to the supernatant (1.5 mL/L of culture), and the solution was nutated at 4 °C for 1 h. The protein-resin mixture was loaded onto a gravity flow column, and proteins were eluted with increasing concentrations of imidazole in buffer A (50 mM HEPES, pH 8.0, 300 mM NaCl). Purified proteins were concentrated and buffer exchanged into HEPES buffer (50 mM HEPES, pH 8.0, 100 mM NaCl) with Vivaspin® Turbo centrifugal concentrators. The final proteins were flash-frozen in liquid nitrogen and stored at –80 °C. Protein concentrations were determined by NanoDrop. The approximate recombinant protein yields for JamA homologs from *Pseudomonas fluorescens* Pf-5 (68 kDa), *Pseudomonas entomophila* L48 (63 kDa), *Burkholderia pseudomallei* K96243 (63 kDa), *Cyanotheca* sp. ATCC 51142 (70 kDa), *Teredinibacter turnerae* T7901 (72 kDa),

Inquilinus limosus DSM 16000 (94 kDa) and *Schlesneria paludicola* DSM 18645 (109 kDa) were 19.8 mg L⁻¹, 3.8 mg L⁻¹, 3.8 mg L⁻¹, 4.7 mg L⁻¹, 1.7 mg L⁻¹, 1.2 mg L⁻¹ and 3.2 mg L⁻¹, respectively. The approximate protein yields for JamC homologs from *Pseudomonas fluorescens* Pf-5 (13 kDa), *Pseudomonas entomophila* L48 (10 kDa), *Burkholderia pseudomallei* K96243 (11 kDa), *Cyanothece* sp. ATCC 51142 (12 kDa), *Teredinibacter turnerae* T7901 (10 kDa), *Inquilinus limosus* DSM 16000 (10 kDa) and *Schlesneria paludicola* DSM 18645 (10 kDa) were 7.0 mg L⁻¹, 4.6 mg L⁻¹, 4.5 mg L⁻¹, 3.4 mg L⁻¹, 4.0 mg L⁻¹, 5.9 mg L⁻¹ and 5.3 mg L⁻¹, respectively.

For expression and purification of JamB homologs, plasmid containing *jamB* homolog or the empty pETDuet-1 vector was transformed into *E. coli* BL21-Gold (DE3) cells. The cells were grown at 37 °C in LB medium with 100 µg mL⁻¹ carbenicillin to an OD₆₀₀ of 0.4. The cells were then induced with 0.1 mM IPTG for 16 h at 16 °C. The cells were harvested by centrifugation (4424 × g, 15 min, 4 °C), resuspended in HEPES buffer (50 mM HEPES, pH 8.0, 100 mM NaCl), and lysed by homogenization on ice. Cell debris was removed by centrifugation (10,000 × g, 10 min, 4 °C). The supernatant was further processed by ultracentrifugation (100,000 × g, 1 h, 4 °C) to collect the cell membrane. The membrane fractions from *ttuB*-overexpressing strain or *E. coli* transformed with a empty vector were used in the *in vitro* assays immediately.

ACP loading assay

For the initial screening of possible substrates for acyl-ACP synthetases, a typical reaction mixture contained 50 mM HEPES (pH 8.0), 2 mM MgCl₂, 2 mM ATP, 1 mM TCEP, 5 mM fatty acids and 10–100 µM holo-ACP. The reactions were initiated by the addition of 10–20 µM acyl-ACP synthetase. All of the reactions were incubated at room temperature overnight. For the kinetic studies of TtuA, a typical assay contained, in a total volume of 100 µL, 50 mM HEPES (pH 8.0), 1 mM TCEP, 2 mM ATP, 2 mM MgCl₂, 25 µM TtuC, 2 or 5 µM TtuA (2 µM of TtuA for assays with nonanoic acid, decanoic acid and 10-undecenoic acid, 5 µM of TtuA for assays with 7-octenoic acid and lauric acid). Reactions were initiated by the addition of 25 µM of acid substrates. At regular time intervals (0, 3, 5, 10, 20 min for most assays except for assays with 7-octenoic acid in which 0, 10, 30, 60, 120 min were used), 10 µL aliquots were quenched with 10 µL of 25% (v/v) formic acid. The mixture was then diluted to a final concentration of 1 µM of ACP and filtered by syringe filters (0.2 µm) to remove any precipitate. The resulting protein sample was processed by nanocapillary LC-HRMS analysis using a chip column (Agilent Zorbax 300SB-C18, 5 µm; separation, 43 mm × 75 µm; enrichment, 4 mm 40 nL) in-line with a QTOF (Agilent 6510 Q-TOF LC/MS). A linear gradient of 3–95% CH₃CN (v/v) over 8 min in H₂O with 0.1% (v/v) formic acid at a flow rate of 0.6 µL min⁻¹ was used for analysis. MassHunter Qualitative Analysis software was used for data analysis, and the intact protein masses were obtained using a maximum entropy deconvolution algorithm.

Type III PKS assays on acyl-ACPs

For the study of type III PKSs activities on acyl-ACPs, a typical reaction mixture contained 50 mM HEPES (pH 8.0), 2 mM MgCl₂, 2 mM ATP, 1 mM TCEP, 1 mM lauric acid, 100 µM holo-ACP, 10 µM acyl-ACP synthetase, 1mM malonyl CoA and 20 µM type III PKS. A

typical positive control reaction mixture contained 50 mM HEPES (pH 8.0), 1 mM TCEP, 1 mM lauryl CoA, 2 mM malonyl CoA and 20 μ M type III PKS. All of the reactions were incubated at room temperature overnight. The reaction was quenched by the addition of trichloroacetic acid to a final concentration of 5% (v/v). Precipitated proteins were removed by centrifugation. LC-HRMS analysis was performed on an Agilent Technologies 6520 Accurate Mass QTOF LC/MS with an Agilent Eclipse Plus C18 column (4.6 $\times$ 100 mm) using a linear gradient of 2–95% CH₃CN (v/v) over 20 min followed by 15 min in 95% CH₃CN supplemented with 0.1% (v/v) formic acid at a flow rate of 0.5 ml min⁻¹.

TtuB activity assay on 9-decenoyl-TtuC

9-Decenoyl-TtuC was first generated *in situ* by TtuA using the assay condition described above and then further processed in the TtuB activity assay. Particularly, 0.1 mM Fe(NH₄)₂(SO₄)₂, 500 u of catalase, and the TtuB-containing membrane fraction or the membrane fraction purified from the *E. coli* strain transformed with an empty vector were added to the reaction mixtures. A control reaction without substrate was set up under similar conditions except that TtuC was not provided in the assay. The membrane fraction was removed by centrifugation (21,000 $\times$ g, 4 °C) after 20 min of incubation. The supernatant was then diluted to a final concentration of 1 μ M of ACP, followed by nanocapillary LC-HRMS analysis as described above.

In vivo screening for active JamB-like acetylenases

For the initial screening, BAP1 strains transformed with the appropriate plasmids were grown in 25 mL of terrific broth (24 g L⁻¹ yeast extract, 12 g L⁻¹ tryptone, and 10 g L⁻¹ NaCl, 6 mL L⁻¹ glycerol, pH 7.4) with the appropriate antibiotics at 37 °C to an OD₆₀₀ of 0.8–1.0, followed by the induction with 0.5 mM IPTG and supplementation of 1 mM fatty acid. For the detailed investigation of TtuA, B, and C activity *in vivo*, BAP1/pXZ92+pXZ101, BAP1/pXZ24+pXZ101, BAP1/pXZ92+pXZ97, BAP1/pXZ92+pXZ107, BAP1/pXZ166+pXZ172, BAP1/pXZ159+pXZ172, BAP1/pXZ174+pXZ172 and BAP1/pXZ175+pXZ172 cells were grown in 125 mL of LB medium with 100 μ g mL⁻¹ carbenicillin and 100 μ g mL⁻¹ spectinomycin at 37 °C to an OD₆₀₀ of 0.4–0.6. The cells were harvested and concentrated 5-fold into 25 mL fresh F1 medium (1L contains 3 g KH₂PO₄, 6.62 g K₂HPO₄, 4 g (NH₄)₂SO₄, 150.5 mg MgSO₄, 5 g glucose, 1.25 mL trace metal solution, 100 μ M Fe(NH₄)₂(SO₄)₂, and 10 mL 100 $\times$ vitamin solution) supplemented with 100 μ g mL⁻¹ carbenicillin, 100 μ g mL⁻¹ spectinomycin, 0.5 mM IPTG, and 1 mM fatty acid. After 48 h of growth at 20 °C, compounds were extracted from cell-free supernatant using ethyl acetate. The solvent was removed by rotary evaporation and the residue was re-dissolved in methanol (0.5 mL) and analyzed by LC-UV-MS and LC-HRMS (20 μ L injection) with an Agilent Eclipse Plus C18 column (4.6 $\times$ 100 mm). A linear gradient of 2–95% CH₃CN (v/v) over 20 min and 95% CH₃CN for a further 15 min in H₂O supplemented with 0.1% (v/v) formic acid at a flow rate of 0.5 mL min⁻¹ was used for LC-UV-MS and LC-HRMS analysis. LC-UV-MS analysis was performed on an Agilent Technologies 6120 Quadrupole LC-MS (with DAD). LCHRMS analysis was performed on an Agilent Technologies 6520 Accurate Mass QTOF LC-MS.

Calculation of TtuB catalytic efficiency in *E. coli*

The feeding of decanoic acid led to the production of **1**, **2** and **3** in a 3:1:3 ratio. 100% of compounds **1** and **2** were derived from 9-decynoyl-TtuC and 9-decenoyl-TtuC, respectively. Therefore, the actual conversion of 9-decenoyl-TtuC to 9-decynoyl-TtuC catalyzed by TtuB in *E. coli* was ~75%. The feeding of 9-decenoic acid led to the production of **1** and **2** in a 1:1 ratio, and part of compound **2** was derived from 9-decenoyl-CoA prepared by an acyl-CoA ligase endogenous to *E. coli*. Based on the 75% conversion efficiency of TtuB to generate 9-decynoyl-TtuC from 9-decenoyl-TtuC, we estimated that ~33% of the uptaken 9-decenoic acid was converted to 9-decenoyl-CoA in forming **2** when 9-decenoic acid was fed into the culture. By assuming that ~33% of the uptaken decanoic acid was also converted to decanoyl-CoA in forming **3** in the decanoic acid feeding experiment, we estimated that ~78% of compound **3** was from decanoyl-CoA starter unit, while only ~22% of **3** was from the leftover decanoyl-TtuC prepared by TtuA. The actual catalytic efficiency of TtuB from decanoyl-TtuC to 9-decenoyl-TtuC in *E. coli* was thus calculated to be ~86%.

Large-scale production, purification and characterization of compound **1**

For large-scale production, a 50 mL seed culture of XZ2 was inoculated into 5 L of LB medium supplemented with spectinomycin (100 $\mu\text{g mL}^{-1}$) and carbenicillin (100 $\mu\text{g mL}^{-1}$). After induction with IPTG, the strains continued to grow in F1 minimal medium supplemented with 1 mM of 9-decenoic acid at 20°C for two days. Compound **1** was extracted from the cell-free supernatant using ethyl acetate. The solvent was removed by rotary evaporation and the residue was re-dissolved in methanol. Compound **1** was purified using reverse-phase high-performance liquid chromatography (RP-HPLC) through an Inertsil ODS-4 column (4.6 mm i.d., 250 mm L, GL Sciences Inc.) with a linear gradient of 50–95% CH_3CN (v/v) over 20 min followed by 5 min in 95% CH_3CN at a flow rate of 0.8 mL min^{-1} . Fractions containing **1** were collected manually, concentrated under vacuum, and dried to a white solid using a lyophilizer. These fractions were further purified by RP-HPLC with an Agilent Eclipse Plus C18 column (4.6 $\times$ 100 mm) using an isocratic program of 40% CH_3CN (v/v) in H_2O at a flow rate of 0.5 mL min^{-1} . The resulting purified compound **1** was dried and analyzed by HRMS and NMR. NMR spectra of **1** were recorded on a Bruker Biospin 900 MHz spectrometer with a cryoprobe in *d*6-dimethyl sulfoxide (*d*6-DMSO; Cambridge Isotope Laboratories).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

This research was financially supported by the Pew Scholars Program and the National Institutes of Health (DP2AT009148). We thank I. Abe (the University of Tokyo) for providing *hspksI*, J. Zhan (Utah State University) for providing *csyA*, H. Zhao (the University of Illinois, Urbana-Champaign) for providing *oras*, and M. Chang (UC Berkeley) for providing pSV272.1. We also thank S. Bauer (UC Berkeley) for assisting with LC-HRMS analysis, J. Pelton (UC Berkeley) for helping with NMR spectroscopic analysis and J. Liu (UC Berkeley) for valuable suggestions in the preparation of the manuscript.

References

1. Binder WH, Kluger C. Azide/alkyne-“click” reactions: applications in material science and organic synthesis. *Curr. Org. Chem.* 2006; 10:1791–1815.
2. Hein CD, Liu X-M, Wang D. Click chemistry, a powerful tool for pharmaceutical sciences. *Pharm. Res.* 2008; 25:2216–2230. [PubMed: 18509602]
3. Sletten EM, Bertozzi CR. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed. Engl.* 2009; 48:6974–6998. [PubMed: 19714693]
4. Grammel M, Hang HC. Chemical reporters for biological discovery. *Nat. Chem. Biol.* 2013; 9:475–484. [PubMed: 23868317]
5. Thirumurugan P, Matosiuk D, Jozwiak K. Click chemistry for drug development and diverse chemical-biology applications. *Chem. Rev.* 2013; 113:4905–4979. [PubMed: 23531040]
6. Zhu X, Zhang W. Tagging polyketides/non-ribosomal peptides with a clickable functionality and applications. *Frontiers in Chemistry.* 2015; 3:11. [PubMed: 25815285]
7. Prescher JA, Bertozzi CR. Chemistry in living systems. *Nat. Chem. Biol.* 2005; 1:13–21. [PubMed: 16407987]
8. Shanklin J, Guy JE, Mishra G, Lindqvist Y. Desaturases: emerging models for understanding functional diversification of diiron-containing enzymes. *J. Biol. Chem.* 2009; 284:18559–18563. [PubMed: 19363032]
9. Shanklin J, Cahoon EB. Desaturation and related modifications of fatty acids. *Annu. Rev. Plant Phys.* 1998; 49:611–641.
10. Lee M, Lenman M, Bana A, Bafor M, Singh S, Schweizer M, Nilsson R, Liljenberg C, Dahlqvist A, Gummeson P-O, Sjö Dahl S, Green A, Stymne S. Identification of non-heme diiron proteins that catalyze triple bond and epoxy group formation. *Science.* 1998; 280:915–918. [PubMed: 9572738]
11. Minto RE, Blacklock BJ. Biosynthesis and function of polyacetylenes and allied natural products. *Prog. Lipid Res.* 2008; 47:233–306. [PubMed: 18387369]
12. Blacklock BJ, Scheffler BE, Shepard MR, Jayasuriya N, Minto RE. Functional diversity in fungal fatty acid synthesis: the first acetylenase from the Pacific golden chanterelle *Cantharellus formosus*. *J. Biol. Chem.* 2010; 285:28442–28449. [PubMed: 20606235]
13. Haritos VS, Horne I, Damcevski K, Glover K, Gibb N, Okada S, Hamberg M. The convergent evolution of defensive polyacetylenic fatty acid biosynthesis genes in soldier beetles. *Nat. Commun.* 2012; 3:1150. [PubMed: 23093187]
14. Ross C, Scherlach K, Kloss F, Hertweck C. The molecular basis of conjugated polyene biosynthesis in phytopathogenic bacteria. *Angew. Chem. Int. Ed. Engl.* 2014; 53:7794–7798. [PubMed: 24898429]
15. Fritsche K, van den Berg M, de Boer W, van Beek TA, Raaijmakers JM, van Veen JA, Leveau JH. Biosynthetic genes and activity spectrum of antifungal polyynes from *Collimonas fungivorans* Ter331. *Environ. Microbiol.* 2014; 16:1334–1345. [PubMed: 24588891]
16. Zhu X, Liu J, Zhang W. *De novo* biosynthesis of terminal alkyne-labeled natural products. *Nat. Chem. Biol.* 2015; 11:115–120. [PubMed: 25531891]
17. Edwards DJ, Marquez BL, Nogle LM, McPhail K, Goeger DE, Roberts MA, Gerwick WH. Structure and biosynthesis of the jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*. *Chem. Biol.* 2004; 11:817–833. [PubMed: 15217615]
18. Dorrestein PC, Blackhall J, Straight PD, Fischbach MA, Garneau-Tsodikova S, Edwards DJ, McLaughlin S, Lin M, Gerwick WH, Kolter R, Walsh CT, Kelleher NL. Activity screening of carrier domains within nonribosomal peptide synthetases using complex substrate mixtures and large molecule mass spectrometry. *Biochemistry.* 2006; 45:1537–1546. [PubMed: 16460000]
19. Jones AC, Monroe EA, Podell S, Hess WR, Klages S, Esquenazi E, Niessen S, Hoover H, Rothmann M, Lasken RS, Yates JR 3rd, Reinhardt R, Kube M, Burkart MD, Allen EE, Dorrestein PC, Gerwick WH, Gerwick L. Genomic insights into the physiology and ecology of the marine filamentous cyanobacterium *Lyngbya majuscula*. *Proc. Natl. Acad. Sci. U. S. A.* 2011; 108:8815–8820. [PubMed: 21555588]

20. Pfeifer BA, Admiraal SJ, Gramajo H, Cane DE, Khosla C. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science*. 2001; 291:1790–1792. [PubMed: 11230695]
21. Terpe K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* 2003; 60:523–533. [PubMed: 12536251]
22. Wanibuchi K, Zhang P, Abe T, Morita H, Kohno T, Chen G, Noguchi H, Abe I. An acridone-producing novel multifunctional type III polyketide synthase from *Huperzia serrata*. *The FEBS journal*. 2007; 274:1073–1082. [PubMed: 17250741]
23. Funa N, Awakawa T, Horinouchi S. Pentaketide resorcylic acid synthesis by type III polyketide synthase from *Neurospora crassa*. *J. Biol. Chem.* 2007; 282:14476–14481. [PubMed: 17374612]
24. Rubin-Pitel SB, Zhang H, Vu T, Brunzelle JS, Zhao H, Nair SK. Distinct structural elements dictate the specificity of the type III pentaketide synthase from *Neurospora crassa*. *Chem. Biol.* 2008; 15:1079–1090. [PubMed: 18940668]
25. Saxena P, Yadav G, Mohanty D, Gokhale RS. A new family of type III polyketide synthases in *Mycobacterium tuberculosis*. *J. Biol. Chem.* 2003; 278:44780–44790. [PubMed: 12941968]
26. Yu D, Zeng J, Chen D, Zhan J. Characterization and reconstitution of a new fungal type III polyketide synthase from *Aspergillus oryzae*. *Enzyme Microb. Technol.* 2010; 46:575–580.
27. Zhang W, Li Y, Tang Y. Engineered biosynthesis of bacterial aromatic polyketides in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 2008; 105:20683–20688. [PubMed: 19075227]
28. Wada H, Avelange-Macherel MH, Murata N. The *desA* gene of the cyanobacterium *Synechocystis* sp. strain PCC6803 is the structural gene for delta 12 desaturase. *J. Bacteriol.* 1993; 175:6056–6058. [PubMed: 8376352]

Figure 1.

Examples of gene clusters containing *jamABC* homologs. (A) Three gene clusters have been identified for the biosynthesis of acetylenic natural products with a terminal alkyne functionality. (B) Examples of the uncharacterized gene clusters that contain *jamABC* homologs. The genes studied in this paper are boxed.

Figure 2. Schematic of the strategy used for the reconstitution of JamA, B, and C homologs. Substrate specificities of JamA homologs were determined by *in vitro* assays, and *in vivo* reporting systems were used for the study of JamB homologs.

Figure 3.

A cladogram of JamB and its homologs from cyanobacteria (red), proteobacteria (blue), actinobacteria (gray), and planctomycetes (green). Representative *jamB* homologs from different clades, and their neighboring *jamA* and *jamC* homologs were targeted for screening. The microorganisms containing these *jamABC* homologs are shown next to each clade.

Figure 4.

Substrate-dependent ACP loading activity of JamA homologs. (A) Biochemical characterization of the substrate specificities of JamA homologs towards fatty acids with different chain lengths. The highest conversion by each JamA homolog is shown as a relative activity of 100%. (B) A phylogenetic analysis of JamA and its homologs. Group A contains the homologs that prefer medium-chain fatty acids, and group B contains the ones that prefer medium- to long-chain fatty acids.

Figure 5. Biosynthesis of **1** using the *in vivo* reporting system. (A) Schematic of **1** formation in *E. coli*. (B) LC-UV-MS analysis (285 nm) showing the production of **1** upon the co-expression of *ttuABC* and *hspks1* and the feeding of 9-decenoic acid (trace i). **2** was a major byproduct. The absence of *ttuA* (trace ii), *ttuB* (trace iii), *ttuC* (trace iv), or *ttuAC* (trace v) significantly decreased or abolished the production of **1** (Supplementary Figure 6).

Figure 7.

In vivo characterization of TtuA, B, and C. LC-UV-MS analysis (285 nm) showing the production of **1** and **2** by the strain XZ2 upon the feeding of decanoic acid (trace i). **3** was the major byproduct. The absence of *ttuB* (trace ii) abolished the production of **1** and **2**. Culture extracts of the strain XZ2 fed with 9-decenoic acid is also shown for comparison (trace iii).

Table 1

The substrate specificity constants of TtuA towards fatty acids of different chain lengths.

Fatty acids	7-Octenoic acid	Nonanoic acid	Decanoic acid	10-Undecenoic acid	Lauric acid
k_{cat}/K_m ($\text{min}^{-1}\text{mM}^{-1}$)	0.20 ± 0.03	3.32 ± 0.24	27.73 ± 2.36	22.40 ± 2.43	18.82 ± 1.17