Title
Engineered biosynthesis of alkyne-tagged polyketides.

Authors
Gu, Di
Zhang, Wenjun

Publication Date
2022

DOI
10.1016/bs.mie.2021.11.013

Peer reviewed

HHS Public Access

Author manuscript

Methods Enzymol. Author manuscript; available in PMC 2023 January 10.

Published in final edited form as:

Methods Enzymol. 2022 ; 665: 347–373. doi:10.1016/bs.mie.2021.11.013.

Engineered Biosynthesis of Alkyne-tagged Polyketides

Di Gu[†], **Wenjun Zhang**[‡,*,1]

[†]Department of Chemistry, University of California, Berkeley, CA, United States

[‡]Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA, United States

[*]Chan Zuckerberg Biohub, San Francisco, CA, United States

Abstract

Polyketides have demonstrated their significance as therapeutics, industrial products, pesticides, and biological probes following intense study over the past decades. Tagging polyketides with a bioorthogonal functionality enables various applications such as diversification, quantification, visualization and mode-of-action elucidation. The terminal alkyne moiety, as a small, stable and highly selective clickable functionality, is widely adopted in tagging natural products. *De novo* biosynthesis of alkyne-tagged polyketides offers the unique advantage of reducing the background from feeding the biorthogonal moiety itself, leading to the accomplishment of *in situ* generation of a clickable functionality for bioorthogonal reactions. Here, we introduce several engineering strategies to apply terminal alkyne biosynthetic machinery, represented by JamABC, which produces a short terminal alkyne-bearing fatty acyl chain on a carrier protein, to functions with different downstream polyketide synthases (PKSs). Successful results in engineering type III and type I PKSs provide engineering guidelines and strategies that are applicable to additional PKSs to produce targeted alkyne-tagged metabolites for chemical and biological applications.

Keywords

terminal alkyne; polyketides; starter unit engineering; extender unit engineering; click chemistry; docking domain; site-directed mutagenesis

1. Introduction

Polyketides are a large family of structurally diverse secondary metabolites derived from natural sources such as animals, plants, fungi and bacteria, showing a broad range of bioactivities as drugs, pesticides, and biological probes (McDaniel et al., 1994; Wang et al., 2020). Some well-known examples of these polyketides include the antibiotic erythromycin and the immunosuppressant rapamycin, both of which were initially isolated from bacterial sources and have been approved for clinical use for decades (Cottens et al., 2019; Hertweck, 2009; Jelić & Antolović, 2016). The process for discovery, diversification, quantification, visualization and mode-of-action elucidation of polyketides remains challenging and

[1]Corresponding author: wjzhang@berkeley.edu.

time-consuming, although it has been improved in recent years due to many technical advancements (Porterfield et al., 2020). One such technology is to tag polyketides with a clickable functionality, which has been demonstrated to facilitate the study of polyketide biosynthesis, biology, and pharmacology through bioorthogonal chemistry (DeGuire et al., 2015; Kalkreuter et al., 2019; Seidel et al., 2019). Among all kinds of clickable moieties, the terminal alkyne functionality has emerged as one of the most potent tools in chemical biology due to its chemical stability in biological environments, its rarity in biology, and its selectivity in bioorthogonal reactions. Polyketides labeled with terminal alkyne enable the imaging and mode-of-action studies through the copper-catalyzed azide-alkyne cycloaddition reaction or directly using stimulated Raman scattering microscopy (Wei et al., 2014). The incorporation of a terminal alkyne tag into polyketides can be accomplished by total synthesis (Staub & Sieber, 2008), semi-synthesis (DeGuire et al., 2015; Seidel et al., 2019), precursor-directed biosynthesis (Moschny et al., 2020; Musiol-Kroll et al., 2017), and *de novo* biosynthesis (Zhu et al., 2015a). It is challenging to generate alkyne-tagged polyketides through total synthesis or semi-synthesis because of the structural complexity, the chemical lability and the limited supply of most polyketides. In regard to precursor-directed biosynthesis based on the promiscuity of biosynthetic machinery, the coexistence of diffusible precursors and final products with the same chemical handle leads to a substantial background in the system, making it incompatible with *in situ* bioorthogonal reactions (Zhu et al., 2015a). Thus, the development of *de novo* biosynthesis of alkyne-tagged polyketides without the feeding of alkynoic precursors is imperative.

In contrast to the broad applications of the alkyne, few biosynthetic enzymes for terminal alkyne formation have been characterized so far. Our group (Zhu et al., 2015a) studied carrier protein-dependent terminal alkyne biosynthetic machinery in microbes inspired by the polyketide natural product jamaicamide B, isolated from the marine cyanobacteria *Moorea producens* (formerly classified as *Lyngbya majuscula*) (Figure 1a). The terminal alkyne functionality of jamaicamide B is introduced as a short-chain alkynoic starter unit by a three-gene operon, *jamABC*, which encode a homologue of fatty acyl-acyl carrier protein (ACP) synthetase, a membrane-bound fatty acid bifunctional desaturase/acetylenase, and an ACP, respectively. JamA is known to activate and load hexanoic acid onto JamC, and the resulting hexanoyl-JamC is modified by JamB to yield 5-hexynoyl-JamC as a starter unit for the downstream polyketide synthase/nonribosomal peptide synthetase (PKS/NRPS) assembly line. Although JamA is able to activate fatty acids of different lengths and unsaturation patterns, JamB was demonstrated to have more stringent substrate specificity towards both the acyl chain length and the ACP. Genome mining revealed at least 80 gene clusters encoding homologues of JamABC across diverse bacterial genera, providing a great potential to expand this molecular toolbox (Zhu, et al., 2015b). For example, we discovered a homologous terminal alkyne biosynthetic pathway comprised of TtuA, -B, and -C from *Teredinibacter turnerae* T7901. The biosynthetic machinery encoded by *ttuABC* is very similar to that of *jamABC*, with the major difference in the chain length of fatty acid substrates. Apart from the biosynthetic enzymes originating from fatty acid synthase (FAS)/PKS, the Chang group discovered a series of enzymes, BesABCDE, in the biosynthesis of a terminal-alkyne amino acid, β-ethynylserine (βes) (Marchand et al., 2019). First, the halogenase BesD chlorinates the C_γ position of L-lysine, followed by an unusual oxidative

C-C bond cleavage by the oxidase BesC to yield a terminal alkene. Then, the PLP-dependent enzyme BesB catalyzes γ-elimination of chloride and isomerization to form the terminal alkyne. BesA and BesE continue to modify L-propargylglycine to form the final product, βes. Notably, the presence of a γ-halogen as a leaving group is required to form the terminal alkyne, which suggests the chlorination happens prior to the acetylenation step.

The above biosynthetic enzymes for terminal alkyne formation serve as a good start point for *de novo* biosynthesis of alkyne-tagged polyketides, of which combinatorial biosynthesis has been studied extensively. (Awakawa et al., 2018; Barajas et al., 2017; Kalkreuter & Williams, 2018; Koryakina et al., 2017). While the BesBCD system to generate a terminal-alkyne amino acid is useful to label peptides (Marchand et al., 2019), its utility for tagging polyketides is not straightforward. The current work has thus mainly focused on the application of JamABC, which were discovered from a natural polyketide biosynthetic pathway, for *in situ* generation and incorporation of terminal alkynes into various molecular scaffolds of polyketides through biosynthetic pathway engineering (Porterfield et al., 2020; Porterfield & Zhang, 2020; Zhu et al., 2015a; Zhu, et al., 2016; Zhu et al., 2015b). In this chapter, we first describe general methods of gene expression, protein purification, and *in vitro* assays during engineering. They are followed by detailed examples of how to engineer JamABC and desired downstream PKSs through starter unit or extender unit engineering to generate alkyne-tagged polyketides, mainly focusing on type III and type I PKSs. At last, this chapter concludes with current challenges and future directions in this field.

2. Enzyme expression, protein purification and *in vitro* assays

2.1 Heterologous expression

2.1.1 Heterologous expression for protein purification—Since the native hosts for *jamABC* and its homologues are underdeveloped for protein expression and purification, a heterologous host is often employed. *Escherichia coli* is one of the most widely used heterologous expression hosts for recombinant proteins due to its fast growth rate, ample genetic tools, and well-characterized primary metabolism (Rosano & Ceccarelli, 2014; Zhu & Zhang, 2018). To express enzymes in *E. coli* for the purpose of protein purification, the commercially available strain *E. coli* BL21 (DE3) (Studier & Moffatt, 1986) in combination with pET vectors is typically used for heterologous expression of hexahistidine-tagged recombinant proteins. Commonly used pET vectors include pET24b that carries a *C*-terminal hexahistidine tag, pET28a that carries both *N*- and *C*-terminal hexahistidine tags. pSV272 vector encoding an *N*-terminal His_6-maltose-binding protein tag is occasionally used to enhance the solubility of the target recombinant protein in *E. coli* by promoting protein folding via a maltose-binding protein fusion (Nallamsetty & Waugh, 2006; Zhu et al., 2015b). The gene of interest can be cloned into the vectors under a T7 promoter either through restriction enzyme-dependent or ligation-independent cloning strategies. A T7 RNA polymerase under an isopropyl β-D-1-thiogalactopyranoside (IPTG)- inducible *lac*UV5 promoter is encoded in the genome of *E. coli* BL21 (DE3) and is required for the expression of target recombinant protein cloned on the pET vectors. Codon optimization of the target genes is sometimes required to avoid adverse effects due to differences in tRNA distribution between natural and heterologous hosts (Gustafsson et al., 2004).

It is notable that heterologous ACPs expressed in *E. coli* BL21 (DE3) are often in *apo*-forms, and a promiscuous phosphopantetheinyl transferase such as Sfp from *Bacillus subtilis* can be used to promote ACP post-translational modification to *holo*-forms (Quadri et al., 1998). Thus, a modified *E. coli* BL21 (DE3) with a chromosomal copy of *sfp* under the control of a T7 promoter (named as *E. coli* BAP1 strain) is usually used to express *holo*-ACPs (Pfeifer et al., 2001). A limitation of BAP1 system is that both *apo*- and *holo*-forms of ACPs sometimes can be co-purified due to incomplete conversion of *apo*-ACPs to *holo*-ACPs catalyzed by Sfp. Alternatively, *apo*-ACPs can still be purified from typical *E. coli* BL21 (DE3) strain, although they need to be coupled with phosphopantetheinyl transferases during *in vitro* assays.

Here is a general protein heterologous expression procedure for alkyne biosynthetic enzymes. The plasmid containing the gene of interest is transformed into electrocompetent or chemically competent *E. coli* BL21 (DE3) or BAP1 through electroporation or heat-shock transformation, respectively. A single colony is picked from LB agar plates supplemented with appropriate antibiotics and grown in 10 mL starter culture of LB plus appropriate antibiotics at 37 °C overnight. 1 mL of the overnight culture is mixed with 1 mL of 40% glycerol in a cryovial and saved at −80 °C for long-term storage. 8 mL of the remaining overnight seed culture is used to inoculate 800 mL of LB supplemented with appropriate antibiotics (~1:100 inoculation ratio). The culture is shaken at 250 rpm at 37 °C until OD_{600} reaches 0.4–0.6 and then placed on ice for 10 min, followed by the induction with a final concentration of 0.12 mM IPTG. The resulting culture is shaken at 220 rpm at 16°C for 16 h. The induction time, IPTG concentration, culture temperature, and duration can be adjusted for optimal expressions of different proteins.

2.1.2 Heterologous expression for polyketide biosynthesis—JamABC homologues are often identified from Gram-negative bacteria, and thus *E. coli* can serve as a representative model organism suitable to mimic the native membrane environment for functional expression of JamB homologues (Zhu et al., 2015b). In addition, the well-studied metabolism, genetic tractability, and fast growth rate of *E. coli* render it attractive for engineering the biosynthesis of alkyne-tagged metabolites. Notably, we preferred an *in vivo* system to produce alkyne-tagged polyketides rather than an *in vitro* system due to the limited activity of the semi-purified membrane fractions containing JamB homologues and the uncertainty for the electron donors and reducing equivalents for this acetylenation reaction.

To properly express enzymes required to produce alkyne-tagged polyketides in *E. coli*, multiple genes encoding JamABC homologues and PKSs should be introduced in a single *E. coli* cell. In terms of choices of plasmid vectors, commercially available Duet vectors (Novagen) are typically used for the coexpression of multiple genes in *E. coli*. The commonly used vectors include pRSFDuet-1, pETDuet-1, pCDFDuet-1, pCOLADuet-1, and pACYCDuet-1. Each Duet vector carries two T7 promoter-driven operons, and thus is suitable to coexpress at least two proteins. The vectors carry compatible replication origins and antibiotic resistance genes to effectively maintain multiple plasmids in a single *E. coli* cell. Notably, the target protein expression level can be affected by plasmid copy numbers, and therefore, it is necessary to optimize the combination of Duet vectors when

multiple plasmids are required for *in vivo* pathway reconstitution. A high plasmid copy number may result in extra metabolic burden and reduce bacterial growth rate (Rosano & Ceccarelli, 2014), and therefore, pRSFDuet-1 should be chosen with caution in a coexpression system. In our laboratory, pETDuet-1 and pCDFDuet-1 are generally used in a two-plasmid system, pCOLADuet-1 is often added when a three-plasmid system is required, and pACYCDuet-1 may be applied as a fourth plasmid (Zhu & Zhang, 2018). In an *E. coli*-PKS system, JamABC homologues and a type III or type I PKS are coexpressed on pETDuet-1 and pCDFDuet-1 vectors in *E coli* BAP1, with each gene controlled by a T7 promoter (Porterfield et al., 2020; Zhu et al., 2015a; Zhu et al., 2015b; Zhu et al., 2016).

Regarding culture conditions, F1 minimal medium is used for high-level polyketide production (Pfeifer et al., 2002; Zhang et al., 2008; Zhu et al., 2015a). Typically, LB medium is not chosen for polyketide production in the *E. coli*-PKS system. 1 L of F1 medium contains 3 g KH_2PO_4, 6.62 g K_2HPO_4, 4 g $(NH_4)_2SO_4$, 150.5 mg $MgSO_4$, 5-20 g glucose, 1.25 mL trace metal solution (optional), 100 μM $Fe(NH_4)_2(SO4)_2$, and 10 mL 100 × vitamin solution (optional). Trace metal solution is composed of 40 mg $ZnCl_2$, 200 mg $FeCl_3 \cdot 6H_2O$, 10 mg $CuCl_2 \cdot 2H_2O$, 10 mg $MnCl_2 \cdot 4H_2O$, 10 mg $Na_2B_4O_7 \cdot 10H_2O$, 10 mg $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ in 1 L water. 100 × vitamin solution contains 100 mg/L biotin, 100 mg/L choline chloride, 100 mg/L D-calcium pantothenate, 100 mg/L folic acid, 100 mg/L nicotinamide, 100 mg/L pyridoxal hydrochloride, 10 mg/L riboflavin, 100 mg/L thiamine hydrochloride, 200 mg/L i-inositol, and 8.5 g/L NaCl.

A general protein coexpression procedure for polyketide biosynthesis is described as followed. The plasmids are cotransformed into *E. coli* BAP1 strain, and a single colony is picked and grown in 3 mL starter culture of LB supplemented with appropriate antibiotics at 37 °C overnight. 0.5 mL of the overnight culture is mixed with equal volume of 40% glycerol in a cryovial and saved at −80 °C for long-term storage. 2 mL of overnight seed culture is used to inoculate 200 mL of LB supplemented with appropriate antibiotics. The culture is shaken at 250 rpm at 37 °C to an OD_{600} of 0.4–0.6. The cells are pelleted at 4000 × g at 4 °C for 20 min and resuspended in 40 mL fresh F1 medium supplemented with appropriate antibiotics, 0.5 mM IPTG, and 1 mM fatty acid for polyketide production. The culture is shaken at 220 rpm at 20 °C for ~2 days. The culture broth and the cell pellets are then separated by centrifugation at 4000 × g at 4 °C for 20 min, and polyketides are extracted from the supernatant using organic solvents such as ethyl acetate. 1% acetic acid can be added into organic solvents for a better extraction if polyketides are stable enough in an acidic environment. After removal of the organic solvent by rotary evaporation, the residue is redissolved in methanol and subjected to high-performance liquid chromatography (HPLC) or liquid chromatography–high-resolution mass spectrometry (LC-HRMS) for further analysis.

2.2 Protein purification

2.2.1 Soluble protein purification—

The cells are harvested by centrifugation at ~4500 × g at 4 °C for 15 min, resuspended in 30 mL of lysis buffer (50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 8.0, 300 mM NaCl, 10 mM imidazole), and lysed by sonication on ice. The lysis protocol we use for 1 L of *E.*

coli culture is 35% amplitude for 3 minutes with three seconds on and three seconds off (Sonicator 700 Watts, Qsonica). Cell debris is pelleted using centrifugation at 15,000 × g at 4 °C for 1 h, and the supernatant is filtered by syringe filters (0.45 μm). Ni-NTA agarose resin is then added to the filtrate (1.5 mL/L of culture), and the mixture is nutated at 4 °C for 1 h and loaded onto the gravity flow column. Proteins are eluted with increasing concentrations of imidazole (ranging from 10 to 250 mM) in buffer A (50 mM HEPES, pH 8.0, 100 mM NaCl). The cell debris and all fractions are collected for sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) analysis. An Amicon Ultra filter with an appropriate molecular weight cutoff is used to concentrate and buffer exchange the purified protein with buffer A (50mM HEPES, pH 8.0, 100 mM NaCl). Reducing agents (e.g., dithiothreitol (DTT)) and stabilizing elements (e.g., glycerol) can also be added in protein purification buffers, and the buffering system, pH, and salt concentration can be optimized for different proteins. Protein concentration is determined by measuring the absorbance at 280 nm or Bradford assay. The final protein is flash-frozen in liquid nitrogen and stored at −80 °C for further usage. Depending on different expression levels, the purity of the resulting proteins ranges from ~50% to 90%. Additional chromatographic methods (e.g., ion exchange or size exclusion chromatography) may be used for further purification when necessary.

2.2.2 Membrane-bound desaturase/acetylenase purification—

Although the purification of the alkyne-forming bifunctional desaturase/acetylenase has not been reported, homologous nonheme diiron-dependent membrane-bound desaturases or hydroxylases have been purified for biochemical and structural studies (Alonso & Roujeinikova, 2012; Bai et al., 2015; Shen et al., 2020; Wang et al., 2015). We here focus on the procedure of semi-purification of membrane fractions containing the desaturase/acetylenase which can also be used in biochemical assays. The plasmid bearing JamB homologues and the empty plasmid vector (negative control) are transformed into *E. coli* BL21 (DE3) cells, separately. The cells are grown at 37 °C in 800 mL of LB medium with appropriate antibiotics to an OD_{600} of 0.4–0.6, and then induced with 0.12 mM IPTG and culture for 16 h at 16 °C. The cells are harvested by centrifugation (~4,500 g, 15 min, 4 °C), resuspended in 30 mL of HEPES buffer (50 mM HEPES, pH 8.0, 100 mM NaCl) and lysed by homogenization or sonication on ice. Cell debris is removed by centrifugation (10,000 g, 10 min, 4 °C). The supernatant was further ultracentrifugated (100,000 g, 1 h, 4 °C) to collect the cell membrane. The membrane fractions can be used for *in vitro* assays immediately or stored at −80 °C for further usage.

2.3 *In vitro* product formation assays

When applying JamABC and their homologues to different downstream PKSs, one important factor is the communication between the ACP (JamC and its homologues) and the domains in PKSs to promote the recognition and translocation of the alkynoic unit. As mentioned above, the ACP encoded in the tri-gene cassette plays a significant role in acetylenation because JamB and its homologues have very stringent substrate specificity towards carrier proteins. Not only critical to alkyne biosynthesis, but carrier proteins in general are also central hubs in the biosynthesis of polyketide natural products. Selective and programmed carrier protein recognition by multiple biosynthetic enzymes is essential for

the precise assembly of complex natural product scaffolds, and combinatorial biosynthesis using simple mix-and-match strategy often yields significantly impaired assembly lines due to carrier protein incompatibility. Thus, to evaluate the compatibility of the ACP and downstream PKSs in engineering, a convenient and efficient assay for detecting the formation of *in vitro* products is performed.

The assay is typically carried out at room temperature for 30 min to 4 h in 50 mM HEPES (pH 8.0) buffer containing 2 mM $MgCl_2$, 5 mM fatty acid, 5 mM ATP, 10 μM fatty acyl-ACP ligase (e.g., JamA), 50 μM holo-ACP (e.g., JamC), 2 mM malonyl-CoA, 1 mM TCEP, and 20 μM PKS. Other cofactors such as 1 mM NADPH can be added on demand based on catalytic domains in PKS. Malonyl-CoA can also be generated *in situ* with 10 μM MatB, 20 mM sodium malonate, and 1 mM CoA (An & Kim, 1998).

3. Type III PKS-based alkyne-tagged polyketide biosynthesis

Unlike the mega-enzyme systems of type I PKSs, type III PKSs, as simple homodimers, catalyze the whole series of decarboxylation, condensation, and cyclization reactions with a single active site. Type III PKSs perform C–C bond forming reactions by iterative Claisen-type condensation of CoA thioesters and cyclization of the poly-β-keto intermediates, to produce pharmaceutically and biologically important aromatic polyketides, such as chalcone, stilbene, and curcumin. (Abe & Morita, 2010; Morita et al., 2011). In addition, Type III PKSs often exhibit remarkable substrate tolerance towards different acyl-CoA starter units (Yu et al., 2012), representing an excellent platform in generating novel polyketides and making starter unit engineering feasible.

3.1 Type III PKS selection and starter unit engineering

Type III PKSs typically use an acyl-CoA as a starter unit, it is thus essential to select a promiscuous PKS that recognizes both the alkyne-bearing acyl group and the ACP for elongation. An *in vitro* product formation assay described above can be used to screen for a suitable type III PKS compatible with both the chain length of the acyl group and the ACP (Zhu et al., 2015b). Followed is the description of how type III PKSs such as HsPKS1, ORAS and PKS18 for the biosynthesis of alkyne-tagged polyketides are selected and engineered.

After the discovery of JamABC, we first selected a type III PKS from the plant *Huperzia serrate* named HsPKS1 to assemble polyketide backbones using malonyl-CoA extender units. HsPKS1 is able to accept a variety of starter units, including aromatic and aliphatic (C_6–C_{10}) CoA thioesters (Morita et al., 2011), suggesting that the recognition of the 5-hexynoyl starter unit would be likely. Two plasmids encoding *jamABC* and *hspks1*, with each gene regulated by a T7 promoter, were transformed into the *E. coli* BAP1 strain. We excluded either *jamA*, *jamB*, *jamC* or *hspks1* individually to yield four control strains. The resulting strains were grown under the culture conditions described above. Analyses of these culture extracts by HPLC and LC/HRMS revealed the production of an alkyne-tagged polyketide **1** (figure 1b), with the retention time, UV spectrum and mass patterns precisely matching those of the standard, giving the yield of ~0.3 mg/L, which is comparable to the typical polyketide titer in *E. coli* using similar culture conditions (Zhang et al., 2008; Zhu et

al., 2015a). Compound **1** was presumably formed by condensing the 5-hexynoyl JamC with two malonyl-CoAs followed by spontaneous intramolecular cyclization. It is notable that, in addition to **1**, a major byproduct **2** bearing a terminal alkene functionality retained from the fed alkene precursors was also produced at a titer of 4–5 fold of that of **1** (Figure 1b).

The production of **1** by coexpressing *jamABC* and *hspks1* confirmed the function of this tri-gene cassette (*jamABC*) in the model organism *E. coli* and demonstrated the feasibility of synthesizing terminal alkyne-tagged natural products by starter unit engineering *de novo*. Exclusion of JamC completely abolished the production of **1**, which further confirmed the necessity of this carrier protein for the acetylenase activity of JamB. Deletion of JamA decreased the titer of **1** by ~30-fold, indicating that endogenous acyl-ACP synthetases of *E. coli* could complement the activity of JamA, albeit with a much lower efficiency.

Based on the successful application of HsPKS1 and the discovery of various JamABC homologues through genome mining, we extended the *E. coli* platform to include a variety of type III PKSs with the corresponding chain length specificity of the starter unit to generate more alkyne-tagged polyketides (Zhu et al., 2015b). Considering the substrate selectivity of JamA homologues, ranging from fatty acids of medium-chain length to long-chain length (C_6-C_{18}), we continued using the established platform of HsPKS1 pairing with the JamA homologues activating medium-chain fatty acids and searched for type III PKSs with activities toward 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* (Funa et al., 2007; Rubin-Pitel et al., 2008; Saxena et al., 2003; Yu, et al., 2010). 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 C_{18} fatty acyl starter units. The substrate tolerance of these type III PKSs toward ACPs was also investigated *in vitro*. Several lauryl-ACPs were prepared *in situ* using enzymatic reactions containing lauric acid, acyl-ACP synthetases, and 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 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. Consequently, we selected ORAS and PKS18 to replace the role of HsPKS1 in the *in vivo* system to generate alkyne-tagged polyketides containing long-chain fatty acyl tails.

To summarize, when selecting type III PKSs for the formation of alkyne-tagged polyketides, we need to consider both the tolerance of the fatty acyl chain length and the recognition of the ACP. Promiscuous type III PKSs that satisfy these two requirements can serve as candidates to be the downstream acceptors of the alkynyl-ACP starter unit.

3.2 Applications of type III PKS-based alkyne tagging

3.2.1 Discovery and characterization of JamABC homologues—

The *E. coli*-type III PKS platform has various applications besides the production of alkyne-tagged

polyketides. First, it can function as an *in vivo* reporting system to probe the substrate specificities of JamB and to discover other JamABC homologues to make terminal alkyne. Specifically, due to the low activity of JamB *in vitro*, the broad substrate specificity of HsPKS1 and JamA renders the engineered *E. coli* strain containing *jamABC* and *hspks1* an excellent reporting system for further probing the fatty acyl substrate specificity of JamB. We fed selected fatty acids, such as 4-pentenoic, 5-hexenoic acid, 6-heptenoic and 7-octenoic acids, to the *E. coli* cultures and analyzed culture extracts by both selected MS ion monitoring and MS-based comparative metabolomics between the engineered strain and the negative control strain without *jamB*. This extensive analysis revealed that the engineered strain failed to produce alkyne-tagged products using these precursors except 5-hexenoic acid, consistent with the *in vitro* biochemical results showing the strict substrate specificity of JamB (Zhu et al., 2015a). Furthermore, together with genome mining, this *in vivo* reporting system was used to discover and reconstitute the activities of JamABC homologues with the promiscuous type III PKSs HsPKS1, ORAS, or PKS18. TtuABC from *Teredinibacter turnerae* T7901 was discovered accordingly. Compound **3** was produced with a titer of ~0.5 mg/L when the cultures containing *ttuABC* and *hspks1* were fed with 9-decenoic acid, and exclusion of *ttuB* completely abolished the production of **3** (Figure 1c). Compound **3** was presumably assembled by the type III PKSs which catalyze the condensation of the 9-decynoyl moiety generated by TtuABC with two malonyl-CoAs, followed by spontaneous intramolecular cyclization. In addition to **3**, a major byproduct **4** that retains a terminal alkene functionality from the 9-decenoyl precursor was observed, which was similar to the results of JamABC with HspKS1. Additionally, deletion of *ttuA* decreased the titer of **3** by ~100-fold, similar to the behavior of the *jamABC* system. To our surprise, omitting *ttuC* did not completely abolish the production of **3**; instead, the titer of **3** was decreased by ~100-fold. Furthermore, the omission of both *ttuA* and *ttuC* still resulted in producing a trace amount of **3**. 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. Moreover, when the cultures were fed with decanoic acid, products **3**, **4**, and **5** were produced in a ~3:1:3 ratio in *E. coli* (Figure 1c), compared to JamB by which terminal alkyne was formed undetectably if fed with hexanoic acid. These observations indicated that although both JamB and TtuB are bifunctional desaturase/acetylenase to catalyze the terminal alkyne formation in a stepwise manner, TtuB had a much higher activity of acetylenase than JamB in *E. coli*. Regarding the substrate specificity of TtuB, the results from the *in vivo* reporting system demonstrated that TtuB has a strict chain length specificity toward C_{10} (Zhu et al., 2015b). Apart from HsPKS1, ORAS and PKS18 can be used to study the activities of JamB homologues that recognize long fatty acyl chains.

3.2.2 Optimization of JamB activity—Second, the *E. coli*-type III PKS system can be utilized to optimize the activities of JamB homologues coupled with a fluorogenic platform. To obtain a more efficient enzymatic tool for alkyne biosynthesis, random mutagenesis was used to optimize the desaturase/acetylenase activity of JamB in *E. coli*. We constructed a screening method based on the quantification of alkyne-tagged metabolites directly in their producing cell cultures in the *E. coli*-type III PKS system to quickly

assess the protein engineering outcomes of JamB. The design of the quantification platform has two components: the tagging of secreted polyketides through engineered biosynthesis with an alkyne moiety, followed by a click reaction with an azido fluorogenic probe. Fluorogenic probes are endowed with a functionality that suppresses fluorescence. Its transformation during the click reaction creates a new functionality that no longer quenches the fluorescence of the underlying system, resulting in a fluorescence enhancement (Shieh et al., 2015; Zhu et al., 2016). Such probes offer significant advantages over fluorescent probes for a quick readout of a tagged metabolite when washing away unreacted probes is not feasible. In particular, the reaction of an azido fluorogenic probe with alkyne metabolites in cell cultures is expected to result in a fluorescence enhancement that is dependent on the amount of these metabolites. The quantification of the alkynes works in both large-scale shake flasks and small-scale 96-well culture plates. Appropriate azido fluorogenic probes should be selected first to consider reaction specificity, selectivity, and sensitivity in complex cell cultures for extracellular metabolite quantification. Among all the probes tested, **6** gave the most consistent and reliable response and was thus chosen for further optimization of JamB (Figure 1d). We next constructed a library of JamB variants by diversifying *jamB* randomly using error-prone PCR. The mutant library of *jamB* was introduced into *E. coli* BAP1 strain, coexpressing *jamC*, *jamA*, and *hspks1*. After screening ~1000 clones, we isolated three clones that showed higher fluorescence signals than the control strain expressing the wild-type *jamB*. Further LC-HRMS quantification confirmed increased titers of **1** from these mutants. To examine the contributions of each amino acid mutation, we constructed the corresponding single mutant through site-directed mutagenesis. The best mutant, JamB-M5T, increased efficiencies in producing **1** by ~20-fold, compared to that of the wild-type upon the feeding of hexanoic acid. It is also notable that JamB mutants seem to be expressed at the level comparable to the wild-type enzyme based on the SDS-PAGE analysis. Our comparison of JamB with its homologues CamB (Zhu et al., 2015a) (identity/similarity = 95%/98%) and TtuB (identity/similarity = 44%/60%) also suggested that their activities in *E. coli* were affected by the amino acids in the non-homologous *N*-terminal region. This was further confirmed by the construction of chimeric JamB mutants with the *N*-terminal region replaced by the one from the more efficient desaturase/acetylenase TtuB or the less efficient CamB. The resulting JamB–TtuB and JamB–CamB chimeras showed increased and decreased activity, respectively, which support the sensitivity of enzymatic activity toward the *N*-terminal amino acids by enhancing the binding of JamB to ferredoxin, the potential electron donor, since the crystal structure of the membrane-bound homologous desaturase SCD1 suggests that the N-terminus lies along the electron donor-binding groove (Bai et al., 2015).

3.2.3 Probing the interactions between JamB and ACPs—Third, the *E. coli*-type III PKS platform can serve as a facile reporting system when studying the interactions between the desaturase/acetylenase and the ACP. As is mentioned above, JamB has very stringent substrate specificity towards JamC, suggesting the protein-protein interactions between the desaturase/acetylenase and the ACP may significantly affect the efficiency of the formation of the alkyne-tagged products. We first studied the ACP recognition of JamA by *in vitro* ACP loading assays (Su et al., 2018). JamC homologues (TtuC, CyACP, BpACP, and PeACP) found in putative alkyne biosynthetic machinery exhibited a high

degree of compatibility with JamA. Then we probed the ACP interactions with JamB based on the established *in vivo* reporting system with HsPKS1 due to difficulties in working with membrane-bound proteins *in vitro*. JamB activities toward TtuC and CyACP decreased 2–3 fold compared to JamC based on the formation of **1**. No reliable production of **1** was observed from the culture extracts of the strains that co-expressed *bpACP* or *peACP*, suggesting weak/no recognition of JamB with these two ACPs. To improve ACP recognition by JamB, we rationally engineered ACP based on the predominant electrostatic interactions between the ACP and the soluble stearoyl-CoA desaturase (Guy et al., 2011). We mutated the key amino acid residue on helix 3 of TtuC, BpACP and PeACP according to the results of sequence alignment and structural modeling. TtuC-K69E mutant displayed higher alkyne production efficiency comparable to JamC, indicating the electrostatic interactions play a critical role in recognition of ACP by JamB. However, the rational engineering of BpACP and PeACP failed to improve the recognition by JamB, suggesting that engineering the electrostatic interactions was not sufficient for proper interactions between JamB and these ACPs. Error-prone PCR was thus chosen to evolve PeACP. Hot spots located right before PeACP helix 2 or at the beginning of PeACP helix 3 were identified to be important for the recognition by JamB. The engineered PeACP had ~16-fold improved activity of producing terminal alkyne compared to the wild type.

Overall, the PKS starter unit engineering is a feasible strategy to install the fatty alkynyl starter unit generated by JamABC onto polyketide scaffolds, such as those generated by promiscuous type III PKSs, which recognize both the acyl group and the acyl carrier (JamC). The platform for the biosynthesis of alkyne-tagged polyketides is a useful *in vivo* activity reporting system for characterization of new enzyme homologues and engineering of the JamABC biosynthetic machinery. However, to generalize this strategy to other polyketide scaffolds, particularly those synthesized by modular type I PKSs, additional model systems and protein engineering methods need to be explored, which will be discussed below.

4. Type I PKS-based alkyne-tagged polyketide biosynthesis

Type I PKSs are multifunctional enzymes containing linearly arranged and covalently fused domains with defined functions separated by short-space regions. The distinct domains work cooperatively to catalyze the carbon chain elongation and functional group modification, and act as an assembly line to select and incorporate building blocks into polyketide scaffolds (Porterfield et al., 2020; Porterfield & Zhang, 2020; Wang et al., 2020). The starter units of type I PKSs can be various acyl-ACP complexes typically incorporated by the starting acyltransferase (AT) domain and ACP into assembly lines (Sattely et al., 2008). The building blocks used for extension, typically malonyl- or methylmalonyl-CoA, are recognized by AT domains, and the carbon-carbon bond is formed through decarboxylative Claisen condensations catalyzed by the ketosynthase (KS) domains. The mega-synthases themselves have been investigated and have undergone extensive engineering efforts due to their modular structures that have captured scientists' imagination with the possibility of producing on-demand, designed molecules (Wlodek et al., 2017). Many of these engineering strategies have included efforts geared towards the inclusion of functional chemical handles for subsequent drug discovery or chemical biology studies (Kalkreuter

et al., 2019; Koryakina et al., 2017). The *de novo* biosynthesis of alkyne-tagged polyketides can potentially be achieved using JamABC to provide a building monomer, and several engineering strategies will be described below.

4.1 Type I PKS selection and starter unit engineering

The hexanoyl-JamC moiety is naturally incorporated into type I PKS assembly line as a starter unit to generate jamaicamide B (Edwards et al., 2004). To explore a generalized method of starter unit engineering, we employed two well-studied type I PKSs, LipPKS1 and DEBSM6, to explore starter unit engineering strategies to make alkyne-tagged polyketides. LipPKS1 is the first module in lipomycin biosynthesis that natively utilizes an isobutyl starter unit presented by a loading ACP (Bihlmaier et al., 2006). DEBSM6 is the last PKS module from the erythromycin biosynthetic pathway (Rawlings, 2001). In addition, engineered LipPKS1 and DEBSM6 have been obtained to utilize malonyl-CoA instead of methylmalonyl-CoA as the extender unit with the promiscuous DEBS thioesterase to promote the acid product release as demonstrated from both *in vitro* biochemical studies and in *E. coli* (Yuzawa et al., 2017). These two engineered modules are thus simple and convenient systems for in-depth assessment of the interaction between representative module PKSs and JamABC for alkyne-tagged polyketide biosynthesis. Considering the known critical role of the cognate ACP (JamC) in the alkyne biosynthetic machinery (Su et al., 2018), the recognition of JamC by PKSs is expected to play a critical role in alkyne-tagged polyketide synthesis and therefore is the focus of the engineering.

To probe the possible recognition of the 5-hexynoyl-JamC by PKSs, *in vitro* assays with 5-hexynoic acid were initially performed using the engineered LipPKS1 and DEBSM6 modules. The reported engineered LipPKS1 was further modified by removing the AT and ACP loading domains to create a truncated version to facilitate the alternative starter unit incorporation. We hypothesized that these engineered PKSs (termed LipPKS1* and DEBSM6*, Figure 2a) without the loading domains would result in JamC acting in *trans* to selectively load and extend JamC-linked acyl chains. JamB activity for alkyne biosynthesis was assessed *in vivo* due to the difficulty of obtaining active and purified membrane proteins. The expected product, 3-hydroxy-7-octynoic acid (**7**), was successfully produced by both engineered PKSs (Figure 2b). Interestingly, replacement of 5-hexynoyl-JamC by 5-hexynoyl-CoA, which was generated *in situ* using a promiscuous acyl-CoA ligase Orf35 (Zhang, Bolla, Kahne, & Walsh, 2010), dropped the formation of **7** to trace amounts, demonstrating a preference of these two PKSs toward JamC over CoA as the acyl carrier.

4.1.1 Docking Domain Strategy to Improve JamC-PKS Interactions—

Engineering strategies should be mainly focused on increasing the communication and recognition between JamC and type I PKS to ensure the translocation efficiency. First, as protein-protein interactions are known to dominate the turnover of chimeric PKS assembly lines (Klaus et al., 2016), improved communication between the upstream JamC and the downstream KS could lead to a more efficient alkyne-tagged polyketide biosynthesis. Docking domains, often found on the *C* terminus of ACPs (dd^{ACP}) and the *N* terminus of KSs (dd^{KS}), have been shown to be important for protein-protein interactions in PKSs (Gokhale et al., 1999; Tsuji et al. , 2001; Zeng et al., 2016). We then fused known docking

domains to the *C* terminus of JamC and the *N* terminus of the LipPKS1*/DEBSM6* KS domains to improve protein recognition. In particular, the class 2 docking domains from the cyanobacterial curacin pathway were chosen as the pair $_{C}dd^{CurK}$ (dd^{ACP}) and $_{N}dd^{CurL}$(dd^{KS}) was shown to be modular and portable (Whicher et al., 2013). The related docking domain pair $_{C}dd^{JamK}$ (dd^{ACP}) and $_{N}dd^{JamL}$(dd^{KS}) from the original jamaicamide pathway was also a candidate as adopting docking domains within pathways was shown to be more successful than inter-pathway swapping (Klaus & Grininger, 2018; Klaus et al., 2016; Whicher et al., 2013).

The fusion of these docking domains to JamC and PKSs did not significantly impact the expression and folding of these proteins. *In vitro* product formation assays demonstrated the success of this strategy in generating product **7**. The adoption of the pair of $_{C}dd^{JamK}$ and $_{N}dd^{JamL}$ had minimal effect on the production of **7**, whereas the pair of $_{C}dd^{CurK}$ and $_{N}dd^{CurL}$ led to significantly more amount of **7** in both PKS systems (3-fold for LipPKS1* and 40-fold for DEBSM6*). Control experiments using only one of the docking domains produced fewer products than using the pair for $_{C}dd^{CurK}$ and $_{N}dd^{CurL}$. In addition, the poor production of **7** with the docking domain fused to JamC excluded the possibility of improved recognition of modified JamC by JamA, indicating that the improved communication between the engineered JamC and KS due to the docking domains is the main contributor for higher production of **7** *in vitro*.

4.1.2 Site-Directed Mutagenesis of JamC to Improve JamC-PKS Interactions

—In addition to docking domain engineering, a less-intensive engineering strategy such as site-directed mutagenesis to improve JamC-PKS communication was studied. Mutating JamC without perturbation to the large mega-synthase would make this strategy more easily adaptable to different systems. From the well-studied DEBS system, it has been shown that direct ACP-KS protein-protein interactions during translocation are selective, and key residues within helix I of ACP have been identified that contribute to chain translocation specificity (Kapur et al., 2012; Klaus et al., 2016). Inspired by the previous successful studies, we identified the corresponding residue in JamC (E32) that may play an important role in ACP-KS interactions through sequence alignments and structural modeling.

To mimic the native upstream ACP, we chose the mutations E32T for LipPKS1* and E32H for DEBSM6* based on alignments to the respective ACPs found upstream in the native systems. These two JamC mutants were cloned, overexpressed, and purified from BAP1 with a similar yield to the wild-type protein. *In vitro* product formation assays showed that the formation of **7** increased approximately 7-fold with LipPKS1* and 2-fold with DEBSM6*. These fold increases demonstrated the effectiveness of this strategy in improving the production of alkyne-tagged polyketides *in vitro*, most likely due to an improved JamC communication with modular PKSs.

4.1.3 Perturbation of JamB Activity by JamC Engineering

—Besides these two engineering strategies, the stronger protein-protein interaction between JamB and JamC also needs to be considered to retain the activity of terminal alkyne formation. *In vitro* biochemical assays demonstrated the success of protein engineering in improving the recognition of JamC by PKSs to promote the translocation of the alkynyl starter unit.

However, the potential impact of JamC modification on the activity of JamB, the desaturase/acetylenase that functions on a JamC-tethered substrate to form a terminal alkyne, is unclear. As it is difficult to reconstitute and quantify the activity of the membrane-bound JamB *in vitro*, we then tried to implement the biosynthetic machinery of alkyne-tagged polyketides in *E. coli* to assess the possible impact. In addition, the titers of relevant products were also quantified in *E. coli* to probe the effectiveness of two engineering strategies to improve JamC-PKS interactions *in vivo*. A single mutation in JamB (M5T) identified in previous work was used in all strains to increase the alkyne titer in *E. coli* (Zhu et al., 2016). All strains were grown supplemented with 5-hexenoic acid, followed by extraction and quantification of 3-hydroxy-7-octenoic acid (**8**) and 3-hydroxy-7-octynoic acid production (**7**), by fitting to a standard curve of synthesized standards generated through LC-HRMS analysis. Product **8** was expected to be a side product due to the activities of JamA, C, and PKS without the action of JamB (Figure 2b). Other possible products were also analyzed, as it is conceivable that the PKSs accept different fatty acyl starter units *in vivo* via JamC or other acyl carriers.

An initial investigation of the titer of compound **7** produced by the co-expression of JamA, B, C, and LipPKS1*/DEBSM6* demonstrated that DEBSM6* produced compound 1 (0.014 mg/L) significantly less than LipPKS1* (0.071 mg/L). Much higher amounts of products other than **7** and **8** with a longer acyl chain were generated by DEBSM6* *in vivo*, consistent with the native acyl chain length accepted by LipPKS1*/DEBSM6*. Therefore, DEBSM6* would not be an effective *in vivo* model system to probe the activity of JamB due to complicated product profiles and thus limited the *in vivo* study to LipPKS1*.

The products **7** and **8** were produced by LipPKS1* in an approximately 1:5 ratio, and this efficiency was set to be a relative JamB activity of 100%. This product ratio was dropped 4-fold when either the docking pair was used, suggesting that the fusion of a docking domain to JamC affected its recognition by JamB. In contrast, the E32T point mutation of JamC had minimal effect on the product ratio while increasing the titer of **7** 6-fold to 0.42 mg/L, consistent with previous observations that the helix I of ACP did not play an important role in interacting with JamB (Su et al., 2018). The combined product titer of alkyne **7** and alkene **8** was probed to assess the effectiveness of the two engineering strategies in their ability to improve JamC-LipPKS1* interactions *in vivo*. Consistent with the trends observed *in vitro*, the combined titer improved more than 10- and 20-fold using docking domains $_{C}dd^{CurK}/_{N}dd^{CurL}$ and $_{C}dd^{JamK}/_{N}dd^{JamL}$, respectively, and 10-fold using JamC (E32T), demonstrating the success of either strategy in improving JamC-PKS interactions *in vivo*.

Finally, the two engineering strategies (Figure 2c) have synergistic effects in improving the alkyne-tagged polyketide biosynthesis *in vivo*. We observed additive effects of docking domains and the JamC point mutation in improving JamC-LipPKS1* interactions. The combined titer of **7** and **8** roughly equaled the sum of that when either engineering strategy was used. The maximum amount of product obtained was 16 mg/L from JamC(E32T)-$_{C}dd^{JamK}/_{N}dd^{JamL}$-LipPKS1*, an approximately 39-fold increase from unmodified JamC/LipPKS1*. However, due to the expected disruption of JamB activity when the docking domain is fused to JamC, the absolute titer of the alkyne product **7** was not increased when

using both engineering strategies compared with the JamC mutagenesis alone. These results further highlight the importance of JamB efficiency in *de novo* alkyne synthesis, which remains a limiting step in the production of alkyne-tagged polyketides.

4.1.4 Type I PKS selection—Although the above results demonstrate a great potential of *de novo* biosynthesizing alkyne-tagged polyketides by engineering both the alkyne biosynthetic machinery and modular type I PKSs, the strategy is limited to incorporate an alkynyl acyl group, which needs to be tolerated by PKSs. It is expected to work well with PKSs with a native starter unit resembling the alkyne-containing acyl group presented by the alkyne biosynthetic machinery, such as in the case of LipPKS1, but may not work with PKSs recognizing very different starter units, such as in the case of DEBSM6. This is particularly exemplified by the *in vivo* results of DEBSM6, in which a complex metabolic background significantly decreased the efficiency of alkyne-tagged polyketide biosynthesis by these PKSs. Therefore, when selecting type I PKSs for the *de novo* biosynthesis of alkyne-tagged polyketides, we need to consider the compatibility of the acyl group generated by the alkyne biosynthetic machinery in addition to the ACP.

4.2 Type I PKS extender unit engineering

The extender unit engineering is an alternative strategy for incorporating an alkynoic extender unit into the molecular scaffolds of PKs besides starter unit engineering. Although most polyketides use malonyl-CoA or methylmalonyl-CoA as extender units, longer-chain fatty acyl extender units have been identified in a few modular assembly lines (Wilson & Moore, 2012). We applied the JamABC cassette to the biosynthetic machinery for antimycin-type depsipeptides, where a minimum set of enzymes (AntCDEFGM) is required for the generation of the antimycin dilactone scaffold (Zhu et al., 2015a; Sandy et al., 2012). The hybrid NRPS/PKS assembly line features the formation and incorporation of atypical PKS extender units by the actions of a reductase/carboxylase homologue AntE and an AT domain embedded in AntD. It has been shown that an alkynoic extender unit can be recognized by the promiscuous AntE and AntD-AT to synthesize an alkyne-labeled antimycin by PDB in *Streptomyces* (Yan et al., 2013). To eliminate the requirement of alkynoic precursors and *de novo* biosynthesize alkyne-tagged antimycin analogs, we coexpressed *antCDEFGM* and *jamABC* in the *E. coli* BAP1 strain, which led to the production of an alkyne-bearing antimycin analog **9** (Figure 3). As expected, the deletion of *jamB* or *antC* completely abolished the production of **9**. The biosynthesis of **9** suggested that 5-hexynoyl-JamC could enter *E. coli* fatty acid metabolic pathways for additional processing. However, the extender unit engineering of incorporating alkynoic moiety is overall not efficient for a few reasons. First, the PKS assembly line must have promiscuous catalytic domains to recognize the alkynyl acyl chain as an extender unit. More importantly, the alkynyl-ACP would need to enter fatty acid metabolic pathways to undergo a series of modifications, such as offloading from ACP to CoA and β-desaturation to be recognized by a reductase/carboxylase homologue to generate a malonyl-CoA derivative.

5. Perspectives

The ACP-dependent terminal alkyne biosynthetic pathway has the potential to be employed to tag a broad spectrum of biomolecules through biosynthetic pathway engineering to circumvent the need of alkyne precursor feeding. The major advantage of this biosynthetic machinery is eliminating the diffusible alkynyl group with the alkyne functionality tethered on ACPs, ensuring the precise tagging of polyketides. Even though only limited accomplishment has been achieved and it remains a challenge to achieve diverse biosynthesis of alkyne-tagged natural products, whose achievement relies on the cross talk of terminal alkyne biosynthetic machinery with biosynthetic pathways of natural products of interest, the successful starter unit and extender unit engineering of PKS and PKS/NRPS build a milestone in the development of *de novo* production of alkyne-tagged polyketides. Starter unit engineering could be generalized to tag polyketides and lipopeptides that naturally contain fatty acyl starter units, and extender unit engineering could lead to the site-selective introduction of a terminal alkyne into additional PKs or PK/NRP hybrids inside living cells. In particular, the engineering strategies for truncated type I PKSs, docking domain engineering, and site-directed mutagenesis, demonstrated the capability of the carrier protein-dependent alkyne biosynthetic machinery working as a *trans*-loading system for downstream non-cognate PKSs, which can serve as generalized methods for the engineering of other PKSs. The key is to enhance the recognition of the ACP (e.g., JamC) by PKSs whereas maintain the recognition of the ACP by the membrane bifunctional desaturase/acetylense. Meanwhile, the preference of downstream PKSs to the alkynyl acyl group is also important, in particular for *in vivo* production.

To address the challenges, it is helpful to discover more JamABC homologues in generating different alkynyl-ACPs to suit different PKS pathways. Additionally, investigating the mechanism of JamB homologues for terminal alkyne catalysis can help improve their activity to decrease the amount of side product formation and hopefully broaden their specificities towards ACPs to diminish the limitations of ACP engineering. Moreover, other engineering strategies, for example, domain swapping of PKSs, can be explored to change the substrate specificity of PKSs toward alkynyl moieties to generate more alkyne-tagged polyketides.

Acknowledgements

The relevant research in Zhang lab was financially supported by grants to W.Z. from the NIH (R01GM136758 and DP2AT009148), the American Cancer Society (Grant RSG-17-013-01-CDD), and the Chan Zuckerberg Biohub Investigator Program. The authors thank Dr. Rui Zhai for the article proofreading.

References

Abe I, & Morita H (2010). Structure and function of the chalcone synthase superfamily of plant type III polyketide synthases. Natural Product Reports, 27(6), 809–838. [PubMed: 20358127]

Alonso H, & Roujeinikova A (2012). Characterization and Two-Dimensional Crystallization of Membrane Component AlkB of the Medium-Chain Alkane Hydroxylase System from Pseudomonas putida GPo1. Applied and Environmental Microbiology, 78(22), 7946–7953. [PubMed: 22941083]

An JH, & Kim YS (1998). A gene cluster encoding malonyl-CoA decarboxylase (MatA), malonyl-CoA synthetase (MatB) and a putative dicarboxylate carrier protein (MatC) in Rhizobium trifolii. European Journal of Biochemistry, 257(2), 395–402. [PubMed: 9826185]

Awakawa T, Fujioka T, Zhang L, Hoshino S, Hu Z, Hashimoto J, et al. (2018). Reprogramming of the antimycin NRPS-PKS assembly lines inspired by gene evolution. Nature Communications, 9(1), 3534.

Bai Y, McCoy JG, Levin EJ, Sobrado P, Rajashankar KR, Fox BG, et al. (2015). X-ray structure of a mammalian stearoyl-CoA desaturase. Nature, 524, 252–256. [PubMed: 26098370]

Barajas JF, Blake-Hedges JM, Bailey CB, Curran S, & Keasling JD (2017). Engineered polyketides: Synergy between protein and host level engineering. Synthetic and Systems Biotechnology, 2(3), 147–166. [PubMed: 29318196]

Bihlmaier C, Welle E, Hofmann C, Welzel K, Vente A, Breitling E, et al. (2006). Biosynthetic Gene Cluster for the Polyenoyltetramic Acid α-Lipomycin. Antimicrobial Agents and Chemotherapy, 50(6), 2113–2121. [PubMed: 16723573]

Cottens S, Kallen J, Schuler W, & Sedrani R (2019). Derivation of rapamycin: Adventures in natural product chemistry. Chimia, 73(7–8), 581–590. [PubMed: 31431218]

DeGuire SM, Earl DC, Du Y, Crews BA, Jacobs AT, Ustione A, et al. (2015). Fluorescent Probes of the Apoptolidins and their Utility in Cellular Localization Studies. Angewandte Chemie International Edition, 54(3), 961–964. [PubMed: 25430909]

Edwards DJ, Marquez BL, Nogle LM, McPhail K, Goeger DE, Roberts MA, et al. (2004). Structure and biosynthesis of the Jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium Lyngbya majuscula. Chemistry and Biology, 11(6), 817–833. [PubMed: 15217615]

Funa N, Awakawa T, & Horinouchi S (2007). Pentaketide Resorcylic Acid Synthesis by Type III Polyketide Synthase from Neurospora crassa. Journal of Biological Chemistry, 282(19), 14476–14481. [PubMed: 17374612]

Gokhale RS, Tsuji SY, Cane DE, & Khosla C (1999). Dissecting and Exploiting Intermodular Communication in Polyketide Synthases. Science, 284(5413), 482–485. [PubMed: 10205055]

Gustafsson C, Govindarajan S, & Minshull J (2004). Codon bias and heterologous protein expression. Trends in Biotechnology, 22(7), 346–353. [PubMed: 15245907]

Guy JE, Whittle E, Moche M, Lengqvist J, Lindqvist Y, & Shanklin J (2011). Remote control of regioselectivity in acyl-acyl carrier protein-desaturases. Proceedings of the National Academy of Sciences, 108(40), 16594–16599.

Hertweck C (2009). The Biosynthetic Logic of Polyketide Diversity. Angewandte Chemie International Edition, 48(26), 4688–4716. [PubMed: 19514004]

Jelić D, & Antolović R (2016). From Erythromycin to Azithromycin and New Potential Ribosome-Binding Antimicrobials. Antibiotics, Vol. 5, p. 29. [PubMed: 27598215]

Kalkreuter E, Keeler AM, Malico AA, Bingham KS, Gayen AK, & Williams GJ (2019). Development of a Genetically Encoded Biosensor for Detection of Polyketide Synthase Extender Units in Escherichia coli. ACS Synthetic Biology, 8(6), 1391–1400. [PubMed: 31134799]

Kalkreuter E, & Williams GJ (2018). Engineering enzymatic assembly lines for the production of new antimicrobials. Current Opinion in Microbiology, 45, 140–148. [PubMed: 29733997]

Kapur S, Lowry B, Yuzawa S, Kenthirapalan S, Chen AY, Cane DE, & Khosla C (2012). Reprogramming a module of the 6-deoxyerythronolide B synthase for iterative chain elongation. Proceedings of the National Academy of Sciences, 109(11), 4110–4115.

Klaus M, & Grininger M (2018). Engineering strategies for rational polyketide synthase design. Natural Product Reports, 35(10), 1070–1081. [PubMed: 29938731]

Klaus M, Ostrowski MP, Austerjost J, Robbins T, Lowry B, Cane DE, et al. (2016). Protein-Protein Interactions, Not Substrate Recognition, Dominate the Turnover of Chimeric Assembly Line Polyketide Synthases. Journal of Biological Chemistry, 291(31), 16404–16415. [PubMed: 27246853]

Koryakina I, Kasey C, McArthur JB, Lowell AN, Chemler JA, Li S, et al. (2017). Inversion of Extender Unit Selectivity in the Erythromycin Polyketide Synthase by Acyltransferase Domain Engineering. ACS Chemical Biology, 12(1), 114–123. [PubMed: 28103677]

Marchand JA, Neugebauer ME, Ing MC, Lin C-I, Pelton JG, & Chang MCY (2019). Discovery of a pathway for terminal-alkyne amino acid biosynthesis. Nature, 567(7748), 420–424. [PubMed: 30867596]

McDaniel R, Ebert-Khosla S, Fu H, Hopwood DA, & Khosla C (1994). Engineered biosynthesis of novel polyketides: influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. Proceedings of the National Academy of Sciences, 91(24), 11542–11546.

Morita H, Yamashita M, Shi S-P, Wakimoto T, Kondo S, Kato R, et al. (2011). Synthesis of unnatural alkaloid scaffolds by exploiting plant polyketide synthase. Proceedings of the National Academy of Sciences, 108(33), 13504–13509.

Moschny J, Lorenzen W, Hilfer A, Eckenstaler R, Jahns S, Enke H, et al. (2020). Precursor-Directed Biosynthesis and Fluorescence Labeling of Clickable Microcystins. Journal of Natural Products, 83(6), 1960–1970. [PubMed: 32464061]

Musiol-Kroll EM, Zubeil F, Schafhauser T, Härtner T, Kulik A, McArthur J, et al. (2017). Polyketide Bioderivatization Using the Promiscuous Acyltransferase KirCII. ACS Synthetic Biology, 6(3), 421–427. [PubMed: 28206741]

Nallamsetty S, & Waugh DS (2006). Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. Protein Expression and Purification, 45(1), 175–182. [PubMed: 16168669]

Pfeifer BA, Admiraal SJ, Gramajo H, Cane DE, & Khosla C (2001). Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of E . coli. Science, 291(March), 1790–1793. [PubMed: 11230695]

Pfeifer B, Hu Z, Licari P, & Khosla C (2002). Process and Metabolic Strategies for Improved Production of Escherichia coli-Derived 6-Deoxyerythronolide B. Applied and Environmental Microbiology, 68(7), 3287–3292. [PubMed: 12089005]

Porterfield WB, Poenateetai N, & Zhang W (2020). Engineered Biosynthesis of Alkyne-Tagged Polyketides by Type I PKSs. IScience, 23(3), 100938. [PubMed: 32146323]

Porterfield WB, & Zhang W (2020). Mutasynthesis for Natural Product Bioengineering. In Comprehensive Natural Products III (3rd ed., Vol. 6).

Quadri LEN, Weinreb PH, Lei M, Nakano MM, Zuber P, & Walsh CT (1998). Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carder protein domains in peptide synthetases. Biochemistry, 37(6), 1585–1595. [PubMed: 9484229]

Rawlings BJ (2001). Type I polyketide biosynthesis in bacteria (Part A—erythromycin biosynthesis). Natural Product Reports, 18(2), 190–227. [PubMed: 11336289]

Rosano GL, & Ceccarelli EA (2014). Recombinant protein expression in Escherichia coli: advances and challenges. Frontiers in Microbiology, Vol. 5, p. 172. [PubMed: 24860555]

Rubin-Pitel SB, Zhang H, Vu T, Brunzelle JS, Zhao H, & Nair SK (2008). Distinct structural elements dictate the specificity of the type III pentaketide synthase from Neurospora crassa. Chemistry & Biology, 15(10), 1079–1090. [PubMed: 18940668]

Sandy M, Rui Z, Gallagher J, & Zhang W (2012). Enzymatic Synthesis of Dilactone Scaffold of Antimycins. ACS Chemical Biology, 7(12), 1956–1961. [PubMed: 22971101]

Sattely ES, Fischbach MA, & Walsh CT (2008). Total biosynthesis: In vitro reconstitution of polyketide and nonribosomal peptide pathways. Natural Product Reports, 25(4), 757–793. [PubMed: 18663394]

Saxena P, Yadav G, Mohanty D, & Gokhale RS (2003). A New Family of Type III Polyketide Synthases in *Mycobacterium tuberculosis*. Journal of Biological Chemistry, 278(45), 44780–44790. [PubMed: 12941968]

Seidel J, Miao Y, Porterfield W, Cai W, Zhu X, Kim S-J, et al. (2019). Structure–activity–distribution relationship study of anti-cancer antimycin-type depsipeptides. Chemical Communications, 55(63), 9379–9382. [PubMed: 31317975]

Shen J, Wu G, Tsai AL, & Zhou M (2020). Structure and Mechanism of a Unique Diiron Center in Mammalian Stearoyl-CoA Desaturase. Journal of Molecular Biology, 432(18), 5152–5161. [PubMed: 32470559]

Shieh P, Dien VT, Beahm BJ, Castellano JM, Wyss-Coray T, & Bertozzi CR (2015). CalFluors: A Universal Motif for Fluorogenic Azide Probes across the Visible Spectrum. Journal of the American Chemical Society, 137(22), 7145–7151. [PubMed: 25902190]

Staub I, & Sieber SA (2008). β-Lactams as Selective Chemical Probes for the *in vivo* Labeling of Bacterial Enzymes Involved in Cell Wall Biosynthesis, Antibiotic Resistance, and Virulence. Journal of the American Chemical Society, 130(40), 13400–13409. [PubMed: 18781750]

Studier FW, & Moffatt BA (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. Journal of Molecular Biology, 189(1), 113–130. [PubMed: 3537305]

Su M, Zhu X, & Zhang W (2018). Probing the acyl carrier protein–Enzyme interactions within terminal alkyne biosynthetic machinery. AIChE Journal, 64(12), 4255–4262. [PubMed: 30983594]

Tsuji SY, Cane DE, & Khosla C (2001). Selective Protein–Protein Interactions Direct Channeling of Intermediates between Polyketide Synthase Modules. Biochemistry, 40(8), 2326–2331. [PubMed: 11327852]

Wang H, Klein MG, Zou H, Lane W, Snell G, Levin I, et al. (2015). Crystal structure of human stearoyl–coenzyme A desaturase in complex with substrate. Nature Structural & Molecular Biology, 22(7), 581–585.

Wang J, Zhang R, Chen X, Sun X, Yan Y, Shen X, et al. (2020). Biosynthesis of aromatic polyketides in microorganisms using type II polyketide synthases. Microbial Cell Factories, 19(1), 1–11. [PubMed: 31898497]

Wei L, Hu F, Shen Y, Chen Z, Yu Y, Lin C-C, et al. (2014). Live-cell imaging of alkyne-tagged small biomolecules by stimulated Raman scattering. Nature Methods, 11(4), 410–412. [PubMed: 24584195]

Whicher JR, Smaga SS, Hansen DA, Brown WC, Gerwick WH, Sherman DH, et al. (2013). Cyanobacterial polyketide synthase docking domains: A tool for engineering natural product biosynthesis. Chemistry and Biology, 20(11), 1340–1351. [PubMed: 24183970]

Wilson MC, & Moore BS (2012). Beyond ethylmalonyl-CoA: The functional role of crotonyl-CoA carboxylase/reductase homologs in expanding polyketide diversity. Natural Product Reports, 29(1), 72–86. [PubMed: 22124767]

Wlodek A, Kendrew SG, Coates NJ, Hold A, Pogwizd J, Rudder S, et al. (2017). Diversity oriented biosynthesis via accelerated evolution of modular gene clusters. Nature Communications, 8, 1206.

Yan Y, Chen J, Zhang L, Zheng Q, Han Y, Zhang H, et al. (2013). Multiplexing of Combinatorial Chemistry in Antimycin Biosynthesis: Expansion of Molecular Diversity and Utility. Angewandte Chemie International Edition, 52(47), 12308–12312. [PubMed: 24123503]

Yu D, Xu F, Zeng J, & Zhan J (2012). Type III polyketide synthases in natural product biosynthesis. IUBMB Life, 64(4), 285–295. [PubMed: 22362498]

Yu D, Zeng J, Chen D, & Zhan J (2010). Characterization and reconstitution of a new fungal type III polyketide synthase from *Aspergillus oryzae*. Enzyme and Microbial Technology, 46(7), 575–580.

Yuzawa S, Deng K, Wang G, Baidoo EEK, Northen TR, Adams PD, et al. (2017). Comprehensive in Vitro Analysis of Acyltransferase Domain Exchanges in Modular Polyketide Synthases and Its Application for Short-Chain Ketone Production. ACS Synthetic Biology, 6(1), 139–147. [PubMed: 27548700]

Zeng J, Wagner DT, Zhang Z, Moretto L, Addison JD, & Keatinge-Clay AT (2016). Portability and Structure of the Four-Helix Bundle Docking Domains of trans-Acyltransferase Modular Polyketide Synthases. ACS Chemical Biology, 11(9), 2466–2474. [PubMed: 27362945]

Zhang W, Bolla ML, Kahne D, & Walsh CT (2010). A Three Enzyme Pathway for 2-Amino-3-hydroxycyclopent-2-enone Formation and Incorporation in Natural Product Biosynthesis. J. Am. Chem. Soc, 132(18), 6402–6411. [PubMed: 20394362]

Zhang W, Li Y, & Tang Y (2008). Engineered biosynthesis of bacterial aromatic polyketides in *Escherichia coli*. Proceedings of the National Academy of Sciences, 105(52), 20683–20688.

Zhu X, Liu J, & Zhang W (2015a). De novo biosynthesis of terminal alkyne-labeled natural products. Nature Chemical Biology, 11(2), 115–120. [PubMed: 25531891]

Zhu X, Su M, Manickam K, & Zhang W (2015b). Bacterial Genome Mining of Enzymatic Tools for Alkyne Biosynthesis. ACS Chemical Biology, 10, 2785–2793. [PubMed: 26441143]

Zhu X, Shieh P, Su M, Bertozzi CR, & Zhang W (2016). A fluorogenic screening platform enables directed evolution of an alkyne biosynthetic tool. Chemical Communications, 52(75), 11239–11242. [PubMed: 27561030]

Zhu X, & Zhang W (2018). Terminal Alkyne Biosynthesis in Marine Microbes. In Methods in Enzymology (1st ed., Vol. 604).

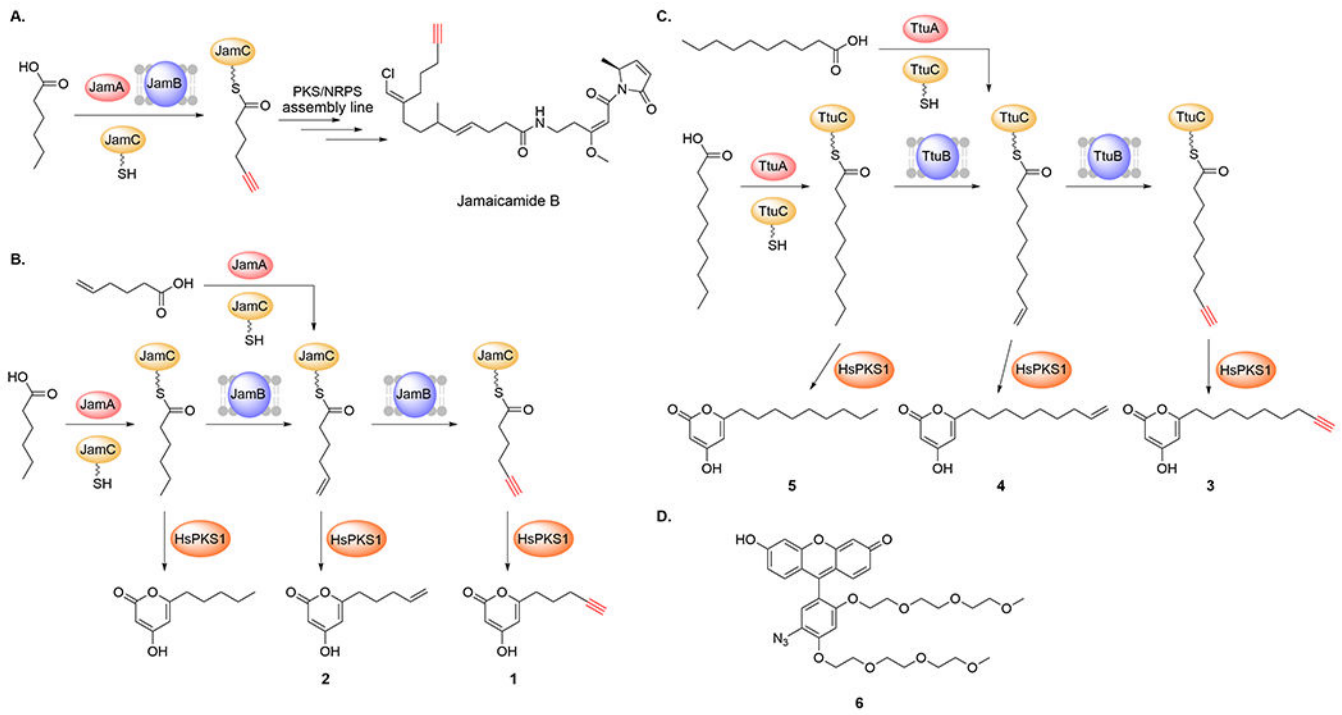

Figure 1.

Type III PKS starter unit engineering, **(a)** The natural pathway of JamABC in the biosysnthesis of Jamaicamide B. **(b)** The biosynthetic machinery of JamABC and starter unit engineering of HsPKS1 in generating alkyne-tagged polyketides. **(c)** The biosynthetic machinery of TtuABC and starter unit engineering of HsPKS1 in generating alkyne-tagged polyketides. **(d)** The azido fluorogenic probe used in optimization of JamB activity.

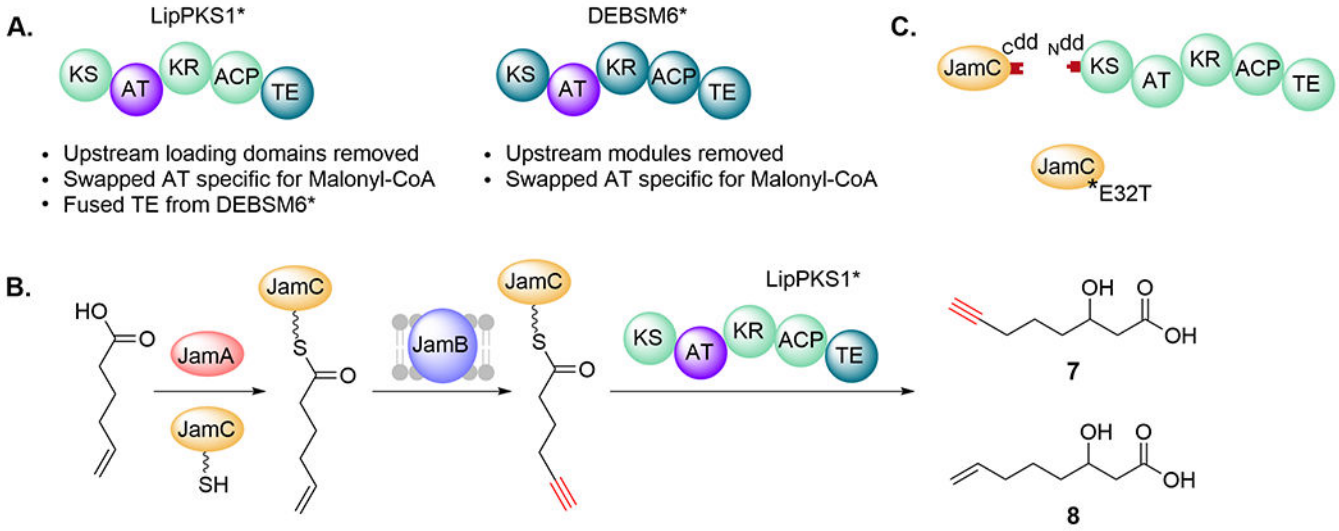

Figure 2.

Type I PKS starter unit engineering, **(a)** Engineered LipPKS1* and DEBSM6* used as representative type I modular PKSs for alkyne-tagged polyketide biosynthesis. KS, ketosynthase; AT, acyltransferase; KR, ketoreduction; ACP, acyl-carrier protein; TE, thioesterase. **(b)** Starter unit engineering of LipPKS1* in generating alkyne-tagged polyketides. **(c)** Two strategies in type I PKS starter unit engineering: docking domain adoption and site-directed mutagenesis.

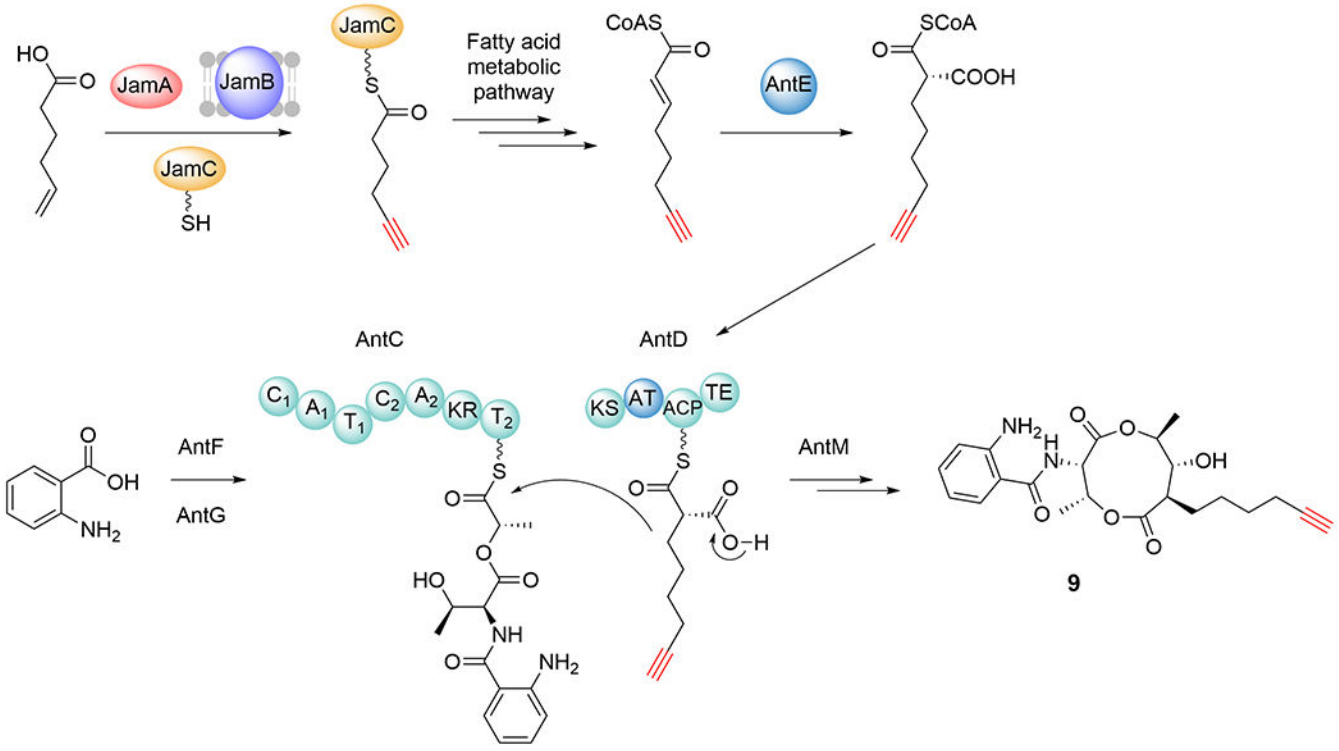

Figure 3.

Type I PKS extender unit engineering. Promiscuous AntE and AntD-AT enable the incorporation of alkyne-tagged extender unit. C, condensation; A, adenylation; T, thiolation.