Biosynthesis of alkyne-containing natural products

Xinyang Li, Jian-Ming Lv, Dan Hu and Ikuro Abe*

Alkyne-containing natural products are important molecules that are widely distributed in microbes and plants. Inspired by the advantages of acetylenic products used in the fields of medicinal chemistry, organic synthesis and material science, great efforts have focused on discovering the biosynthetic enzymes and pathways for alkyne formation. Here, we summarize the biosyntheses of alkyne-containing natural products and introduce de novo biosynthetic strategies for alkyne-tagged compound production.

1. Introduction

The alkyne is an important chemical functional moiety found in numerous natural products widely distributed in terrestrial bacteria, fungi, insects, and plants.1 Marine organisms, such as microbes, algae and sponges also represent rich sources for producing acetylenic natural products.2 Many of them display promising biological activities, such as anti-tumor, anti-parasitic, anti-malarial, and anti-HIV activities.1 In addition, the alkyne moiety is a valuable synthetic building block in organic chemistry. Conjugated enynes, including the 1,3-enyne motif and its derivatives found in many bioactive molecules,3,4 are important intermediates in the syntheses of organic compounds and macromolecules.4–7 The terminal alkyne is particularly useful for the alkyne–azide cycloaddition (“click” reaction), which enables the installation of a covalent bond of two chemical fragments to form complex molecular scaffolds for applications in pharmaceutical, cell biology and material sciences research.8–10 The alkyne can be used as an important bio-orthogonal tag and chemically or metabolically installed into biologically and medicinally important molecules, such as proteins, lipids, nucleic acids, and glycans.11 This handle will enable the visualization and quantification of the alkyne-tagged molecules through bio-orthogonal chemistry. Additionally, the modes of action...
and cellular dynamics of these alkyne-tagged molecules can be investigated through click chemistry or directly by stimulated Raman scattering microscopy.\textsuperscript{12,13}

Given the importance of alkynes in both chemistry and biology, chemical synthesis and precursor directed biosynthesis have been used for alkyne-tagged compound production.\textsuperscript{14–17} Although total and semi-synthesis are powerful tools for chemical production, they are often restricted by the yield of target compounds and the availability of starting materials. For the precursor directed biosynthesis, a set of structural analogues can be obtained by feeding precursor variants to an organism; however, it is often challenging to incorporate the desired precursors into the biosynthetic pathway due to the substrate specificity of the enzymatic machinery. In addition, both the shunt transformation pathway and co-existence of precursors and target compounds pose challenges to selectively probe the target compounds in the system, thus impeding broader applications.\textsuperscript{18} Alternatively, \textit{de novo} biosynthesis is another promising strategy for producing alkyne-tagged compounds, without feeding alkyne precursors. For this purpose, the biosynthesis of alkyne formation should be thoroughly understood. The discovery of the biosynthetic machineries for alkyne moieties has facilitated some illuminating progress in recent years. Hitherto, only a few types of genes have been identified as being involved in alkyne biosynthesis. Especially, the discovery of fatty acid/acyl acetylenases, a unique group of desaturases that catalyze alkyne formation in acetylenic fatty acids, is regarded as the first well studied example of the alkyne-forming enzymatic machinery.\textsuperscript{19,20} Furthermore, studies on the discovery of enzymes that catalyze alkyne formation in \beta-ethylserine and alkynylated mero-terpenoids have expanded the enzyme inventory and offered a potential route for the engineered biosynthesis of alkyne-tagged compounds.\textsuperscript{21–23} These findings have increased our understanding of the enzymatic machineries participating in acetylenic natural product generation. Previously, Minto et al.\textsuperscript{24} and Haritos et al.\textsuperscript{25} summarized the biosynthesis of polyacylenes and allied natural products. Herein, we update the recent research progress in the biosynthesis of alkyne-containing natural products and introduce \textit{de novo} biosynthesis for \textit{in situ} generating alkyne-tagged natural products.

2. Biosynthesis of acetylenic natural products

2.1 Fatty acids

Acetylenic fatty acids containing one or more C–C triple bonds are widespread metabolites in plants, fungi, marine sponges, and insects.\textsuperscript{26–28} Crepenynic acid (1), stearolic acid (2), and tariric acid (3) are general precursors to many acetylenic fatty acid-derived metabolites (Fig. 1A).\textsuperscript{29,30} Bio-synthetic studies of crepenynic acid from \textit{Crepis alpina} seed oil led to the characterization of the bifunctional enzyme Crep1, responsible for the double bond and subsequent acetylenic bond formation.\textsuperscript{31} The microsomal fraction from developing \textit{C. alpina} seeds converted \textsuperscript{14C}linoleic acid into \textsuperscript{14C}crepenynic acid in the presence of NADH, clearly establishing that the Δ12 triple bond formation employs an acyl chain with a double bond at Δ12 position as a precursor (Fig. 1B). Furthermore, the \textit{in vivo} heterologous expression of the \textit{crep1} gene in \textit{Saccharomyces cerevisiae} (YN94-1 strain) afforded results consistent with those from the \textit{in vitro} experiment. The results reflected the recognition of a special fatty acid desaturase (FAD), designated as an acetylenase, which is capable of catalyzing the conversion of olefinic bonds to acetylenic bonds and is distinct from enzymes that only convert single bonds to double bonds.\textsuperscript{28,31,32}

Although the exact triple bond-forming mechanism by the acetylenase remains unclear due to the lack of a membrane desaturase crystal structure, the desaturase-mediated
formation of a double bond in a fatty acid is reportedly diiron-dependent,33–36 thus providing an important reference for understanding the triple C–C bond-forming reaction catalyzed by Crep1. Crep1 is composed of 375 amino acid residues and shares high sequence similarity to a widely distributed group of membrane-bound, non-heme diiron enzymes in living organisms.15 As shown in Fig. 2, the selected sequences include mouse stearoyl-CoA desaturases 1 and 3, and other fatty acid desaturases from mammals, yeast, Arabidopsis thaliana, etc. Among them, the crystal structure of mouse stearoyl-CoA desaturase 1 (SCD1) (protein data bank ID 4YMK), an integral membrane protein, clearly shows the dizinc-center with longer distance between the two metals which is different from previously solved structures of soluble di-iron enzymes.33,37 The other obvious difference is the absence of glutamate residue in SCD1 which serves as bridging ligand of the two iron ions in the reported soluble enzymes.33 As Zn$^{2+}$ has an ionic radius of 0.88 Å, it may serve as a reasonable substitute for Fe$^{2+}$ (0.92 Å) and didn’t alter the structure of SCD1.33 From the sequence alignment results (Fig. 2), the histidine motifs for di-iron center ligation in SCD1 are consistent with the reported desaturases. In addition to the two HxxHH and HxxxxH motifs (marked in red frames) that are also conserved in Crep1, there are two other potential histidine-rich motifs (HHTH and HVMHH,
shadowed in red), indicating that Crep1 possesses a similar manner of metal chelation.

SCD1 catalyzes the C9–C10 cis-double bond formation from steroyl-CoA to oleyl-CoA.33 This step requires oxygen, which is activated by the metal center. The electrons needed for the desaturation reaction are obtained from cytochrome b₅, which in turn gained electrons from NAD[P]H via cytochrome b₅ reductase.33 Fatty acid desaturation in plants also relies on cytochrome b₅ reductase to transfer electrons.33 Crep1 may undergo two rounds of oxidation to catalyze the triple-bond formation in a saturated fatty acid. The kinetic isotope effect (KIE) study, using Crep1 expressed in yeast, supported a conventional two-step mechanism from alkene to alkyne.39 The large primary deuterium isotope effect (k_H/k_D = 14.6 ± 3.0) for C₁₂–H bond cleavage, and the minimal isotope effect of the C₁₃–H bond breaking step (k_H/k_D = 1.25 ± 0.08), suggested that crepenynic acid is produced via an initial H-atom abstraction at C₁₂ of a linoleoyl substrate (Fig. 3).39

Intriguingly, a homolog of Crep1 has been discovered in the biosynthetic gene cluster of falcarindiol (5),41 a dietary metabolite from edible crops such as carrots, tomatoes and celery, that exhibits antifungal activity and cytotoxicity against several human cancer cell lines. The conversion of linoleic acid (4) to crepenynic acid (5) by a Crep1-like acetylenase was proposed to be involved in an early catalytic step in the falcarindiol biosynthetic pathway in tomato (Solanum lycopersicum) (Fig. 4).41 As only a few amino acid changes distinguish acetylenases from other desaturases, it is difficult to find orthologs of Crep1 by a homology search of genome databases. Instead, researchers correlated the expression levels of 40 corresponding transcripts from their RNA sequencing datasets with the falcarindiol accumulation in tomato leaves, to prioritize the putative acetylenases responsible for falcarindiol alkyne installation. Four sequentially located genes, with expression levels showing the best correlations with falcarindiol accumulation, were predicted to encode a cluster of enzymes for falcarindiol alkyne formation. Two genes (ACET1a and ACET1b) from the candidate cluster were determined to possess acetylenase activity by analyses of their overexpression in the heterologous host Nicotiana benthamiana. The results revealed that ACET1a and ACET1b encode identical acetylenases that catalyze the conversion of linoleic acid to crepenynic acid in the first step of falcarindiol biosynthesis in tomato. In addition to ACET1a and ACET1b, Soly12g100250 and Soly12g100270, which respectively encode a desaturase and a putative decarboxylase, are also involved in the biosynthesis of falcarindiol (Fig. 4).41

2.2 Polyketides

Many microbial and marine polyacetylenes are formed from polyketides or peptides. Enediyne natural products, mainly produced by soil and marine microorganisms, have an unusual bicyclo[7.3.0]-dodecadienediyne core that contains the 1,5-diyne-3-ene unit embedded in a 9-membered ring or the bicyclo[7.3.1]-tridecadiynene core in a 10-membered ring.42,43 These kinds of natural products feature special molecular architectures with promising antibiotic and antitumor
activities that are attracting considerable attention.\textsuperscript{44,45} The subunit six-carbon enediyne was investigated to determine whether it is the key structural element responsible for the biological activity. The enediyne moiety can form a reactive diradical intermediate by Bergman cyclisation. The highly reactive diradical abstracts hydrogen atoms from the DNA and generates a carbon-centered radical, leading to double stranded DNA cleavage in the presence of O\textsubscript{2}.\textsuperscript{45,46}

Early experiments feeding 9-membered neocarzinostatin\textsuperscript{47} and 10-membered esperamicin\textsuperscript{48} with \textsuperscript{13}C-labeled sodium acetate to their producers revealed that each was formed by the head-to-tail assembly of acetate units. In 2002, Shen and co-workers identified the biosynthetic gene cluster of C-1027 (6), a chromoprotein antibiotic isolated from \textit{Streptomyces globisporus}, by chromosomal walking from the \textit{sgcAB} locus, which is essential for 6 production in \textit{S. globisporus}.\textsuperscript{49,50} At the same time, Thorson and co-workers identified the biosynthetic gene locus for calicheamicin $\gamma$ (7) by screening a genomic library for cosmids that confer calicheamicin resistance to \textit{Micromonospora echinospora}.\textsuperscript{51} Both gene clusters contain the five-gene cassette (the minimal enediyne PKS gene cassette or warhead gene cassette) proposed to be responsible for enediyne core formation (Fig. 5A). Genome mining of the gene clusters producing other enediyne natural products and sequencing of some microbial genomes in the following years further confirmed the conservation of the warhead gene cassettes.\textsuperscript{52–56}

Shen’s group co-expressed the PKS (SgcE) for C-1027 with its cognate thioesterase TE (SgcE10) in \textit{E. coli}, which produced a 15-carbon polyene (9).\textsuperscript{57} Further \textit{in vivo} experiments showed that SgcE can be incorporated into the biosynthetic pathways of other 9-membered enediynes, supporting the idea that 9 is the precursor to the 9-membered enediyne cores. Meanwhile, Liang and co-workers obtained a linear carbonyl-conjugated polyene (11), from the enzymatic assay of the PKS (CalE8) of 7 with the assistance of its TE (CalE7).\textsuperscript{58,59} Using mass spectrometry, Townsend and co-workers proposed that 11 was derived...
from a β-keto carboxylic intermediate (10) (Fig. 5B).60 The enediyne core and the other polyketide moieties of the enediyne-containing compounds were considered to have different biosynthetic origins.51,52 Interestingly, by gene deletion and 18O-labeling experiments, Townsend’s group recently demonstrated that the enediyne PKS, DynE8, plays a dual role in the biosynthesis of the enediyne core and the anthraquinone scaffold in dynemicin A (8).61,62 Furthermore, Liang proposed that CalU15, encoded by a gene in the warhead gene cassette of 7 and sharing minimal sequence identity with the families of diiron containing desaturases or acetylenases, catalyzes the acetylenic bond formation in the 1,5-diyne-3-ene unit.58 However, the enzymatic activity of the protein must still be proved.

2.3 Polyketide/non-ribosomal peptide hybrids

The first gene clusters for terminal alkyne biosynthesis were identified from jamaicamide B (12) and carmabin A (13), which are terminal alkyne-bearing polyketide–nonribosomal peptide (PK–NRP) hybrids produced by the marine cyanobacteria Moorea producens JHB and M. producens 3L, respectively.20,63 A bioinformatics analysis revealed the conserved trigene cassettes, jamABC and camABC, with high amino acid sequence identities of 95.7–97.4%.64 JamA is a fatty acyl-CoA ligase, while JamB is a membrane associated fatty acid desaturase with strict substrate specificity for 5-hexanoic acid, and together with JamC, an acyl carrier protein (ACP), they are responsible for alkyne formation in a PKS–NRPS system.20 In addition, genome mining revealed more than 80 gene operons homologous to jamABC across diverse bacterial genera, indicating that this gene operon is likely to be employed in the synthesis of a variety of natural products containing an alkyne motif.19 Further experimental studies determined the biosynthetic pathway of the terminal alkyne generated by the three proteins. Hexanoic acid was first activated with ATP by JamA, and then loaded onto the ACP JamC. Subsequent dehydrogenation of the terminal alkenoic moiety was accomplished by the membrane associated fatty acid desaturase/acetylenase JamB to generate the terminal alkyne (Fig. 6B).65,66 Zhang’s group further demonstrated that desaturase/acetylenase (JamB) has a more stringent interaction with ACP than the acyl-ACP ligase (JamA) by both in vitro and in vivo assays, which provided new insights into the role of ACP in alkyne biosynthesis.67

The discovery of the first terminal alkyne biosynthetic pathway has inspired further characterizations and applications of this unique biosynthetic route. Genome mining of JamABC homologs led to the discovery of the new terminal alkyne biosynthetic cassette ttuABC, from Teredinibacter turnerae T7901.19 TtuABC shows an analogous mode of action with JamABC to form a terminal alkyne moiety. The medium-chain fatty acid is activated by acyl-CoA ligase TtuA and load onto carrier protein TtuC. One obvious difference between operons TtuABC and JamABC is that the desaturase/acetylenase TtuB is specific for decanoic acid (C10 chain length) rather than hexanoic acid (C6 chain length) for substrate selection (Fig. 6C).19

Gerwick and co-workers discovered vatiamides A–D (14–17) from Moorea producens, which feature a fatty acid tail harboring a terminal alkyne or bromoalkyne that is identical to the moiety found in jamaicamides.20,68 Genome mining of jamABC
homologs identified the PKS/NRPS gene cluster \( vatA-O \) responsible for the formation of vatiamides A and B. Efforts toward finding the genes for the biosynthesis of vatiamides C and D illuminated a 40 kb extension of NRPS and PKS modules downstream of \( vatO \) in the original 50 kb cluster. This research characterized the first example of a special combinatorial non-collinear PKS/NRPS system in which the upstream PKS cassette interacts with three separate NRPS partners (Fig. 7B). Terminal alkynes can undergo a copper-catalyzed azide–alkyne cycloaddition (CuAAC) reaction to allow sensitive alkyne identification in a crude extract.28 The researchers synthesized a fluorogenic “click”-based probe that exhibited a potent fluorescence increase at 490 nm in the triazole adduct formation. Three alkynylated lipopeptides and their brominated analogues vatiamides A–F (14–19), which are predicted to be produced by the unique non-collinear PKS/NRPS system, were discovered with the use of this fluorogenic azide-based click probe (Fig. 7A).68

2.4 Amino acids
A series of terminal alkyne-containing amino acids are reportedly produced by fungi and bacteria.69,70 Early studies revealed that some terminal alkynes originate from fatty acid or polyketide biosynthetic pathways.24,65 Recently, Chang and co-workers characterized a unique pathway to produce a terminal alkyne-containing amino acid, \( \beta \)-ethynylserine (\( \beta \)es) (20), from Streptomyces cattleya.21 Knock outs of three fatty acid desaturases (SCAT_4823, SCAT_p1525 and SCAT_0184) did not disrupt \( \beta \)es production, thus suggesting a different route than canonical desaturation for target compound production. The putative biosynthetic gene cluster (BesA-F), related to amino acid metabolism, was identified by a genomic comparison of two Streptomyces spp. that produce acetylenic amino acids (Fig. 8A) with 26 other control Streptomyces spp. genomes.21 In contrast to jamaicamide B, terminal alkyne installation in \( \beta \)es is catalyzed by three crucial enzymes, including BesB, a pyridoxal-5’-phosphate (PLP)-dependent cystathionine-\( \beta \)-lyase.
homolog, BesC, which shares homology with the non-heme iron oxidase HemeO superfamily, and BesD, a non-heme Fe/α-KG-dependent oxygenase. The pathway involves multiple unique functional groups to yield the terminal alkyne, including the BesD-mediated 4-chlorination of L-lysine (23), followed by oxidative cleavage via BesC to produce a chlorinated terminal alkene (24), with ammonia and formaldehyde as co-products. BesB catalyzes the conversion of 24 to 21 with the aid of a PLP cofactor. To probe the mechanism of BesB, 24 was used as substrate in D₂O and lead to the formation of [M + 2D]⁺ Pra which suggested at least two deprotonation–protonation events occurring during the BesB catalysis. The Cα-deprotonation giving the 4-Cl-Alg-ketimine intermediate is consistent with the cystathionine γ-synthase mechanism.
Further deprotonation at the $\beta$ position initiated the elimination of chloride to form an allene intermediate (27). The terminal allene was formed by subsequent isomerization of the allene intermediate and release from PLP (Fig. 8B and C). Pra is then ligated to glutamate to form the $\gamma$-glutamyl-$\alpha$-dipeptide (25), which is the substrate of the BesE hydroxylase to produce the $\beta$es dipeptide (26). Notably, the glutamylolation step is considered to be a physiological step that reduces side reactions and facilitates transport. The amide hydrolase is responsible for the release of free Pra (21) and $\beta$es (20).

Searches for BesA-D homologs revealed five additional Streptomyces spp. with the genes encoding the proteins, but their abilities to produce terminal alkyne containing amino acids have not been reported.\(^{21}\) Subsequently, the authors found that three of the strains could produce either Pra or $\beta$es. Their discovery will lead to the identification of enzyme homologs responsible for the formation of a wider range of terminal alkyne amino acid products. In addition, characterization of the $\beta$es biosynthetic pathway will provide an opportunity for the production of non-proteinogenic amino acids, such as halo-, terminal-alkene and terminal-alkyne amino acids, that can be used for downstream applications.\(^{21}\)

### 2.5 Meroterpenoids

Acetylenic meroterpenoids are structurally unique compounds with an alkylnyl moiety in prenyl chains. These compounds display a wide range of properties, such as cytotoxic, antifungal and antibacterial bioactivities.\(^{72,73}\) Although some chemical approaches have been developed to synthesize these acetylenic compounds, little is known about the biosynthesis and molecular basis for the acetylenic prenyl chain formation.\(^{74,75}\)

Recently, two research groups simultaneously reported that the alkynyl moieties in the prenyl chains of biscognienyne B (28) and asperpentyn (29) are generated by cytochrome P450 enzymes, as determined via heterologous expression, feeding experiments and in vitro enzyme reactions.\(^{20,22}\) Furthermore, the two groups demonstrated that the stereochemistry of the double bond in the intermediates is important for substrate recognition.

Gao and co-workers previously isolated 28 and several other acetylenic meroterpenoids from an endolithic fungus Biscogniauxia sp. (71-10-1-1).\(^{76}\) Based on the previously reported prenyltransferase lacE and the cupin protein lacF, which are reportedly associated with the biosynthesis of biscognienyne B,\(^{77}\) they identified the putative gene cluster consisting of 11 genes. Through heterologous expression in Aspergillus oryzae NSAR1, feeding experiments and in vitro assays, they demonstrated that 28 is synthesized from 4-hydroxybenzoic acid (30) by six enzymes (Fig. 9A).\(^{22}\) Particularly, they found that the alkyne in 28 was installed by the cytochrome P450 enzyme BisI, which is different from the other reported enzymes dedicated to acetylenic bond biosynthesis. In order to explore the mechanism and substrate scope of the BisI-mediated biosynthesis of the alkylnyl moiety, the authors chemically prepared the 1-OH methylated form (31) of 4-hydroxy-3-prenylbenzoic acid and performed feeding experiments (Fig. 9B). The results indicated that BisI can convert 31 to the dehydrogenated product (7E)-diene (32). However, further feeding experiments and in vitro assays revealed that BisI cannot accept 32 or its demethylated derivative 33 as a substrate to yield the corresponding acetylenic products, but it shows promiscuous activities towards C5 and C15 prenyl chains.\(^{22}\)

Meanwhile, Lin and co-workers identified the similar cytochrome P450 enzyme AtyI, responsible for the 1,3-enyne formation in

---

**Fig. 9** (A) Proposed biosynthetic pathway for biscognienyne B. (B) The substrate scope of BisI.
asperpentyn (29) (Fig. 10B). To investigate the mechanism of the enyne moiety formation by Atyl, a series of substrate derivatives were synthesized and fed to *Saccharomyces cerevisiae* expressing *atyl*. Feeding (7Z)-alkene (34) to *S. cerevisiae* expressing *atyl* led to the production of compound 35, while (7E)-alkene (37) could not be converted (Fig. 10A). Further experiments feeding 35 and its 4-OH methylated derivative 36 to *S. cerevisiae* expressing *atyl* showed no conversion to the enyne, suggesting that 7-alkyne containing intermediates were not involved in the enyne formation by Atyl. In contrast, when *S. cerevisiae* expressing *atyl* was fed 38, compound 39 was detected (Fig. 10A). The above results indicated that the Atyl-catalyzed formation of the 1,3-enyne moiety occurs through a (7Z)-diene intermediate. This P450 monoxygenase was thus verified to possess dual functions. It catalyzes prenyl chain dehydrogenation to yield a (7Z)-diene from 40, followed by further dehydrogenation to form the 1,3-enyne moiety (Fig. 10B). The formation of the (7Z)-diene intermediate might be conducive to the abstraction of the two hydrogen atoms on the same side, by the active ferryl-oxo complex of the

---

**Fig. 10** Proposed mechanisms of (A) Atyl-catalyzed synthetic substrate analogs 34 and 38 by feeding experiments, and (B) the natural Atyl-catalyzed pathway from 40 to asperpentyn (29).
P450 enzyme. Particularly, Compound I sequentially abstracts the 7-H and 8-H to give the enyne moiety. This mechanism is similar to that of a fatty acid desaturase; however, the iron cofactors are different. This hypothesis also suggests opportunities for discovering novel acetylenic meroterpenoids and expanding the enzymatic tools for the de novo biosynthesis of alkyne-tagged compounds.

3. Engineered biosynthesis of alkyne-tagged products

Tagging natural products with a unique chemical handle facilitates the stabilization, quantification, visualization, and functional analysis of natural products through bioorthogonal chemistry. In particular, the terminal alkyne can be widely employed through the Cu(I)-catalyzed azide–alkyne cycloaddition (CuAAC) ‘click’ reaction, which is one of the most powerful tools in chemical biology. In addition to tagging natural products by semi-synthesis and total synthesis, precursor-directed biosynthesis, which is a more facile and gentle way of chemical synthesis, has been used for azide- or alkyne-tag installation. However, for precursor-directed biosynthesis, a major limitation is the low conversion efficiency of precursor into product, which in turn causes the problem of the coexistence of precursors and products with the same chemical handle. Given the progress achieved toward understanding alkyne biosynthesis, de novo biosynthesis, which is employed for the creation of alkyne-labeled molecules through biosynthetic pathway engineering, has drawn keen attention. This strategy avoids the feeding of alkyne precursors and overcomes the limitation of precursor-directed biosynthesis.

The function of the trigene cassette jamABC for terminal alkyne formation in jamaicamide B was first biochemically characterized by Zhang et al. Based on this result, they generated a JamABC in situ system for introducing terminal alkyne tags into natural products. JamB has strict substrate specificity for 5-hexanoic acid and the acyl carrier protein (JamC, CamC). Therefore, only a 5-hexynoyl moiety is incorporated into alkyne-tagged natural products by the JamABC in situ system. The terminal alkyne biosynthetic cassette was first co-expressed with a type III PKS (HsPKS1) system in a starter unit, with malonyl-CoA as the extension unit, in the E. coli BAP1 strain. After induction by IPTG, the strain was grown for 2 days
in F1 minimal medium supplemented with 5-hexanoic acid. Analysis of the culture extracts revealed the production of the terminal alkyne-tagged PK compound 41 (compound 3 in ref. 65) (Fig. 11A). However, the low activity of JamB leads outcompeted 5-hexenoic acid supplied than the target terminal alkyne. Zhang and co-workers improved the activity of JamB up to 20-fold to convert hexanoic acid to the terminal alkyne in E. coli and developed a rapid and sensitive platform through fluorogenic click reaction for detecting extracellular alkene-tagged metabolites directly in crude cultures.86

Further explorations for incorporating an alkynoic extender unit into the molecular scaffolds of PKs employed the gene cassette antCDEFGM, required for the generation of the antmycin dilactone scaffold,87 and the jamABC cassette. Liu and co-workers demonstrated that an alkynoic extender unit can be accepted by the promiscuous AntE (crotonyl-CoA reductase/carboxylase (CCR)-like protein) and the acyltransferase (AT) domain embedded in AntD to synthesize an alkyne-labeled antmycin by the precursor directed biosynthesis in Streptomyces.88 To avoid feeding alkynoic precursors, co-expression of the two gene cassettes, antCDEFGM and jamABC, in E. coli led to the production of the alkyne-bearing PK/NRP hybrid antmycin analog 42 (compound 5 in ref. 65) (Fig. 11B).85

TtuABC, encoded from the genome of T. turnerae T7901, shows a similar mode of function to JamABC in forming a terminal alkyne moiety. TtuABC has different substrate specificity towards fatty acyl moieties with a C10 chain length, instead of hexanoic acid (C6 chain length). Zhang’s group employed the type III PKS (HsPKS1) in a TtuABC in situ system that specifically recognized the C10 fatty acid of the starter unit to generate the novel terminal alkyne-tagged polyketide compound 43 (compound 1 in ref. 19) in E. coli (Fig. 12A).19

Based on these results, they proved that PKS starter unit engineering is a viable way to install the alkynyl moiety from JamABC onto a polyketide to generate terminal alkyne-tagged products. They further tried to generalize this strategy to other kinds of PKSs by employing two well studied type I PKSs, LipPKS1 and DEBSM6, to produce the target alkyne-tagged polyketides.89 LipPKS1 is the first module in lipomycin biosynthesis and DEBSM6 is the last PKS module in the erthro-mycin biosynthetic pathway.90,91 Engineered LipPKS1 and DEBSM6 have been constructed to utilize malonyl-CoA instead of methylmalonyl-CoA as the extender unit, with the promiscuous DEBS thioesterase, to promote the release of a carboxylic acid product.92 Zhang and co-workers further removed the AT and ACP loading domains of the engineered LipPKS1, to create a truncated version that facilitates the alternative starter unit incorporation. The two engineered modules (LipPKS1* and DEBSM6*), in combination with JamABC working as a trans loading system, were employed to produce the alkyne-tagged polyketide 3-hydroxydec-9-ynoic acid (44) (compound 1 in ref. 89) (Fig. 12B).89 These findings have shed light on the discovery and development of synthetic tools for the de novo biosynthesis of various acetylenic products.

4. Conclusion

The alkyne is a biological and chemical functional moiety found in numerous natural products. Hitherto, a few types of gene cassettes were identified in the biosynthetic pathways of alkyne-containing natural products. The recent progress in identifying the biosynthetic enzymes and pathways of acetylenic products offers a potential route for the de novo
biosynthesis of alkyne-labeled products, for use as powerful tools for diverse downstream applications in drug discovery and chemical biology. Meanwhile, with these enzymes in hand, finding more homologs will lead to an expanded alkyne biosynthetic enzyme inventory capable of efficiently generating acetylenic products.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors would like to express sincere appreciation to an excellent group of coworkers whose contributions are cited in the text. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan (JSPS KAKENHI Grant No. JP16H06443, JP20H00490, and JP20KK013), and Japan Science and Technology Agency (JST SICORP Grant No. JPMJSC1701).

References