

Integrated omics approach to unveil antifungal bacterial polyynes as acetyl-CoA acetyltransferase inhibitors

Ching-Chih Lin^{1,2,#}, Sin Yong Hoo^{1,2,#}, Li-Ting Ma^{1,2,#}, Chih Lin¹, Kai-Fa Huang³, Ying-Ning Ho⁴, Chi-Hui Sun¹, Han-Jung Lee¹, Pi-Yu Chen¹, Lin-Jie Shu¹, Bo-Wei Wang^{1,2,5}, Wei-Chen Hsu^{1,2}, Tzu-Ping Ko³, and Yu-Liang Yang^{1,2,*}

¹ Agricultural Biotechnology Research Center, Academia Sinica, Nankang Dist., Taipei 115, Taiwan

² Biotechnology Center in Southern Taiwan, Academia Sinica, Guiren Dist., Tainan 711, Taiwan

³ Institute of Biological Chemistry, Academia Sinica, Nankang Dist., Taipei 115, Taiwan

⁴ Institute of Marine Biology and Center of Excellence for the Oceans, National Taiwan Ocean University, Jhongjheng Dist., Keelung 202, Taiwan

⁵ Department of Marine Biotechnology and Resources, National Sun Yat-sen University, Gushan Dist., Kaohsiung 804, Taiwan

#These authors contributed equally.

*Corresponding author: Dr. Yu-Liang Yang

Phone number: +886-2-2787-2089

E-mail: ylyang@gate.sinica.edu.tw

Abstract

Bacterial polyynes are highly active natural products with a broad spectrum of antimicrobial activities. However, their detailed mechanism of action remains unclear. By integrating comparative genomics, transcriptomics, functional genetics, and metabolomics analysis, we identified a unique polyynone resistance gene, *masL* (encoding acetyl-CoA acetyltransferase), in the biosynthesis gene cluster of antifungal polyynes (massilin A **1**, massilin B **2**, collimonin C **3**, and collimonin D **4**) of *Massilia* sp. YMA4. Crystallographic analysis indicated that bacterial polyynes serve as covalent inhibitors of acetyl-CoA acetyltransferase. Moreover, we confirmed that the bacterial polyynes disrupted cell membrane integrity and inhibited cell viability of *Candida albicans* by targeting ERG10, the homolog of MasL. Thus, this study demonstrated that acetyl-CoA acetyltransferase is a potential target for the development of antifungal agents.

Introduction

Polyynes or polyacetylenes contain a conformationally rigid rod-like architecture and an electron-rich consecutive acetylene moiety. Hundreds of polyynes have been discovered, out of which compounds have mainly been isolated from terrestrial plants and marine organisms^{1,2}. In contrast to polyynes from plant sources, bacterial polyynes contain a distinct terminal alkyne with conjugated systems, which causes bacterial polyynes to be more unstable. This instability has discouraged surveys of bacterial polyynes using the bioactivity-guided isolation approach. To date, only 15 bacterial polyynes have been recorded in a few species. However, these polyynes have been reported to have a broad spectrum of antimicrobial effects. For instance, cepacins, isolated from *Pseudomonas cepacia* (taxonomically reclassified as a *Burkholderia diffusa*), was reported to have antibacterial activity against the majority Gram-negative bacteria, *Staphylococci* spp., and anti-oomycetal activity against *Pythium ultimum*¹⁻⁴. In addition, Collimonins isolated from *Collimonas fungivorans* Ter331^{5,6} and Sch 31828 isolated from *Microbispora* sp. SCC1438⁷ were reported to have antifungal activity against *Aspergillus niger* and *Candida* spp., respectively. Recently, protegenins produced by *Pseudomonas protegens* were reported to have antifungal, antioomycete, and anti-algae activities⁸⁻¹¹. Despite the apparent bioactivity of these compounds, the active target and mechanism(s) remain unclear.

Here, we used a multi-omic approach and genetic engineering to identify the bioactive polyynes of *Massilia* sp. YMA4, characterized their biosynthesis gene cluster (BGC) named massilin (*mas*) BGC, and delineated the antifungal mechanism of bacterial polyynes. Furthermore, heterologous co-expression of *mas* BGC further illustrated the conserved genes critical to the electron-rich consecutive alkyne moiety and an additional gene responsible for structure derivatives modification of the polyynes. The comparative genomic analysis further uncovered the self-resistance gene and identified bacterial polyynes as antifungal agents that target the first enzyme of ergosterol biosynthesis, acetyl-CoA acetyltransferase. Furthermore, crystallographic analysis unveiled the detailed binding model of polyynes to the acetyl-CoA acetyltransferase.

Results and discussion

Transcriptomics analysis reveals polyynes as antifungal agents and their encoding BGC in *Massilia* sp. YMA4

In antagonism assay of *Massilia* sp. YMA4 against *C. albicans*, a distinct phenotype showed that the antifungal metabolites were produced in potato dextrose agar (PDA) medium but not in yeast malt agar (YMA) medium (**Supplementary Figure 1**). In addition, the antifungal metabolites were unstable in the extract and hard to purify for bioassay using the classic bioactivity-guided isolation approach. Therefore, to reveal the antifungal metabolites and their BGCs, a combined transcriptomics and metabolomics approach was used to explore the compounds produced in the two different media (PDA and YMA). First, the BGCs in the genome of *Massilia* sp. YMA4 were characterized via DeepBGC¹² with default settings and integrated with the criteria: DeepBGC score > 0.7 to characterize 13 BGCs in *Massilia* sp. YMA4 (**Supplementary Data 1**). Then, the differential expression genes (192 upregulated and 226 downregulated with *P* value < 0.05 and |fold-change| > 2) from transcriptomics analysis of PDA vs. YMA cultured cells were mapped into the 13 characterized BGCs. Among them, only one BGC was upregulated in PDA compared to YMA (**Supplementary Data 1**), and we named this gene cluster massilin (*mas*) BGC with 12 transcribed genes (*masA* to *masL*) (**Figure 1a**).

To characterize the metabolites produced from *mas* BGC, we constructed mutant strains through insertion mutation at the *masH* gene locus in *Massilia* sp. YMA4. $\Delta masH$ lost antifungal activity against *C. albicans* (**Supplementary Figure 2**), which demonstrated that *mas* BGC was crucial for the antifungal activity of *Massilia* sp. YMA4. We then conducted mass spectrometry-based untargeted metabolomics to analyze the differential features in wild-type and biosynthesis-deficient mutant strain $\Delta masH$. Via target isolation of differential features, we purified four polyynes (collimonin C/D **1**, **2**, and massilin A/B **3**, **4**) from ethyl acetate extract of *Massilia* sp. YMA4 (**Figure 1b** and **1c**). Their structures were elucidated using high-resolution mass spectrometry (**Supplementary Figure 3** and **4**) and nuclear magnetic resonance (**Supplementary Table 1-5**). Among the four polyynes, collimonin C **1** and collimonin D **2** were isolated from *C. fungivorans* Ter331⁶. In addition, a new compound with an ene-triyne moiety was named massilin A **3** and identified as a racemate at C(6)OH of an unsaturated hexadecanoic acid. Another new ene-diyne-ene compound, massilin B **4**, was presumed to be the precursor of collimonin C **1** or collimonin D **2**. Notably, massilin B **4** was more chemically stable than other polyynes with a terminal alkyne.

In antifungal activity assay (**Supplementary Figure 5**), polyynes with a terminal alkyne showed potent inhibition of *C. albicans* with minimum inhibitory concentrations (MIC): 69.73 μ M (collimonin C **1**), 35.24 μ M (collimonin D **2**), and 2.40 μ M (massilin A **3**). However, massilin B **4** with a terminal alkene moiety had no antifungal activity with MIC > 500 μ M. These results suggest that the terminal alkyne is essential for polyyne antifungal activity and stability.

Phylogenetic analysis of polyene BGCs and *mas* heterologous co-expression revealed the core components for polyene biosynthesis

We further explored 56 bacterial polyene BGCs to reveal the core biosynthesis module. The alignment results suggested the core biosynthesis gene architecture of polyene BGCs contains the fatty acyl-AMP ligase (FAAL) – 2x fatty acid desaturase (FAD) – acyl carrier protein (ACP) – fatty acid desaturase (FAD) – hydrolase/thioesterase (H/TE) – rubredoxin (Rd) (**Supplementary Figure 6**). In order to verify the in-silico result, we use plasmid-derived insertion mutagenesis to gain a serial mutant of *mas* BGC from *masD* (FAAL) to *masL* (**Supplementary Figure 2**). $\Delta masD$, $\Delta masE$, $\Delta masF$, $\Delta masH$, and $\Delta masI$ showed completely disrupted polyene biosynthesis, and $\Delta masL$ (Rd) showed significantly decreased production of polyenes. However, $\Delta masK$ and $\Delta masJ$ did not affect the biosynthesis process. To complement the mutagenesis results, we established an *Escherichia coli* heterologous co-expression system of *mas* genes to determine the genes that participate in polyene biosynthesis. The engineered *E. coli mas*- (heterologous co-expression with *masD* to *masJ*) produced a relatively nonpolar polyene (**Figure 2a**). This new polyene with an ene-triene moiety was identified and named massilin C **5** (**Supplementary Figure 7**). The results demonstrated that *masD* to *masJ* are the core genes in the biosynthesis of the ene-triene moiety. Like the core genes *masD* to *masJ*, the expression of *masB* in *Massilia* sp. YMA4 was upregulated when cultured in PDA medium. We further proposed *masB*, a dioxygenase gene, as a candidate for the hydroxylation of polyenes since similar reactions can be executed by cytochrome P450s or dioxygenases^{13–15}. The *masB* incorporated strain, *mas*+, produced collimonin C/D **1**, **2** (**Figure 2b** and **Supplementary Figure 7**), which suggested the product of co-expression of *masD* to *masJ*, massilin C **5**, is the intermediate of collimonin C/D **1**, **2** and massilin A/B **3**, **4** and would be further hydroxylated by *masB*. Furthermore, the absence of *masJ* significantly decreased the accumulation of polyenes in *mas* expressed *E. coli* (*mas*-' and *mas*+', **Figure 2**), indicating the crucial role of rubredoxin in polyene biosynthesis. Rubredoxin-hydroxylase fusion can increase alkane hydroxylation, and rubredoxin might hypothetically enhance the essential electron transfer¹⁶, which is likely critical to polyene production.

MasL serves as a polyene direct target with a protective function

“Antibiotic resistome” refers to the notion that not only antibiotic resistance genes present in resistant pathogens but also those in the antibiotic producers often harbor resistance genes for self-protection^{17,18}. Sometimes the resistance genes are located within the BGC of antibiotics¹⁹. To unveil the target protein of the bacterial polyenes of interest, we adopted the concept of the antibiotic resistome. We conducted phylogenetic analysis of bacterial polyene BGCs and classified the BGCs into two chemotaxonomy groups: a

palmitate-derived (C16) and a stearate-derived (C18) group (**Figure 3a** and **Supplementary Figure 8**). In the palmitate-derived group BGCs, *ccn* BGC branched out before the most recent common ancestor of *mas/col* BGCs. Interestingly, the gene encoding the major facilitator superfamily (MFS) transporter, which is implicated in multidrug resistance and transports small molecules and xenobiotics²⁰, is preserved in *col* BGC but lost in *mas* BGC (**Figure 3b**, **Supplementary Table 6**, and **Supplementary Data 2**). This implies that *Massilia* sp. YMA4 has an alternative protection mechanism for self-protection from polyynes. Notably, a conspicuous gene *masL*, an acetyl-CoA acetyltransferase gene, remains in *mas* BGC and ancestor *ccn* BGC but is absent in *col* BGC. We, therefore, proposed that *masL* is a potential self-resistance gene (SRG) of polyynes.

To evaluate the protective effect of *masL*, we first heterologously expressed *masL* in polyyn-sensitive *C. albicans*. The expression of *masL* rescued fungal cell viability from polyyn inhibition under minimum inhibitory concentration (**Figure 4**). Furthermore, an *in vitro* MasL inhibition assay showed that polyynes (collimonin C/D **1**, **2**, and massilin A **3**) inhibited the MasL enzyme activity (**Table 1** and **Supplementary Figure 9**). Since the drug target identification could be achieved by gene over-expression for gain-of-function of resistance²¹, we suggest *masL* serves as an SRG in the *mas* BGC, and the acetyl-CoA acetyltransferase it encodes is the target of collimonin C/D **1**, **2** and massilin A **3**. Moreover, this implicated that a deletion event occurred after horizontal gene transfer²². The *ccn* BGC divided into *mas* BGC and *col* BGC independently, each of which contained a different self-protection mechanism.

Polyynes are covalent inhibitors of acetyl-CoA acetyltransferase

To investigate how bacterial polyynes target acetyl-CoA acetyltransferase, we solved the crystal structures of MasL in its apo and two holo (bound collimonin C **1** and collimonin D **2**) forms at 1.78 Å, 1.66 Å, and 1.40 Å resolution, respectively. The asymmetric unit (space group P1 for apo MasL and MasL-collimonin D complex; P21 for MasL-collimonin C complex) of both structures contains a tetramer of the protein (**Figure 5** and **Supplementary Figure 10**), as observed in solution. The monomer of MasL shares the general fold architecture reported in type II biosynthetic thiolase family²³. MasL consists of three domains: an N-terminal α/β domain (N-domain, residues 1–121 and 251–271), a loop domain (L-domain, residues 122–250), and a C-terminal α/β domain (C-domain, residues 272–394) (**Supplementary Figure 11**). The N- and C-domains form a typical five-layered fold (α - β - α - β - α) as observed in the structures of the other type II biosynthetic thiolases including *Zoogloea ramigera* PhaA²³, *Clostridium acetobutylicum* CEA_G2880²⁴, and *Aspergillus fumigatus* ERG10A²⁵. The L-domain displays an α/β fold with a tetramerization loop associated with the C-domain.

Many high-resolution atomic structural models of the substrate-binding complex

revealed the Claisen condensation reaction and binding model within the reaction pocket of acetyl-CoA acetyltransferases. In apo and holo forms of MasL, the substrate-binding pocket was located on the surface of the enzyme facing the opposite dimer of the tetrameric assembly. The pocket was a tunnel shape of ~10 Å depth with ~6-8 Å diameter for the linear pantothenic moiety of coenzyme A (CoA) extending through the reactive center. The reactive center in MasL contained reactive cysteine residues Cys90 and nucleophilic activation residues His350 and Cys380. In the MasL-collimonin C complex, the conjugated polyynes tail extended into the MasL substrate binding site and formed a covalent bond between the terminal carbon (C16) of the alkyne moiety and the reactive cysteine sulfhydryl moiety of Cys90 via nucleophilic addition (**Supplementary Figure 12 and 13**). Thus, the terminal alkyne was essential for polyynes activity through covalent modification on the reactive cysteine sulfhydryl group of acetyl-CoA acetyltransferases.

Although we did not obtain a good X-ray diffraction result surrounding the reactive Cys90 in the MasL-collimonin D complex, the nucleophilic addition mechanism of polyynes to the acetyl-CoA acetyltransferase was confirmed by mass spectrometry analysis (**Supplementary Figure 14 and Supplementary Data 3**). The observation of collimonin C/D (+C₁₆H₁₈O₄)- derived adducts of enzyme reactive cysteine residue provided convincing evidence of protein irreversible S-alkylation via nucleophilic addition to the conjugated terminal alkyne. Furthermore, the observation of massilin A (+C₁₆H₁₈O₃)- derived adducts suggest that massilin A **3** would also covalently inhibit acetyl-CoA acetyltransferase activity. In the further comparative analysis of the polar interactions in the activation pocket of MasL in MasL-collimonin C and MasL-collimonin D complex, the C(7)OH moiety of collimonin C/D **1, 2**, His158 of MasL, and a water molecule formed a strong polar interaction network, including a direct hydrogen bond (3.00-3.16 Å) and a water-mediated hydrogen bond between C(7)OH and His158 (**Figure 5**). Moreover, although there was no significant induced-fit within the pocket, the collimonin C/D **1, 2** caused the Arg135 on the tetramerization loop to swap and form a hydrogen bond interaction and salt bridge, respectively, across the two subunits within the binding site (**Figure 5**). These polar interacting residues for the collimonin C/D **1, 2** (inhibitor) binding were similar to CoA (product) in other thiolase models²². Therefore, the acetyl-CoA acetyltransferase inhibitions by collimonin C/D are both competitive binding models against acetyl-CoA substrate.

The stereochemistry of hydroxyl groups on polyynes affects its binding affinity toward acetyl-CoA acetyltransferase

The C6 chiral center of collimonin C/D **1, 2** is critical to forming different polar interactions. The C(6)OH of collimonin C **1** had more flexibility concerning the spatial direction (with a dihedral angle to C(7)OH from 109° to 170°) and built a sophisticated polar interaction network with the amide of Pro249 in the panthetheine loop, Arg135, and multiple water molecules

(**Figure 5** and **Supplementary Figure 15**). In contrast, the C(6)OH of collimonin D **2** did not form a polar interaction network at the contrasting face and showed fixed spatial direction (dihedral angle to C(7)OH from 32.9° to 42.6°) (**Supplementary Figure 15**). The polar interaction network derived from C(6)OH of collimonin C **1** further affected the distal carboxyl group binding model. Furthermore, the reduced binding strength from the salt bridge (in MasL-collimonin D complex) to the hydrogen bond (in MasL-collimonin C complex) corresponded to their inhibition kinetic, that collimonin D **2** has a lower KI (42.84 μ M) than collimonin C **1** (297.10 μ M). The above information suggests that the stereochemistry of the hydroxyl group on polyynes is vital for initial non-covalent complex affinity. This salt bridge/hydrogen bond formation will dominate the binding affinity in dihydro-type polyynes.

***Candida albicans*. ERG10 is the antifungal target of polyynes**

As the bacterial polyynes target acetyl-CoA acetyltransferase MasL, we further evaluated the inhibition effect of bacterial polyynes to ERG10, an acetyl-CoA acetyltransferase homolog in *C. albicans*. ERG10 is the first enzyme of the ergosterol biosynthesis pathway and is crucial for fungal cell membrane formation. The enzyme activity of ERG10 was inhibited by polyynes (**Supplementary Figure 16**), and the overexpression of ERG10 in *C. albicans* also rescued the cell viability from polyynes inhibition (**Supplementary Figure 17**). Although we did not obtain the X-ray crystallography for the polyynes-ERG10 complex, the mass spectrometry analysis showed that collimonin C/D **1**, **2** and massilin A **3** target the reactive Cys90 of ERG10 (**Supplementary Figure 18**). A previous study reported that the homologous gene in *Aspergillus fumigatus*, *erg10A* is essential for survival and reduced expression leads to increased susceptibility to oxidative and cell wall-perturbing agents²⁵. In the TEM analysis, we also observed that bacterial polyynes disrupted the cell membrane integrity in *C. albicans* (**Supplementary Figure 19**).

Conclusion

The practical application of the antibiotic resistome concept with in-silico genome mining and comparative genetics analysis facilitated drug target identification. Our integrated approaches revealed that bacterial polyynes inhibit fungal growth by disrupting ergosterol biosynthesis. The crystallography analysis revealed that bacterial polyynes inhibit acetyl-CoA acetyltransferase via irreversible S-alkylation of the reactive residue with the functional terminal alkyne and the hydroxyl moiety stereochemistry affects the inhibitor affinity. Through genetic engineering, we characterized the core biosynthesis module and modification genes for polyynes. These findings will aid in the understanding of the structure-activity relationships of the polyynes and their production via microbial fermentation.

Methods

Bioinformatics analysis

The biosynthesis gene clusters (BGCs) in the genome of *Massilia* sp. YMA4 (Genbank accession number **GCA_003293715.1**) were characterized via the command-line program DeepBGC¹² with default settings and integrated with DeepBGC score > 0.7. Then, the *mas* BGC of *Massilia* sp. YMA4 was used to discover the homologous gene clusters in bacteria species using MultiGeneBlast²⁶. The database was built with a bacterial sequences database (BCT, 2020 December 01) and whole-genome sequences of polyene-reported bacterial species from NCBI (**Supplementary Data 2**). A total of 56 bacteria with polyene BGC (Cumulative Blast bit score > 1500) were found. The homologous protein sequences of each bacterial polyene BGC were respectively concatenated (total of five amino sequences, starting from *masD* homolog to *masI* homolog). The concatenated protein sequences were used for alignment (MUSCLE) and the distance (UPGMA, bootstrap 5000 times) between 56 bacteria with *Massilia* sp. YMA4 was identified for phylogenetic tree construction. The analysis was completed by using MEGA 10 with default parameters²⁷. iTOL was used to present the results of phylogenetic analysis²⁸.

The RNA sequencing and transcriptome analysis were performed by using Illumina MiSeq system (Illumina, USA) and CLC genomics workbench (version 11, CLC bio, Denmark), and the reads were mapped to the *Massilia* sp. YMA4 genome for RPKM quantification (Supplementary Data 1). The differential gene expression analysis was analyzed through CLC software with default pipeline and settings. Identification of differentially expressed genes (DEGs) with the expression ($|\text{Fold change (FC)}| \geq 2$ with P value < 0.05) was based on RPKM and analyzed using the Empirical analysis method²⁹. Two biologic replicates of each condition were analyzed.

Construction of *mas* co-expression platform in *Escherichia coli*

Individual *mas* genes were cloned to the pET-28b(+) (Merck, Germany) using restriction enzyme digestion and ligation. The artificial gene clusters were then subcloned to the destination vector and resulted in three modules for better flexibility of adjusting composition, including pET-22a(+)-*masD-masE-masF-masG*, pACYCduet-Bv4687 (*masH* homolog from *B. vietnamiensis* LMG 10929, Genbank accession number WP_011881350)-*masI-masJ*, and pRSFduet-*masB*. Two or three modules were co-transformed to *E. coli* strain C41 (Yeastern Biotech, Taiwan), and selected colonies were then inoculated in LB medium with suitable antibiotics. The expression was induced by adding isopropyl- β -D-thiogalactopyranoside (IPTG) and cultured for two days. The extractions were carried out by adding equal volumes of EA and sonicated for 30 min. Finally, the extracts were concentrated and transferred to DMSO for further analysis.

Minimum inhibitory concentration determination and genetic rescue assay

The minimum inhibitory concentration (MIC) measurement was modified from R J Lambert's method⁸. Different concentrations of collimonin C **1**, collimonin D **2**, massilin A **3**, and amphotericin B were prepared in YPD. The *C. albicans* cell viabilities were seeded with initial OD₆₀₀ 0.05 and incubated at 37°C. After 24 h incubation, the final OD₆₀₀ was recorded by Epoch 2 Microplate (BioTek Instruments, USA) for MIC calculation. The MIC of polyynes was built with cell viability (%) of different concentrations, fitting into the modified Gompertz function³⁰.

For genetic rescue assay, the *ERG10* overexpression and *masL* heterologous expression strains were seeded with OD₆₀₀ 0.05 in YPD treated with MIC of each polyne at 37°C and supplied with 40 µg/mL doxycycline for gene expression. The *C. albicans* cell viabilities were recorded at 24 h. The statistical results were analyzed using GraphPad Prism (Version 8, GraphPad Software, USA) with multiple t-test analyses (FDR < 0.05).

Enzymatic inhibition assays

The enzymatic inhibition assay of acetyl-CoA acetyltransferase was conducted by a modified method³¹. The releasing CoA was monitored using a fluorescent probe. The fluorescence intensity of each experiment was obtained by subtracting the fluorescence intensity of the polyne-free reaction. The measurement and calculation of polyynes inhibition kinetic refer to the previous covalent inhibitor model³². The progress curves used the hyperbolic regression model to calculate the kinetic inhibition parameters using GraphPad Prism (Version 8, GraphPad Software, USA).

Protein Expression and crystallization

The recombinant MasL protein was purified from *Escherichia coli* cells (see **Supplementary Methods** for details). For MasL-collimonins complex, 20 µM MasL was incubated with 100 µM collimonin C **1** or D **2** for further purification and crystallization. A freshly thawed aliquot of MasL, MasL-collimonin C complex, and MasL-collimonin D complex was concentrated to 20 mg/mL for an initial crystallization screening of ca. 500 conditions (Academia Sinica Protein Clinic, Academia Sinica). X-ray diffraction experiments were conducted by National Synchrotron Radiation Research Center (Hsinchu, Taiwan) at the TLS beamline 15A or the TPS beamline 05A. The MasL, MasL-collimonin C, and MasL-collimonin D complex structures were solved by the molecular replacement method using the structure of thiolase from *Clostridium acetobutylicum* (pdb ID **4N44**) as the search model. The molecular figures were produced with Maestro (Schrödinger Release 2021-1: Maestro, Schrödinger, USA).

References

1. Negri, R. Polyacetylenes from terrestrial plants and fungi: Recent phytochemical and biological advances. *Fitoterapia* **106**, 92-109 (2015).
2. Zhou, Z.-F., Menna, M., Cai, Y.-S. & Guo, Y.-W. Polyacetylenes of marine origin: Chemistry and bioactivity. *Chem. Rev.* **115**, 1543-1596 (2015).
3. Parker, W. L. *et al.* Cepacin A and Cepacin B, two new antibiotics produced by *Pseudomonas cepacia*. *J. Antibiot. (Tokyo)* **37**, 431-440 (1984).
4. Mullins, A. J. *et al.* Genome mining identifies cepacin as a plant-protective metabolite of the biopesticidal bacterium *Burkholderia ambifaria*. *Nat. Microbiol.* **4**, 996-1005 (2019).
5. Fritsche, K. *et al.* Biosynthetic genes and activity spectrum of antifungal polyynes from *Collimonas fungivorans* Ter331. *Environ. Microbiol.* **16**, 1334-1345 (2014).
6. Kai, K., Sogame, M., Sakurai, F., Nasu, N. & Fujita, M. Collimonins A–D, unstable polyynes with antifungal or pigmentation activities from the fungus-feeding bacterium *Collimonas fungivorans* Ter331. *Org. Lett.* **20**, 3536-3540 (2018).
7. Patel, M. *et al.* Sch 31828, a novel antibiotic from a *Microbispora* sp. taxonomy, fermentation, isolation and biological properties. *J. Antibiot. (Tokyo)* **41**, 794-797 (1988).
8. Mullins, A. J. *et al.* Discovery of the *Pseudomonas* polyne protegencin by a phylogeny-guided study of polyne biosynthetic gene cluster diversity. *mBio* **12**, e00715-21 (2021).
9. Murata, K., Suenaga, M. & Kai, K. Genome mining discovery of Protegenins A–D, bacterial polyynes involved in the antioomycete and biocontrol activities of *Pseudomonas protegens*. *ACS Chem. Biol.* doi:10.1021/acschembio.1c00276 (2021).
10. Hotter, V. *et al.* A polyne toxin produced by an antagonistic bacterium blinds and lyses a Chlamydomonad alga. *Proc. Natl. Acad. U. S. A.* **118**, e2107695118 (2021).
11. Rose, M. M. *et al.* The bacterium *Pseudomonas protegens* antagonizes the microalga *Chlamydomonas reinhardtii* using a blend of toxins. *Environ. Microbiol.* **23**, 5525-5540 (2021).
12. Hannigan, G. D. *et al.* A deep learning genome-mining strategy for biosynthetic gene cluster prediction. *Nucleic Acids Res.* **47**, e110 (2019).
13. Van Bogaert, I. N. A., Groeneboer, S., Saerens, K. & Soetaert, W. The role of cytochrome P450 monooxygenases in microbial fatty acid metabolism. *FEBS J.* **278**, 206-221 (2011).
14. Ross, C., Scherlach, K., Kloss, F. & Hertweck, C. The molecular basis of conjugated polyne biosynthesis in phytopathogenic bacteria. *Angew. Chem. Int. Ed. Engl.* **53**, 7794-7798 (2014)

15. Martínez, E. & Campos-Gómez, J. Oxylipins produced by *Pseudomonas aeruginosa* promote biofilm formation and virulence. *Nat. Commun.* **7**, 13823 (2016).
16. Nie, Y., Liang, J., Fang, H., Tang, Y.-Q. & Wu, X.-L., Two novel alkane hydroxylase-rubredoxin fusion genes isolated from a *Dietzia* bacterium and the functions of fused rubredoxin domains in long-chain n-alkane degradation. *Appl. Environ. Microbiol.* **77**, 7279-7288 (2011).
17. Chopra, I. Over-expression of target genes as a mechanism of antibiotic resistance in bacteria. *J. Antimicrob. Chemother.* **41**, 584-588 (1998).
18. Hobson, C., Chan, A. N. & Wright, G. D., The antibiotic resistome: A guide for the discovery of natural products as antimicrobial agents. *Chem. Rev.* **121**, 3464-3494 (2021).
19. Yan, Y., Liu, N. & Tang, Y., Recent developments in self-resistance gene directed natural product discovery. *Nat. Prod. Rep.* **37**, 879-892 (2020).
20. Redhu, A. K., Shah, A. H. & Prasad, R. MFS transporters of *Candida* species and their role in clinical drug resistance. *FEMS Yeast Res.* **16**, fow043 (2016).
21. Sugden, C. J., Roper, J. R. & Williams, J. G., Engineered gene over-expression as a method of drug target identification. *Biochem. Biophys. Res. Commun.* **334**, 555-560 (2005).
22. Soucy, S. M., Huang, J. & Gogarten, J. P. Horizontal gene transfer: Building the web of life. *Nat. Rev. Genet.* **16**, 472-482 (2015).
23. Modis, Y. & Wierenga, R. K. A biosynthetic thiolase in complex with a reaction intermediate: the crystal structure provides new insights into the catalytic mechanism. *Structure* **7**, 1279-1290 (1999).
24. Kim, S. *et al.* Redox-switch regulatory mechanism of thiolase from *Clostridium acetobutylicum*. *Nat. Commun.* **6**, 8410 (2015).
25. Zhang, Y. *et al.* *Aspergillus fumigatus* mitochondrial acetyl coenzyme A acetyltransferase as an antifungal target. *Appl. Environ. Microbiol.* **86**, e02986-19 (2020).
26. Medema, M. H., Takano, E. & Breitling, R., Detecting sequence homology at the gene cluster level with MultiGene Blast. *Mol. Biol. Evol.* **30**, 1218-1223 (2013).
27. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547-1549 (2018).
28. Letunic, I. & Bork, P., Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **47**, W256-W259 (2019).
29. Robinson, M. D. & Smyth, G. K. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* **9**, 321-332 (2008).

30. Lambert, R. J. & Pearson, J., Susceptibility testing: accurate and reproducible minimum inhibitory concentration (MIC) and non-inhibitory concentration (NIC) values. *J. Appl. Microbiol.* **88**, 784-790 (2000).
31. Long, T., Sun Y., Hassan, A., Qi, X. & Li X. Structure of nevanimibe-bound tetrameric human ACAT1. *Nature* **581**, 339-343 (2020).
32. Strelow, J. M. A perspective on the kinetics of covalent and irreversible inhibition. *SLAS Discov.* **22**, 3-20 (2017).

Acknowledgements

This research was funded by the Ministry of Science and Technology of Taiwan (MOST 104-2320-B-001-019-MY2). We thank Prof. Chih-Chuang Liaw (National Sun Yat-sen University, Taiwan) and the R/V Ocean Researcher III team for collecting marine sediment. We thank Dr. Chao-Jen Shih (Bioresource Collection and Research Center, Taiwan) for isolating and identifying *Massilia* sp. YMA4. The materials and methods for constructing biosynthetic gene-null mutant strains were generously provided by Prof. Nai-Chun Lin's Lab (National Taiwan University, Taiwan). We thank Prof. Ching-Hsuan Lin and Ms. Chih-Chieh Hsu (National Taiwan University, Taiwan) for providing the material for constructing a tetracycline-inducible expression system. We thank Mr. Ning Lu, Ms. Ying-Mi Lai, and Ms. Chia-Chi Peng for collecting preliminary data. NMR data were collected in the High Field Nuclear Magnetic Resonance Center, Academia Sinica. LC-MS data were collected in the Metabolomics Core Facility, Agricultural Biotechnology Research Center, Academia Sinica, and the Proteomics Core Laboratory, Institute of Plant and Microbial Biology, Academia Sinica. TEM data were collected in the Biological Electron Microscopy Core Facility, Academia Sinica. The EM core facility is funded by the Academia Sinica Core Facility and Innovative Instrument Project (AS-CFII-108-119). We further thank the Protein Crystallization Facility (Academia Sinica Protein Clinic, Academia Sinica) for crystallization preparation; and the National Synchrotron Radiation Research Center (Hsinchu, Taiwan) with beamlines TLS 15A and TPS 05A for assistance in X-ray data collection and access to the synchrotron radiation centers.

Author Contributions

C. -C. L., S. Y. H., L. -T. M., C. L., C. -H. S., H.-J. L., P.-Y. C., L.-J. S., B.-W. W., and W. -C. H. performed the experiments. C. -C. L., S. Y. H., K. -F. H., T. -P. K., and Y. -N. H. carried out the data analysis. C. -C. L., S. Y. H., and L. -T. M. wrote the manuscript. Y. -L. Y. supervised the study and acquired funding to support the work.

Competing interests

There are no conflicts to declare.

Data availability

The transcriptome datasets were deposited in the Sequence Read Archive (SRA) database under the accession number **PRJNA706894**. The refined models of MasL, MasL-collimonin C, and MasL-collimonin D complex were deposited in the Protein Data Bank with pdb codes **7EI3**, **7EI4**, and **7FEA**, respectively. Detailed experimental methods were described in supplementary information. The results of transcriptome analysis, Multigeneblast and bolltom up proteome were listed in **Supplementary Data 1, 2 and 3**.

Figures and Tables

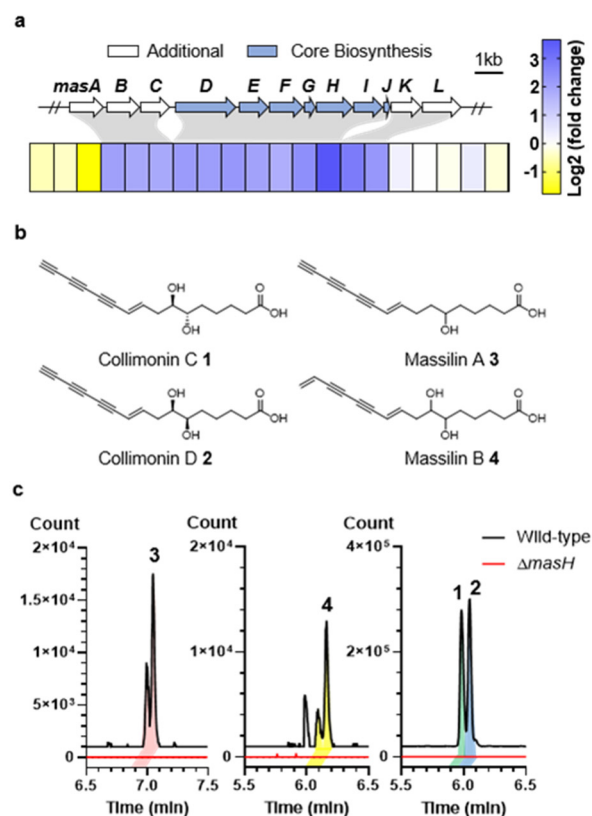


Figure 1. Polyynes and polyynone biosynthesis gene cluster of *Massilia* sp. YMA4. (a) Gene expression profile around the *mas* BGC under polyynone production (PDA) versus non-production (YMA) medium. The expressions of additional and core biosynthesis genes of *mas* BGC are framed in the grey area. (b) Structure of polyynes collimonin C/D **1**, **2** and massilin A/B **3**, **4**. (c) Extraction ion chromatography (EIC) of collimonin C/D **1**, **2** (*m/z* 273.1132), massilin A **3** (*m/z* 257.1183), and massilin B **4** (*m/z* 275.1289) in *Massilia* sp. YMA4 wild type (black) and the biosynthesis null mutant strains ($\Delta masH$) with a 10 ppm mass window.

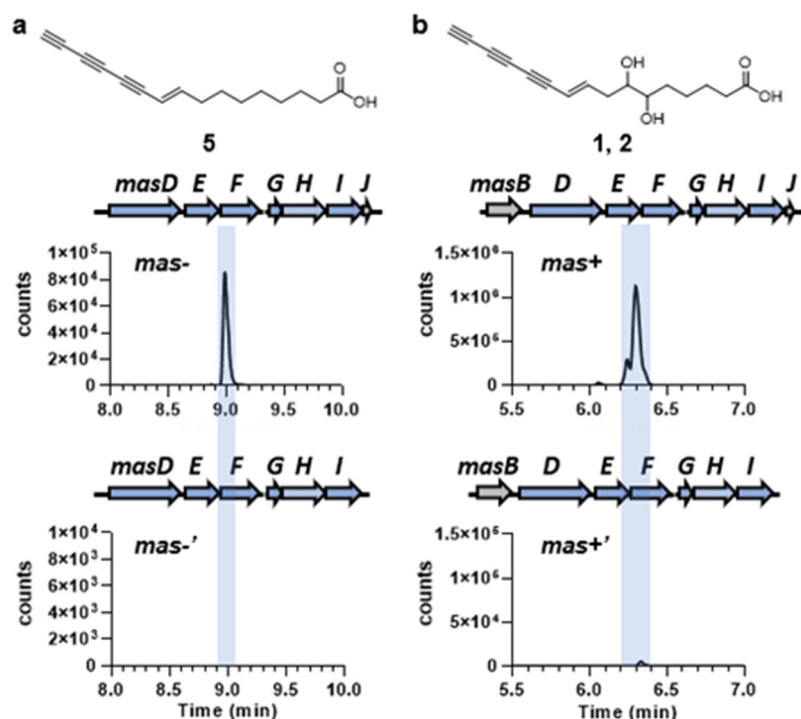


Figure 2. Heterologous co-expression of *mas* genes in *E. coli*. (a) The EIC (m/z 241.1234) of engineered *E. coli* strain, which carried *masD*, *masE*, *masF*, *masG*, *Bv4687* (homolog of *masH* from *Burkholderia vietnamiensis* LMG 10929) and *masI* when co-expressed with (*mas*-, top) or without (*mas*-' , bottom) *masJ*. (b) The EIC (m/z 273.1132) of *E. coli* strain *mas*- and *mas*-' harboring additional dioxygenase (*masB*) were renamed as *mas*+ and *mas*+'. The structures of products are illustrated on top of each chromatogram and were verified by LC-HRMS/MS analysis (Supplementary Figure 7).

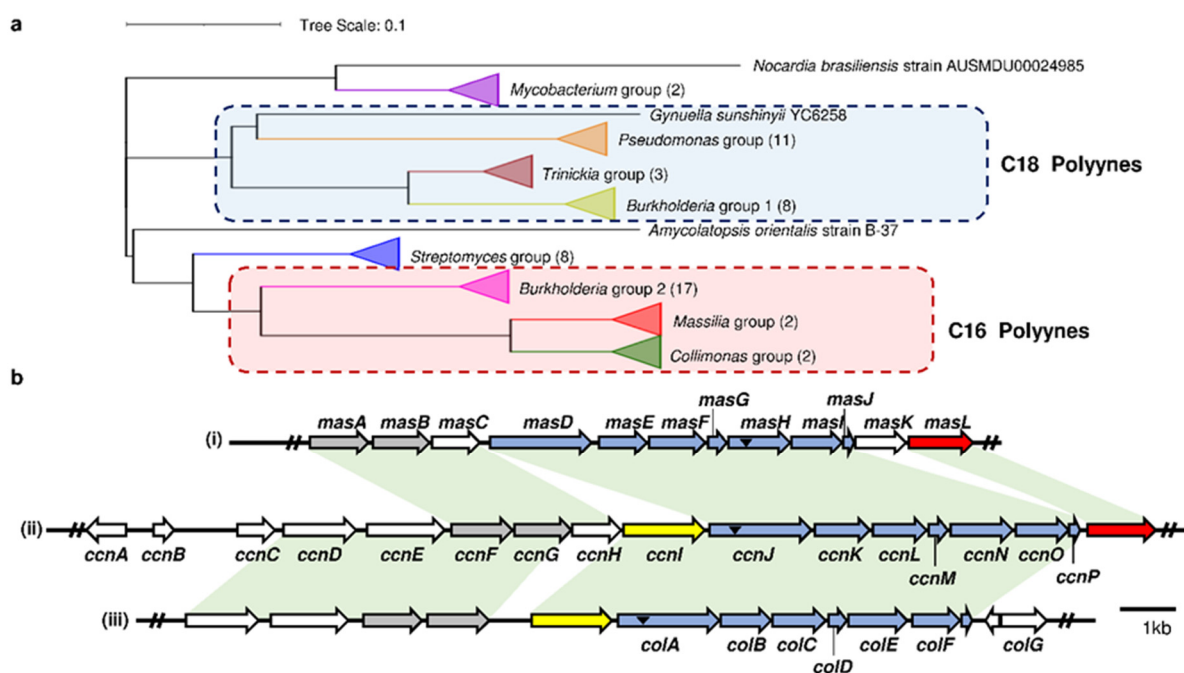


Figure 3. Comparative analysis of polyne biosynthetic gene cluster and the structures of the bacterial polyynes. (a) Phylogenetical analysis of polyne biosynthetic gene clusters (BGCs) in bacteria (**Supplementary Figure 6 and 8**). Species in the red box have been reported to produce palmitate-derived polyynes (C16), and species in the blue box have been reported to produce stearate-derived polyynes (C18). The number of BGCs in each group is shown in parentheses. (b) Comparison of the polyne BGC architectures of *mas* BGC in *Massilia* sp. YMA4 (i), *ccn* BGC in *B. ambifaria* BCC0191(ii)⁴, and *col* BGC in *C. fungivorans* Ter331(iii)⁶. Genes conserved in polyne BGC across the phylogenetic tree are shown in blue, and those conserved in the C16 polyne group are shown in gray. The potential protective genes in the BGC are shown in red for acetyl-CoA acetyltransferase and in yellow for MFS transporter. The corresponding homologs (over 40% identity) in BGCs between two species is shown as the light green area. Black triangles indicate the mutation sites in previous research and this study.

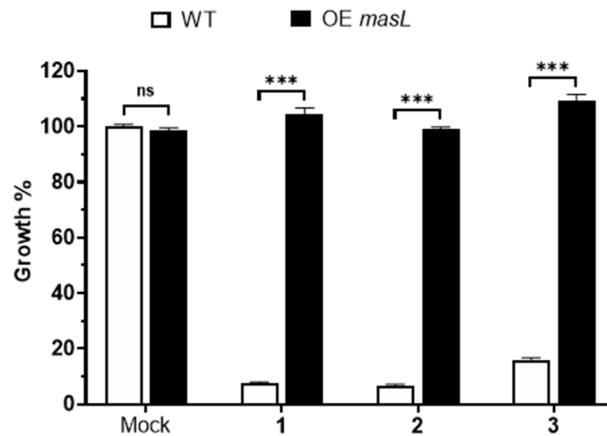


Figure 4. *C. albicans* was rescued by heterologous overexpression of *masL* (OE *masL*) from the minimum inhibitory concentration of collimonin C/D 1, 2 and massilin A 3. The cell viabilities were normalized to the mock treatment. The standard deviation was calculated based on three replicates and the two-tailed Student's t-test for statistical analysis. ***, *P* value < 0.000001. Each experiment contained three biological replicates (*N*=3).

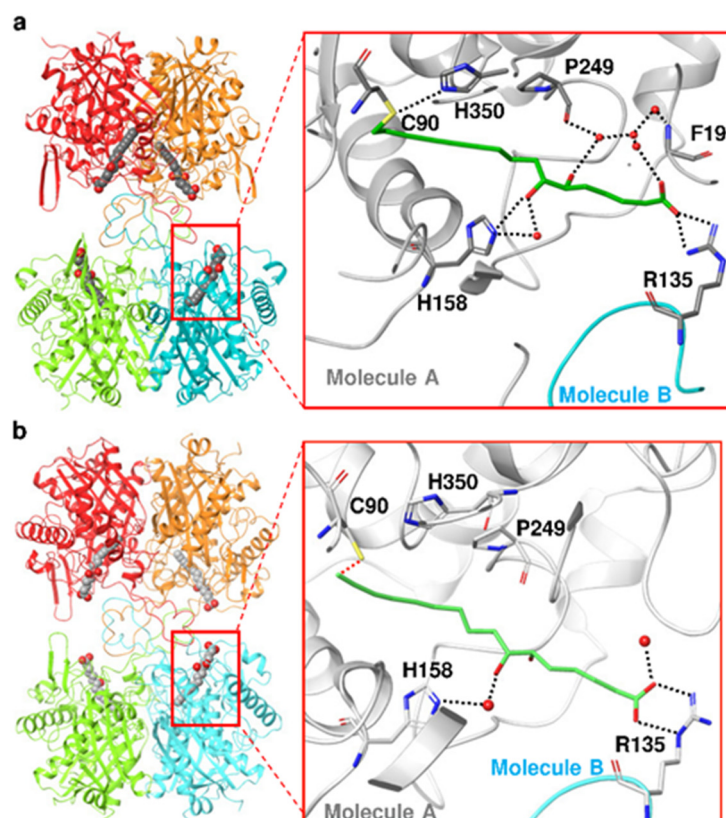


Figure 5. X-ray structures and polar interaction within the binding site of MasL-collimonin C (a) and MasL-collimonin D complex (b). Overall views of the crystal structures of MasL-collimonin C complex. Four subunits per asymmetric cell are shown in red, orange, light green, and cyan, respectively. Collimonin C/D 1, 2 is shown as green space-filling balls. In the magnified view, the residues involved in collimonin C/D 1, 2 interactions are shown as sticks with sequence identities indicated in the main chain molecule shown in gray. The R135 residue of another molecule is shown in cyan. The dotted lines indicate the hydrogen bonds and salt bridges involved in collimonin C/D 1, 2 interactions within the binding pocket.

Table 1. Inhibition kinetic of *Massilia* sp. YMA4 MasL by polyynes.

Polyynes	MasL		
	K_i (μM)	k_{inact} (min^{-1})	k_{inact} / K_i ($\mu\text{M}^{-1} \text{min}^{-1}$)
Collimonin C 1	297.10	9.798×10^{-2}	3.30×10^{-4}
Collimonin D 2	42.84	5.208×10^{-2}	1.22×10^{-3}
Massilin A 3	132.10	3.449×10^{-2}	2.61×10^{-4}

*Each experiment contained three biological replicates ($N=3$).

†Full data was illustrated in **Supplementary Figure 9**.