Title

Synthetic Biology in Natural Product Biosynthesis

Journal

Chemical Reviews, 125(7)

ISSN

0009-2665

Authors

Seshadri, Kaushik
Abad, Abner ND
Nagasawa, Kyle K
et al.

Publication Date

2025-04-09

DOI

10.1021/acs.chemrev.4c00567

Synthetic Biology in Natural Product Biosynthesis

Kaushik Seshadri,[1,#] Abner N. D. Abad,[1,#] Kyle K. Nagasawa,[2,#] Karl M. Yost,[2] Colin W. Johnson,[2] Moriel J. Dror,[3] Yi Tang[1,2,3,*]

1. Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, 420 Westwood Plaza, Los Angeles, CA 90095, USA.
2. Department of Chemistry and Biochemistry, University of California, Los Angeles, 420 Westwood Plaza, Los Angeles, CA 90095, USA.
3. Department of Bioengineering, University of California, Los Angeles, 420 Westwood Plaza, Los Angeles, CA 90095, USA.

# These authors contributed equally.
* Correspondence: yitang@ucla.edu

Abstract

Synthetic biology has played an important role in the renaissance of natural products research during the post-genomics era. The development and integration of new tools have transformed the workflow of natural product discovery and engineering, generating multidisciplinary interest in the field. In this review, we summarize recent developments in natural product biosynthesis from three different aspects. First, advances in bioinformatics, experimental, and analytical tools to identify natural products associated with predicted biosynthetic gene clusters (BGCs) will be covered. This will be followed by an extensive review on the heterologous expression of natural products modeled in bacterial, fungal and plant organisms. The native host-independent paradigm to natural product identification, pathway characterization, and enzyme discovery is where synthetic biology has played the most prominent role. Lastly, strategies to engineer biosynthetic pathways for structural diversification and complexity generation will be discussed, including recent advances in assembly-line megasynthase engineering, precursor-directed structural modification, and combinatorial biosynthesis.

1. Introduction

1.1 Discovery and Biosynthesis of Natural Products

While natural products are often broadly defined as any molecule found in nature, in this review we will follow the definition more traditionally used within the organic and medicinal chemistry communities, which defines natural products as small organic molecules (molecular weight < 1,500 Daltons) produced by dedicated, conditional metabolic pathways.[1] In contrast to the pathways of primary metabolism, conditional (or secondary) metabolic pathways are not conserved across all organisms and are generally activated in response to specific environmental stimuli, enhancing the producing organism's survivability.[2] Despite their canonical functions often remaining unclear, natural products and their diverse biological activities have a rich history of being repurposed for the betterment of humankind, especially in the contexts of human medicine and agriculture. Between 1981 and 2019, natural products, natural product-derived, and natural product-inspired molecules accounted for nearly 65% of newly approved drugs.[3] With their broad spectrum of active pharmacophores and bioactivities, natural products and their derivatives exhibit a wide array of activities and have been used in the treatment or prevention of many debilitating human ailments—Alzheimer's disease, asthma, various infectious diseases, cancer, and hypercholesterolemia, etc. Natural products provide valuable leads for drug development through their richly decorated structures and evolutionarily optimized properties that are unmatched by synthetic compounds.[4] Despite the multitude of success stories during the golden age of natural product-driven drug discovery, significant challenges emerged in the early 1990s that led to a decline in interest by the pharmaceutical industry.[5] In particular, the issue of "rediscovering" already known natural products has been a major roadblock, and new methods to rapidly "dereplicate" metabolites identified from the classic isolation workflow are needed.

There has been revitalized interest in natural products as drug leads in recent years, due to shifting emphasis on finding molecular disruptors of protein-protein interactions, targeted delivery of payloads, and enriching chemical libraries for high-content and phenotypic screening.[6] For example, antibody-drug conjugates (ADCs) have emerged as a clinically relevant family of cancer therapeutics using highly toxic natural products as payloads.[7] By conjugating the natural product to a monoclonal antibody that targets desired tumor cells, off-target cytotoxicity can be minimized. A number of ADCs have already been approved for clinical use, with payloads such as the enediyne polyketide calicheamicin, the semisynthetic monomethyl auristatin E, the macrolide maytansine, etc.[8]

Natural products such as cyclosporine and FK506 led to the first discovery of molecular glues,[9] which have continued to be developed in recent years, exemplified by the functional characterization of the polyketide manumycin as a molecular glue between TP53 and E3 ligase UBR7.[10] Along these lines, natural products have been widely explored as components of targeted protein degradation, such as parts of proteolysis-targeting chimeras (PROTACs).[11] The fungal sesquiterpene ovalicin that inhibits methionine aminopeptidase 2 (MetAP2) was the first protein-targeting compound used in the proof of concept of PROTAC.[12] Since then, many natural products with identified molecular targets have been repurposed in various PROTAC molecules.[11] For example, E3 ligase-targeting natural products, such as the diterpene nimbolide from the Neem tree, was recently used by Nomura et al. as the recruiter component in targeted protein degradation.[13] Collectively, these examples illustrate the important roles natural products can play in therapeutic development—as well as the need for continual discovery of new compounds from Nature.

A major reason for the ongoing renaissance in natural products research is the injection of genomics- and transcriptomics-driven discovery, along with application of state-of-the-art synthetic biology tools.[14,15] With rapid development in both *in silico* and experimental technologies, the discovery and engineering of natural products have turned the entire field on its head in recent years. Classically, natural product isolation and characterization were dominated by natural product chemists, who analyzed extracts from microbes and plants for the presence of new biological activities, followed by enrichment and identification of the metabolites. A working knowledge of botany, bacteriology, and mycology may have been helpful, but details of biochemistry and molecular biology were certainly not required. At the onset of gene-driven investigation of natural product biosynthesis in the 1990s, the inaccessibility of producing organisms, along with high barriers of obtaining fully sequenced biosynthetic gene clusters (BGCs), hindered the entry of researchers from related disciplines into this field.

In the post-genomics era, advances in bioinformatics—identification of BGCs from genomes, prediction of protein structures and functions, computational structural dereplication, as well as availability of publicly accessible databases—have transformed the field into one that is informationally rich, and for the most part, producing-organism independent (as long as genome information is available).[16] At the same time, breathtaking paces of advances in genome sequencing, gene synthesis and assembly, gene editing, and protein engineering have lowered the entry barrier for many to initiate research projects in natural product biosynthesis, and to launch their independent academic careers and entrepreneurial efforts.

Collectively, these "synthetic biology" tools have brought forth multidisciplinary efforts to realize the full potential of Nature's chemical inventory for drug discovery, biomanufacturing, and green synthesis. It is important to emphasize, however, that natural product chemists with expertise in small molecule isolation from complex extracts and in structural determination with sub-milligram quantities of samples remain essential at the backend of the workflow to translate the genetic information into chemical structures. Nevertheless, their skills have also been complemented by the continuing improvement in nuclear magnetic resonance spectroscopy (NMR) and mass spectrometry (MS) instrumentations, as well as computational analytical tools for metabolite detection and structural elucidation.

1.2 Major Classes of Natural Products

Among the more than 500,000 natural products that have been isolated, physiochemical properties, structural features, and biosynthetic origins have been the major criteria by which the compounds are classified. For the purpose of this review, classifications from the recent texts by Walsh and Tang will be used.[1] Figures 1, 3, and 4 list notable examples of natural products from each family, with many more shown throughout the review.

Polyketides are broadly defined as natural products that are assembled through the iterative head-to-tail condensation of acyl units, such as acetate and propionate, and are not always completely reduced at the β-positions after each condensation.[17] Polyketide carbon skeletons are assembled by polyketide synthases (PKS).[18] Fatty acids, which are synthesized by fatty acid synthases (FASs), follow the same chemical logic except the carbon chains are completely reduced.[19] Polyketides can be further subclassified based on structure, including macrolides such as erythromycin (**1**), decalin-containing molecules such as lovastatin (**2**), aromatics, polyethers, polyenes, and hybrids with peptide, terpene, or phenylpropanoid building blocks. The structural diversity of polyketides arises from the incorporation of different acyl building blocks in the form of α-carboxyacyl-CoAs, different sizes as a result of varying extents of polymerization, combinations and permutations in modifications of the α- and β- positions following each condensation, cyclization modes of the polyketide backbones, and post-PKS modifications.

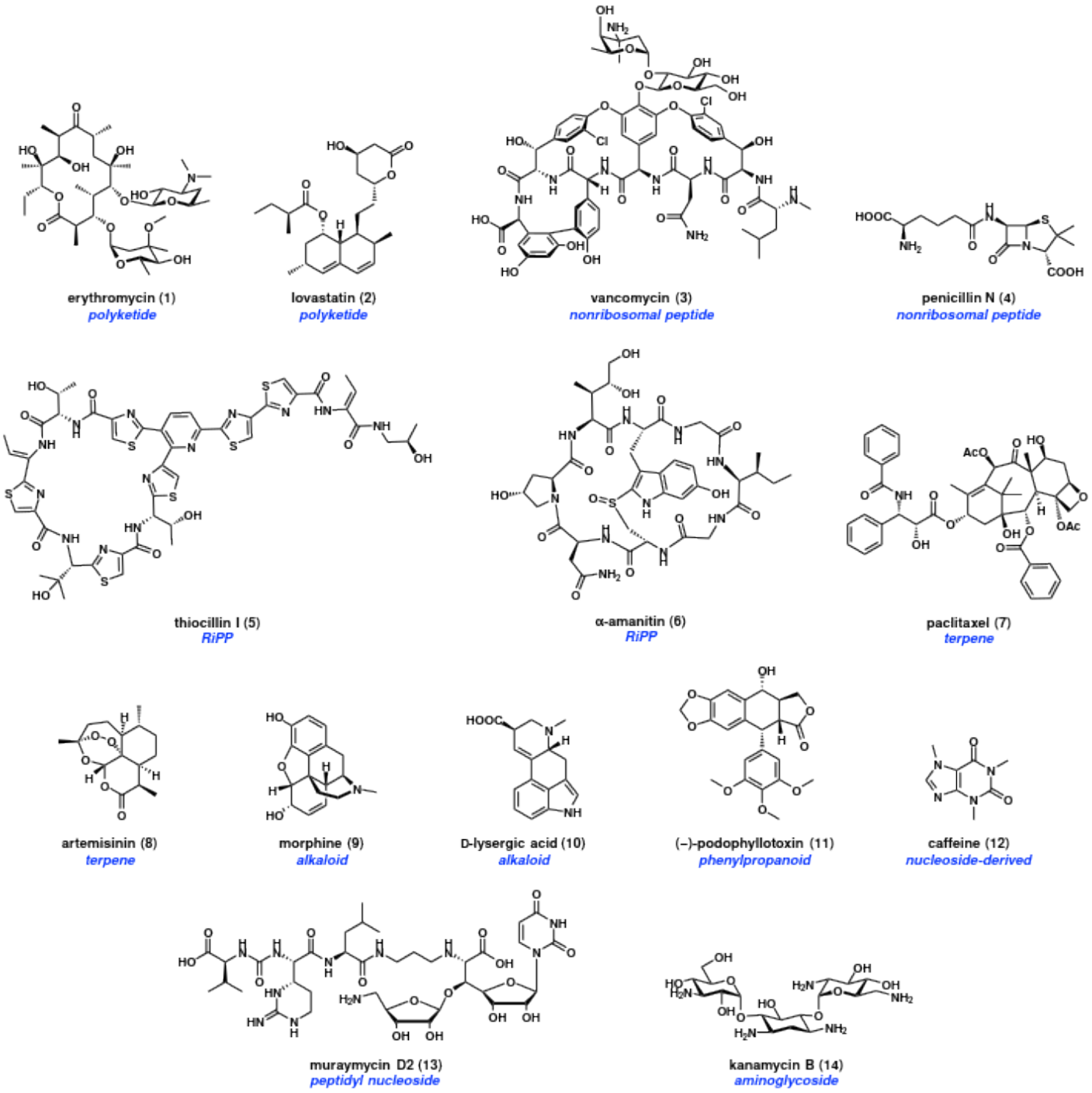


Figure 1. Representative natural products from eight different classes.

Nonribosomal peptides (NRPs) are natural products that contain amino acid building blocks that are polymerized via amide bonds in a ribosome-independent fashion. In addition to the twenty proteinogenic amino acids, unnatural amino acids, hydroxy acids, or fatty acids (often as starter units) can be incorporated into NRPs, giving rise to additional structural diversity. The predominant mode of assembly of NRPs is through the modular megasynthetases known as nonribosomal

peptide synthetases (NRPSs), in which each module is responsible for the activation of an amino acid via adenylation, transfer to the carrier protein as a thioester, and condensation with the aminoacyl thioester from the previous module to elongate the growing intermediate.[20] Notable examples include vancomycin (**3**) and penicillin N (**4**), representing cyclic and linear NRPs, respectively. Both **3** and **4** highlight characteristic features of NRPs, including use of unnatural amino acids, epimerizations at the α-carbon, crosslinking of side chains, oxidative rearrangements, and glycosylations. Other modes of assembling peptides or pseudopeptides are also found in Nature, utilizing discrete amide bond forming ligases, such as ATP-grasp enzymes or amide-bond forming enzymes.[21,22]

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a fast-growing family of peptide natural products that are biosynthesized as precursor peptides at the ribosome, followed by extensive post-translational modification (PTM) steps (Figure 2).[23] The precursor peptide is genetically encoded in a separate open reading frame (ORF) and contains the core peptide that is transformed into the mature RiPP natural product, as well as leader and follower peptide sequences that recruit enzymes for PTMs. The types of PTMs are diverse, and include oxidative cyclization of serine, threonine, and cysteine residues into oxazole, methyloxazole, and thiazole rings, respectively; dehydration of serine and threonine into dehydroalanine and dehydrobutyrine, respectively; macrocyclization via aza-Diels-Alder cyclization, isopeptide formation, radical-SAM mediated side chain crosslinking, etc. Examples of RiPPs include thiocillin I (**5**) which is the product of many of the aforementioned PTMs, and the death-cap mushroom toxin α-amanitin (**6**). Because RiPP precursor peptides are genetically encoded, this family has been the subject of intensive bioinformatics mining and synthetic biology engineering efforts in recent years.[24]

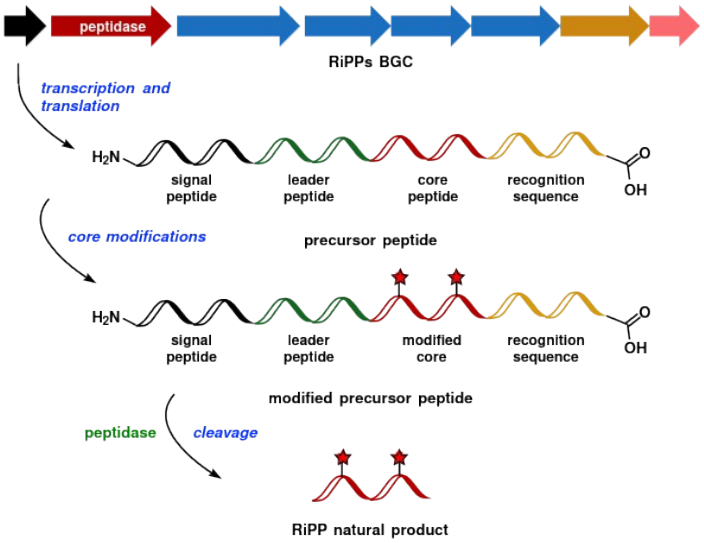


Figure 2. Simplified scheme of RiPP biosynthesis.

Terpenes constitute the largest class of natural products with more than 75,000 known members.[25] Compounds in this family are polymerized from five-carbon isopentenyl pyrophosphate building blocks known as Δ^2-isoprenyl pyrophosphate (also referred to as dimethylallyl pyrophosphate, or DMAPP) and Δ^3-isoprenyl pyrophosphate (IPP). Polymerization and subsequent cyclization steps are driven by formation and quenching of carbocation intermediates, which are distinct from the mechanisms observed in the biosynthesis of PKs and NRPs. The C10, C15, C20, and C25 terpenes are known as monoterpenes, sesquiterpenes, diterpenes and sesterterpenes, respectively, while the head-to-head dimerization of C15 farnesyl pyrophosphate (FPP) or C20 geranylgeranyl pyrophosphate (GGPP) give rise to the C30 squalene and C40 phytoene, respectively. Cyclization of the linear hydrocarbon terpene, followed by oxidative modification, catalyzed mostly by cytochrome P450s, generates tremendous structural diversity in this family of compounds that are found in all kingdoms of life. Many terpenes and terpenoids produced by plants are essential drugs to humans, including the antitumor paclitaxel (Taxol®, **7**) and antimalarial artemisinin (Qing Hao Su, **8**). Reconstitution of plant terpene natural products has seen major advancements in the recent decade following the landmark work on the production of artemisinic acid, the precursor of **8**, in yeast.[26]

Alkaloid biosynthesis has also been a major focus of synthetic biology, aimed at providing more scalable access to many plant-derived pharmaceuticals and psychoactive compounds.[27] Alkaloids, the oldest class of natural products, are named for the presence of basic nitrogen(s) in the molecules. Notable alkaloids include the analgesic morphine (**9**) from opium poppies (*Papaver somniferum*), and D-lysergic acid (**10**), the precursor to fungal ergot alkaloids and the semisynthetic lysergic acid diethylamide (LSD). Alkaloid biosynthesis is commonly initiated with the decarboxylation of amino acids such as lysine, ornithine, tyrosine, tryptophan, etc., a biosynthetic feature that has led to a different classification of alkaloids.[28] Alkaloids also feature various nitrogen-containing heterocycles, including pyrrolidines, indoles, tropanes, pyrrolizidines, etc. With advances in heterologous expression in model yeasts and plants, as well as gene discovery from plant transcriptomics, many alkaloid biosynthetic pathways have been elucidated in the last decade.

In addition to these five major classes of natural products, other notable classes include phenylpropanoids such as (-)-podophyllotoxin (**11**), which are derived from deaminated phenylalanine precursors such as cinnamic acid, *p*-coumaric acid, and the subsequent monolignols such as sinapyl, coniferyl, or coumaryl alcohol. Caffeine (**12**) and muraymycin D2 (**13**)

are examples of natural products that are derived from nucleosides. While **12** is an example of a nucleobase natural product, the bacterial cell wall translocase inhibitor **13** is a member of the large peptidyl nucleosides in which a pseudopeptide is grafted to the C_5' of the ribose ring of uridine.[29] Lastly, glycosylated natural products, in which sugars are connected to the core scaffolds via glycoside linkages, are abundant in Nature, such as **1** and **3**. The sugar moieties are essential for the biological activities of the natural products, serving as recognition elements in binding to the molecular targets. Kanamycin B (**14**) is an example of an aminoglycoside, a category comprising of many important antibiotics.[30]

Nature is able to generate further structural diversity and biological activities by recruiting building blocks from two or more different families to afford hybrid compounds. The glycosylated polyketides **1** and **3** can be considered examples of such merging of biosynthetic pathways. Additional examples are shown in Figure 3. The antitumor epothilone D (**15**) is assembled from a hybrid thio-templated assembly that consists of one NRPS module and the remaining PKS modules.[31] Such combinations of polyketide and peptide in the same compound are abundant in bacteria and fungi. Paxilline (**16**) is an example of a terpene-alkaloid hybrid, in which the electron-rich indole ring is prenylated with GGPP to form the indole diterpene precursor that can be oxidatively cyclized to give the hexacyclic structure.[31] Quercetin (**17**) is an example of a plant flavonoid that is derived from type III PKSs that use phenylpropanoid precursors as starter units. This is a large family of natural products of considerable pharmaceutical and nutraceutical value, resulting in the extensive exploration of their biosynthesis in microbial hosts.[32] Hyperforin (**18**) is a nutraceutical component of St. John's Wort and is a hybrid of polyketide and terpene. The electron-rich polyketide product, a benzenetriol, is tetra-alkylated by prenyl precursors (3x DMAPP, 1x GPP) to arrive at the complex final structure.[33]

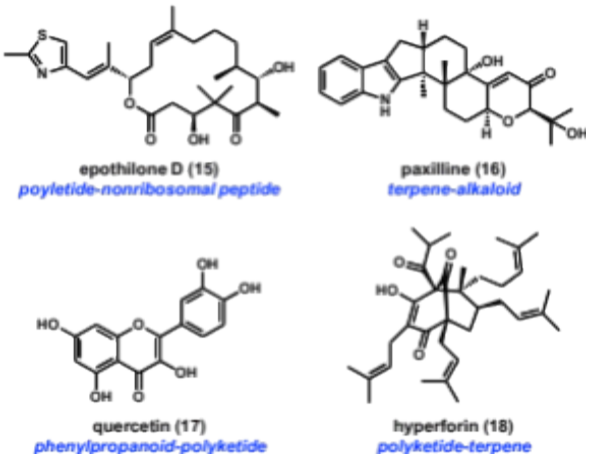


Figure 3. Representative natural products that contain features from multiple classes (hybrids).

Lastly, many natural products are not biosynthesized using signature enzymes from these pathways discussed above. Instead, other enzymes are involved in building the molecular scaffolds from primary metabolites. Notable examples include altemicidin (**19**), an antibiotic whose core is biosynthesized from β-NAD^+ and *S*-adenosylmethionine (SAM) using a pyridoxal-5'-phosphate (PLP)-dependent enzyme;[34] and fluopsin C (**20**), a copper-binding molecule (chalkophore) found in *Pseudomonas* bacteria that is derived from L-cysteine (Figure 4).[35] Natural products that are synthesized from these pathways without an easily identifiable core enzyme (such as PKS, NRPS, etc.) are of significant interest for genome mining and exploration of enzymology.[36]

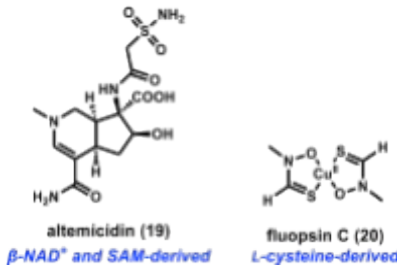


Figure 4. Representative natural products that do not belong to the eight classes.

1.3 Biosynthetic Gene Clusters (BGCs)

In fungal and bacterial organisms, the genes responsible for the biosynthesis of natural products are commonly encoded in biosynthetic gene clusters (BGCs), where the genes responsible for the formation of the natural product are colocalized within the genome (Figure 5). The first BGC identified was the cluster responsible for the production of actinorhodin, identified forty years ago from *Streptomyces* by Sir David Hopwood.[37] In the publication, Sir Hopwood anticipated that: "*the tendency for the genes of antibiotic synthesis to be clustered together… have general significance for this area of biotechnology.*" The concept of BGCs was further demonstrated by publications on the erythromycin BGC in the early 1990s.[38,39] The clustering of biosynthetic genes is not a common feature in plants, although examples of clustering of related genes in plant genomes have emerged.[40] Microbial BGCs can come in many sizes ranging from a few to a few dozen genes. Amongst genes present in a BGC are those encoding core and tailoring enzymes. Core enzymes can be classified into two groups. The first are the polymerizing enzymes mentioned above, including PKSs, NRPSs, and terpene synthases/cyclases (TS/TC). Accompanying these polymerizing core enzymes are enzymes responsible for release or cyclization of the products, which can often be a domain of the core enzyme. The second type of core enzymes are those

that do not perform polymerization, but catalyze reactions that are class-defining, including dimethylallyl tryptophan synthases (DMATs) involved in indole alkaloid biosynthesis, cyclodipeptide synthases (CDPSs) that form diketopiperazines, phenylalanine aminomutases (PAMs) that generate cinnamic acid, deoxysugar and aminosugar modification enzymes that initiate aminoglycoside biosynthesis, ribosomally synthesized and post-translationally modified peptide (RiPP) modifying enzymes, etc.

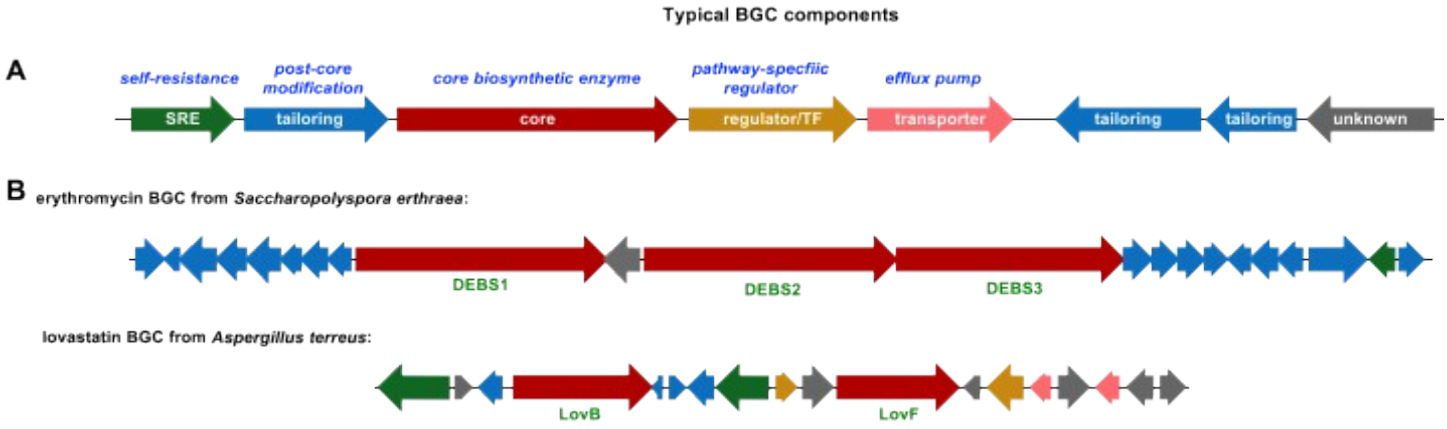


Figure 5. (A) Organization and constituents of a typical microbial biosynthetic gene cluster. (B) Examples of the erythromycin and lovastatin BGCs are shown as examples of bacterial and fungal BGCs, respectively.

Genes encoding the core enzymes in a BGC are flanked by those encoding tailoring enzymes that modify the core structures into functionally rich, structurally diverse, and biologically active natural products. These enzymes include: i) transferases such as acyltransferases, methyltransferases, glycosyltransferases, and aminotransferases, etc.; ii) cyclases and hydrolases that catalyze aldol cyclizations, Dieckmann condensations, pericyclizations, epoxide-mediated cyclizations, etc.; iii) enzymes that catalyze other redox-neutral modifications such as isomerization, epimerization, dehydration, etc.; iv) redox active enzymes, including those that catalyze both two-electron reactions such as ketoreduction, enoyl reduction, flavin-dependent hydroxylations, etc.; and one-electron reactions. The latter are catalyzed by metalloenzymes such as P450 monooxygenases, nonheme-iron-dependent enzymes, radical-SAM enzymes, and others. In addition, BGCs may also encode efflux pumps, self-resistance enzymes (SREs), and regulatory proteins dedicated to the pathways. Lastly, many genes encoding hypothetical proteins (HPs) or those predicted to have domains of unknown function (DUFs) can be found in a BGC. Taken together, the clustering of genes encoding core enzymes, tailoring enzymes, and other ancillary proteins have served as the basis for many of the bioinformatics tools that have emerged in the post-genomic era. With these tools,

it has become clear that the number of BGCs far outnumber natural products that have been identified with classic methods. It is estimated that > 90% of the BGCs are cryptic, silent, or orphaned, depending on the preferential adjective of the researcher.[1] While BGC identification from any sequenced organism, especially microorganisms, is now routine and can be accomplished with several clicks on a computer, one grand challenge is how to prioritize the vast number of BGCs that have emerged for investigation.

1.4 Scope of this Review

In this review, we aim to cover recent approaches in synthetic biology-driven natural product discovery and engineering. This is an expansive topic with an explosion of research activities from many disciplines in the last decade. As a result, a review cannot cover all the reports in scientific literature and omissions are inevitable. We hope the readers can get an appreciation of the multidisciplinary nature of the field and walk away inspired by the exciting advances Sir Hopwood anticipated in his landmark 1984 paper.[37]

A typical workflow for synthetic biology-driven natural product discovery is shown in Figure 6. The workflow can be divided based on whether the project is a genome mining campaign with no initial lead compounds, or one that is focused on elucidating pathways of a known compound for enzyme discovery and pathway engineering. In either case, the genome and/or transcriptome of the producing organism must be sequenced. This information is then analyzed by a suite of computational tools to identify potential BGCs, and/or coregulated transcription of candidate genes. These efforts are assisted by annotation of individual gene functions through homology-based comparison or machine-learning approaches such as AlphaFold.[41] Part of Section 2 of this review will focus on recent developments in these bioinformatics tools.

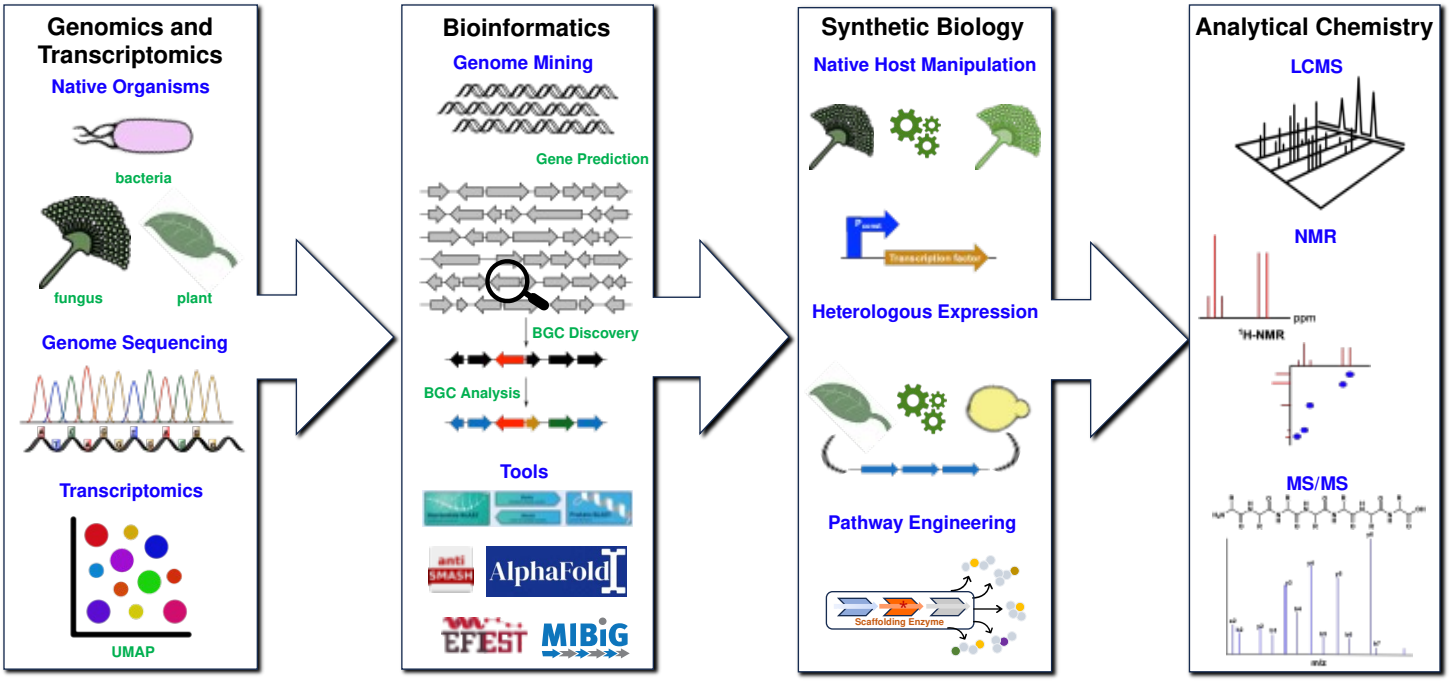


Figure 6. Simplified workflow of synthetic biology approaches to study and engineer natural products.

In the case of genome mining of hitherto unknown compounds, prioritization of BGCs based on potential structural novelty, new enzyme chemistry and/or biological activities can be conducted to arrive at candidate BGCs. The identified candidate BGCs can then be investigated either in the native producing host or a heterologous host. In the case where a native host can be obtained, growing the organism in different media followed by both transcriptomics and metabolomics analysis may lead to mapping of a BGC to a metabolite. If the BGC is not expressed under these conditions, transcriptional activation and/or additional tools can be used to "turn-on" the expression of the BGC followed by metabolite analysis. Alternatively, the BGC of interest can be entirely transplanted into a model heterologous host for direct expression of the genes followed by metabolite detection. This could allow complete refactoring of the BGC in a fast-growing host with a relatively clean metabolic background. New metabolites that emerge can be detected, isolated, and structurally characterized. Section 2 will also briefly cover new approaches in mass spectrometry and NMR that enable more rapid dereplication and structural elucidation.

When a target molecule is known, elucidation of the biosynthetic enzymes starts with a chemically reasonable biosynthetic hypothesis of the pathway. If the metabolite is observed in the native host, the classical approach of genetic inactivation to abolish metabolite production, followed by genetic complementation, remains highly useful given the organism is genetically tractable. In many examples shown in this review, heterologous expression of the core enzyme or

the entire BGC can provide a direct link between the BGC and the metabolite. Heterologous reconstitution of different combinations of biosynthetic genes can rapidly delineate the order of individual steps, as well as assigning an enzyme to each chemical transformation step. In the case of plant compounds of which the biosynthetic genes are not clustered, co-transcriptional analysis in specific tissues is often performed to identify gene candidates that may encode biosynthetic enzymes. This is followed by complete reconstitution of the pathway in a fast-growing organism as an alternative host for harvesting the high value plant compounds. Heterologous hosts can be further optimized for the high titer production of such natural products, often well exceeding titers that can be obtained in the native hosts. Section 3 of the review will focus on recent developments in heterologous expression in bacterial, fungal (including yeast), and plant hosts.

With the pathway established in either native or heterologous hosts for a metabolite of interest, synthetic biology approaches can be applied to engineer the pathway for production of new-to-nature compounds, or derivatives of the natural product that may be of interest. Well-established methods such as precursor-directed biosynthesis (PDB), mutational biosynthesis (mutasynthesis), biotransformation, and pathway engineering have all resulted in considerable expansion of the chemical space associated with natural products. Recent bioinformatics and machine learning approaches have also led to newfound successes in the engineering of modular megasynthases such as PKSs and NRPSs.[42] Section 4 of the review will cover examples in the engineering aspect of natural product biosynthesis.

2. Advances in Tools for Natural Product Discovery

2.1 Bioinformatics and Machine Learning Tools

2.1.1 BGC Prediction

Natural products are produced from bacterial, fungal, plant, and marine organisms by biosynthetic enzymes, which are encoded by biosynthetic genes. The ability to rapidly and cost-effectively sequence genomes and transcriptomes has unlocked the potential of genome mining. By annotating the biosynthetic genes and predicting corresponding biosynthetic enzymes, one can estimate the number and types of natural products that can be produced by a single organism. Advancements in data science and statistical models have enabled the development of many computational tools. These tools, such as GLIMMER,[43] GeneMark,[44] and Augustus,[45] can predict open reading frames and intron-splicing sites in different organisms. They are based on a hidden Markov model (HMM) which is trained with a dataset between known DNA and known genes to assist in accurate prediction of amino acid sequences from DNA sequences. A powerful way to predict the function of proteins encoded in biosynthetic genes is via alignment tools. Similarities in the amino acid sequence to characterized proteins can be used to predict biosynthetic functions. A key algorithm essential for functional prediction is Basic Local Alignment Search Tool (BLAST),[46] which can calculate alignments not only between a set of sequences, but also the entirety of the NCBI database.

The clustering of genes encoding core enzymes, tailoring enzymes, and other ancillary proteins has served as the basis for many of the bioinformatics tools that have emerged in the post-genomic era of natural product research. Many computational tools have been developed to predict both BGCs and their associated products from genomic sequences. Over time, bioinformatics tools have become more sophisticated and more user friendly. A list of widely used computational tools is shown in Table 1 and briefly described below.

SMURF: "**S**econdary **M**etabolite **U**nknown **R**egions **F**inder" is an early tool developed for finding BGCs.[47] The algorithm mines through fungal genomes for BGCs based on the colocalization of a core biosynthetic gene, tailoring and accessory genes in the genome. Core genes are identified using an HMM search for fingerprint domains present in NRPS or PKS enzymes. Genes adjacent to that encoding the core enzyme are collected to form the "hit" BGC. However, in some cases, neighboring genes can be predicted incorrectly to be part of the BGC.

antiSMASH: "**anti**biotics and **S**econdary **M**etabolite **A**nalysis **Sh**ell" is a benchmark genome mining tool that takes genomic data and identifies potential BGCs.[48,49] The tool starts with a workflow that predicts amino acid sequences from genomic sequence. Gene function is then predicted using a profile HMM to align primary sequences to key protein sequences associated with preset domains. The clustering of signature biosynthetic genes near pre-encoded class types defines a BGC. These class types have been pre-identified to 80+ types of chemical classes. After initial identification, the BGC is further analyzed for domain organization, comparison to other known BGCs, and can be graphically visualized. Since its initial debut as a tool to analyze bacterial genomes, antiSMASH has continued to evolve. It has been applied to the fungal genome mining with fungiSMASH.[50] A separate workflow known as plantiSMASH has also been developed which considers the unique clustering that is associated with plant genes along with the deviation of HMM profile that were developed with respects to bacteria and fungi.[51] Overall, because of its robust application, antiSMASH has been implemented into the workflow of many other tools developed later on. A recent study showed the application of antiSMASH for large scale genome mining in *Streptomyces* species (Figure 7).[52] Inputting the available *Streptomyces* genomes into the tool revealed a high diverse set of BGCs. Based on differential patterns in BGCs for already characterized bioactive compounds, a library of novel BGCs that can be explored and characterized was identified.

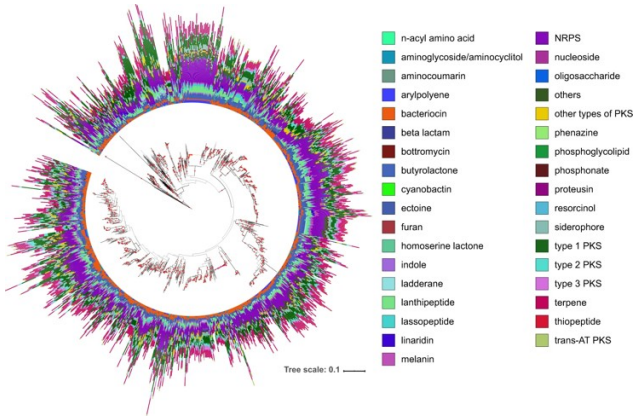


Figure 7. Distribution of more than 30 different natural product classes of BGCs from over 1000 Streptomyces strains enables strain-level comparison between number and classification of BGCs.

Reprinted from Belknap, K. C.; Park, C. J.; Barth, B. M.; Andam, C. P. Genome Mining of Biosynthetic and Chemotherapeutic Gene Clusters in *Streptomyces* Bacteria. Sci Rep 2020, 10 (1), 2003. Copyright 2020 Springer Nature.

Tool	Natural Product Classes	Kingdoms	Gene Prediction	Domain Identification	BGC/Core Identification
SMURF	nonribosomal peptides, polyketides, NRP-PK hybrids, indole alkaloids	fungal	no	pHMM	based on key domains rules
antiSMASH	most major biosynthetically characterized NP classes (81 BGC types)	bacteria, fungal, plant	Glimmer	pHMM	based on key domains rules
PRISM	nonribosomal peptides, polyketides and RiPPs, expanded to β-lactams, alkaloids, phosphonates, cyclodipeptides, bisindoles, and aminoglycosides	fungal, bacterial	ORF	pHMM	based on key domains rules
ClusterFinder	nonribosomal peptides, polyketides, NRP-PK hybrids, terpenes and notable for improving mining of saccharides	bacterial	Glimmer	pHMM	trained two-state HMM
BAGEL	RiPPs	bacterial	ORF	pHMM	alignment against core peptide database
RiPPMiner	RiPPs (including lasso peptides)	n/a	n/a	SVM classifier for class identification	SVM classifier for RiPPs precursor
RODEO	lasso peptide (RiPPs)	dependent on known RiPPs	n/a	pHMM	SVM classifier for RiPPs precursor and structures
DeepRIPP	RiPPs	bacterial (trained)	NLP	NLP	Natural Language Processing (NLP)
GECCO	nonribosomal peptides, polyketides, RiPPs, terpenes, alkaloids, saccharides	bacterial	ORF (Prodigal)	pHMM	For BGCs, NLP: Conditional random field. RF classifiers for class type
DeepBGC	nonribosomal peptides, polyketides, RiPPs, terpenes, alkaloids, saccharides	bacterial	ORF (Prodigal)	pHMM	Bidirectional Long Short-Term Memory recurrent neural network and RF classifiers for class type
BiGCARP	nonribosomal peptides, polyketides, RiPPs, terpenes,	bacterial (trained)	Prodigal (if unannotate	ESM-1b (graph	self-supervised neural network

	alkaloids, saccharides	d)	convolutional neural network)

Table 1: Comparison of genome mining tools developed. Each tool can detect BGCs of classes of natural products from specified species. While older tools use probabilistic modeling, newer tools aim to apply machine learning for identification.

PRISM 4: "**Pr**ediction **I**nformatics for **S**econdary **M**etabolomes" predicts and translates ORFs to predict biosynthetic domains and enzymes. [53] The developers curated a library of HMMs to aid in the identification of enzyme domains that range from those in core biosynthetic enzymes and tailoring enzymes. Similar to antiSMASH, clustering of domains is dependent on a greedy algorithm that iteratively searches along the ends of the core for biosynthetic enzymes. The PRISM algorithm can predict unique structural features of the product of a BGC, such as starter units for assembly-line megasynthases such as PKSs and NRPSs. The ability to encode biosynthetic knowledge has made a tool like PRISM a bridge between genomic data and chemical structures. Earlier versions of PRISM were specialized in structural prediction of polyketides and nonribosomal peptides, but with recent updates to the algorithm and HMM libraries in 2020, its predictive capabilities have expanded to other classes of natural products including alkaloids, phosphonates, cyclodipeptides, and aminoglycosides.

ClusterFinder: ClusterFinder is a tool that mines for BGCs encoding natural products of both known and unknown classes.[54] The previously mentioned tools use algorithms that define BGC based on "rules" that are specific to known classes. In order to decouple from this rule-based approach to find unknown natural product classes, the workflow for ClusterFinder is centered around a two-state HMM, with one hidden state corresponding to BGCs (BGC state) and a second hidden state corresponding to the rest of the genome (non-BGC state). Once a BGC has been found using ClusterFinder, annotation of the BGC can be performed using tools like antiSMASH. For instance, ClusterFinder was recently used to mine lanthipeptide and lassopeptide RiPPs with antimicrobial activities.[55] Mined BGCs were analyzed using antiSMASH and BiG-SCAPE (discussed in a later section) to specifically filter for lanthipeptide producing BGCs. Characterization and antimicrobial studies were performed to reveal a library of such compounds from the human microbiome.

2.1.2 Recent Advances in Prediction Tools

2.1.2.1 Prediction and Annotation of RiPP BGCs

ClusterFinder and antiSMASH have demonstrated capabilities to discover BGCs associated with RiPP natural products. The broader definition of RiPPs makes it a very large family, having members in nearly all kingdoms of life, as well as having many major subfamilies such as lantipeptides, lasso peptides, thiopeptides, sactipeptide and graspetides.[56] Examples of RiPPs from different subfamilies include: nisin A (**21**, lanthipeptide),[57] thiomuracin A (**22**, thiopeptide),[58] microcin J25 (**23**, lasso peptide),[59] chryseoviridin (**24**, graspetide),[60] and thurincin H (**25**, sactipeptide) (Figure 8).[61] Identification of the precursor protein is challenging as there are no generalized rules for the sequence, except in fungi where a single precursor protein may contain multiple repeats of the core peptide.[62] However, the PTM enzymes associated with RiPP maturation can be identified through sequence homology via a RiPP precursor recognition element (RRE).[63] As a result, workflows for mining RiPP BGCs have been modified to computationally code this search strategy.

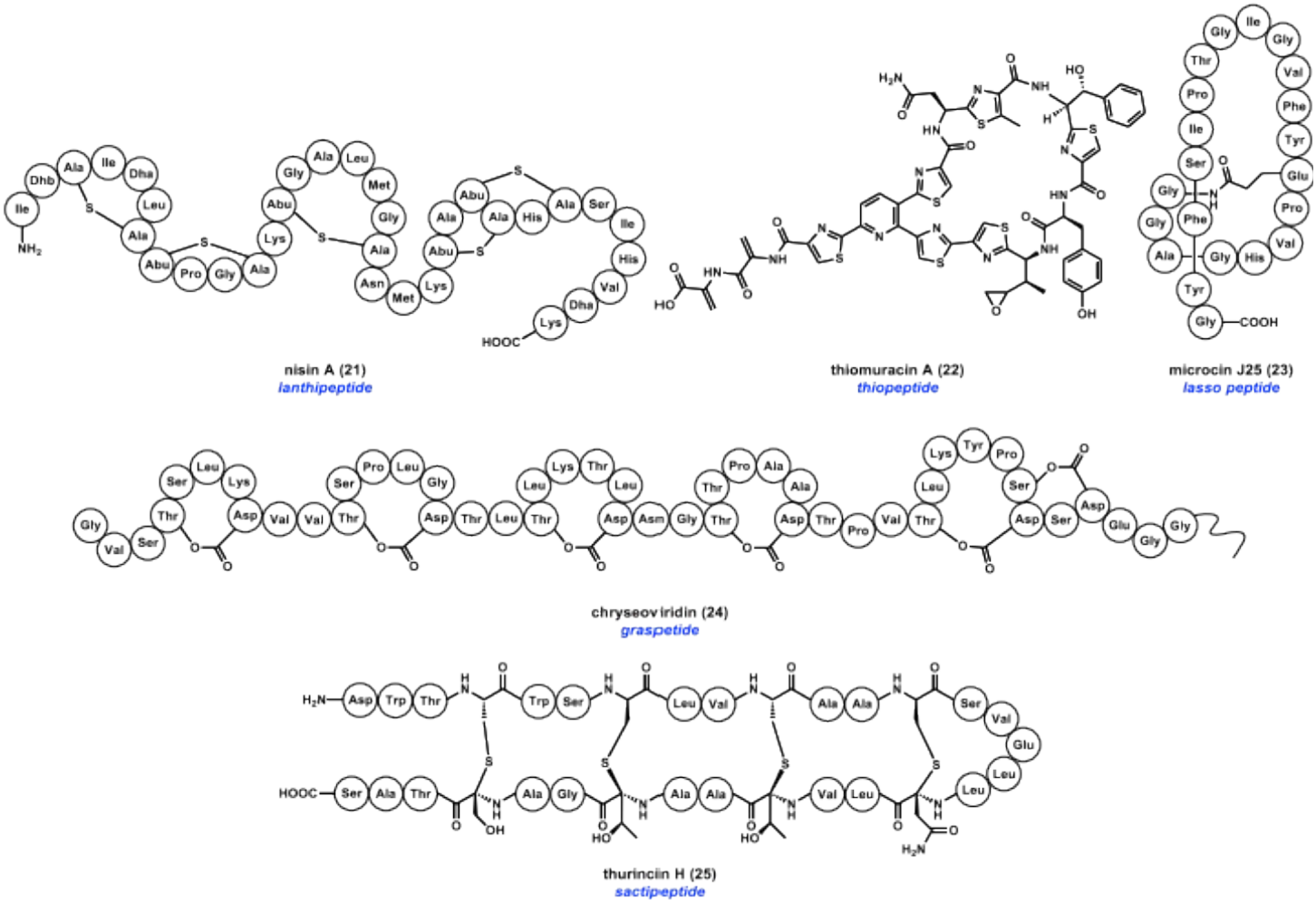


Figure 8. Examples of natural products from different RiPPs subfamilies.

BAGEL: "**BA**cteriocin **GE**nome mining too**L**" is one of the most notable tools developed for RiPP mining.[64] ORFs are predicted from genome data and compared against a preset library to predict putative RiPP BGCs, dependent on the presence of known RiPPs motifs, homology to known RiPPs, and class-specific PFAM domains, including leader sequences. Since its initial launch, the tool has been upgraded to incorporate more features to aid in the discovery of new RiPPs.[65] For instance, ORFs are predicted based on any of the six potential ORFs having a valid start codon and can be supplemented with RNAseq data for analysis. Further modifications to the program include promoter and terminator prediction and alignment to a database of known core peptides.

RiPPMiner: RiPPMiner is an algorithm in which computational genome mining shifts towards the application of machine learning.[66] The model is trained from a database of known RiPPs using a support vector machine (SVM) classifier to model and predict RiPPs precursor genes. Another model is used to predict the specific type of RiPP natural product that will be produced by the precursor. Cleavage sites within the precursor protein can be predicted using a different SVM model using k-mers. RiPPs that contain intrapeptide crosslinks such as lasso peptides can also be predicted by the tool using machine learning approaches.

RODEO: "**R**apid **O**RF **D**escription and **E**valuation **O**nline" focuses on the lasso peptides with a lariat three-dimensional structure.[67] Using the expected features of a BGC encoding a lasso peptide, which include a leader peptidase, lasso cyclase, and others with RRE features, BGCs can be predicted using a pHMM. Structural prediction using a SVM classifier model is incorporated to predict the lasso structures of the products including position of the isopeptide bond and the linear region that threads through the loop. RODEO has also been used for the mining of RiPPs in the sactipeptide [68] and thiopeptide families.[69]

DeepRIPP is a modular platform to mine unknown RiPPs.[70] The first part of the algorithm uses natural language processing that can identify RiPPs based on precursors. Then, the second part uses retrosynthetic processing to align predicted RiPPs to mined genes while also filtering out known RiPPs. The final component predicts the sequence and structure of mature metabolites.

2.1.2.2 Prediction of Biological Activity based on Self-Resistance Genes

One disadvantage of genome-guided mining of natural products compared to the classical, bioactivity-guided isolation, is that the new products are isolated without knowledge of possible biological activities. While screening using biological assays can yield desired activities for some compounds, many genome-mined natural products remain uncharacterized with respect to their true biological targets in Nature. To address this shortcoming, one method of genome mining is to focus on BGCs that encode a colocalized self-resistance enzyme (SRE) that confers resistance to the activities of the natural product. While many modes of self-resistance are known, including efflux pumps, molecular chaperons, etc., one that stands out for bioinformatics application is a co-clustered gene encoding a duplicated, but mutated, form of a housekeeping enzyme that is the target of the natural product. The SRE has the same function as the housekeeping enzyme (target of natural product), but is not sensitive to the natural product, thus allowing the producing cell to survive. The colocalization of such SREs with a small fraction (<5%) of BGCs is a well-known phenomenon and is documented in classic examples of lovastatin (**2**) (HMG-CoA reductase, HMGR),[71] fumagillin (methionine aminopeptidase),[72] and platensimycin (fatty acid synthase),[73] etc. Using the colocalized SRE as a beacon, many recent applications have been reported (Figure 9), including:

i) assigning biological activities to natural products that have known BGCs: griselimycin (**26**) (DNA polymerase),[74] harzianic acid (**27**) (acetohydroxyacid synthase),[75] and (-)-FR901483 (**28**) (phosphoribosyl pyrophosphate amidotransferase, PPAT);[76]

ii) identifying BGCs of natural products with known biological activity: altemicidin (**19**) (Ile-tRNA synthetase),[77] chlorflavonin (**29**) (acetohydroxyacid synthase),[78] and restricticin (**30**) (CYP51);[79]

iii) and more importantly, with a desired biological target as a beacon, discovering new natural products or rediscovering known compounds with no previously assigned biological activities: rediscovery of thiolactomycin (**31**) (fatty acid synthase),[80] aspterric acid (**32**) (hydroxyacid dehydratase),[81] roseopurpurin C (**33**) (CDK2),[82] fellutamide B (**34**) (proteasome);[83] discovery of clipibicyclene (**35**) (caseinolytic protease),[84] pyxidicycline A (**36**) (DNA topoisomerase),[85] and phenelfamycin B (**37**) (EF-Tu).[86]

It is worth noting that in some of these examples, including the BGCs of **28** and **35**, the SRE is the only copy of the gene in the genome, hence serving both housekeeping and self-resistance roles. Through comparison of BGCs to

homologous versions found in related organisms, it is also known that some BGCs have lost the SRE during evolution. In these cases, it can be proposed that the housekeeping enzyme, which is typically encoded elsewhere on the genome, has gained sufficient mutations to confer resistance to the natural product.

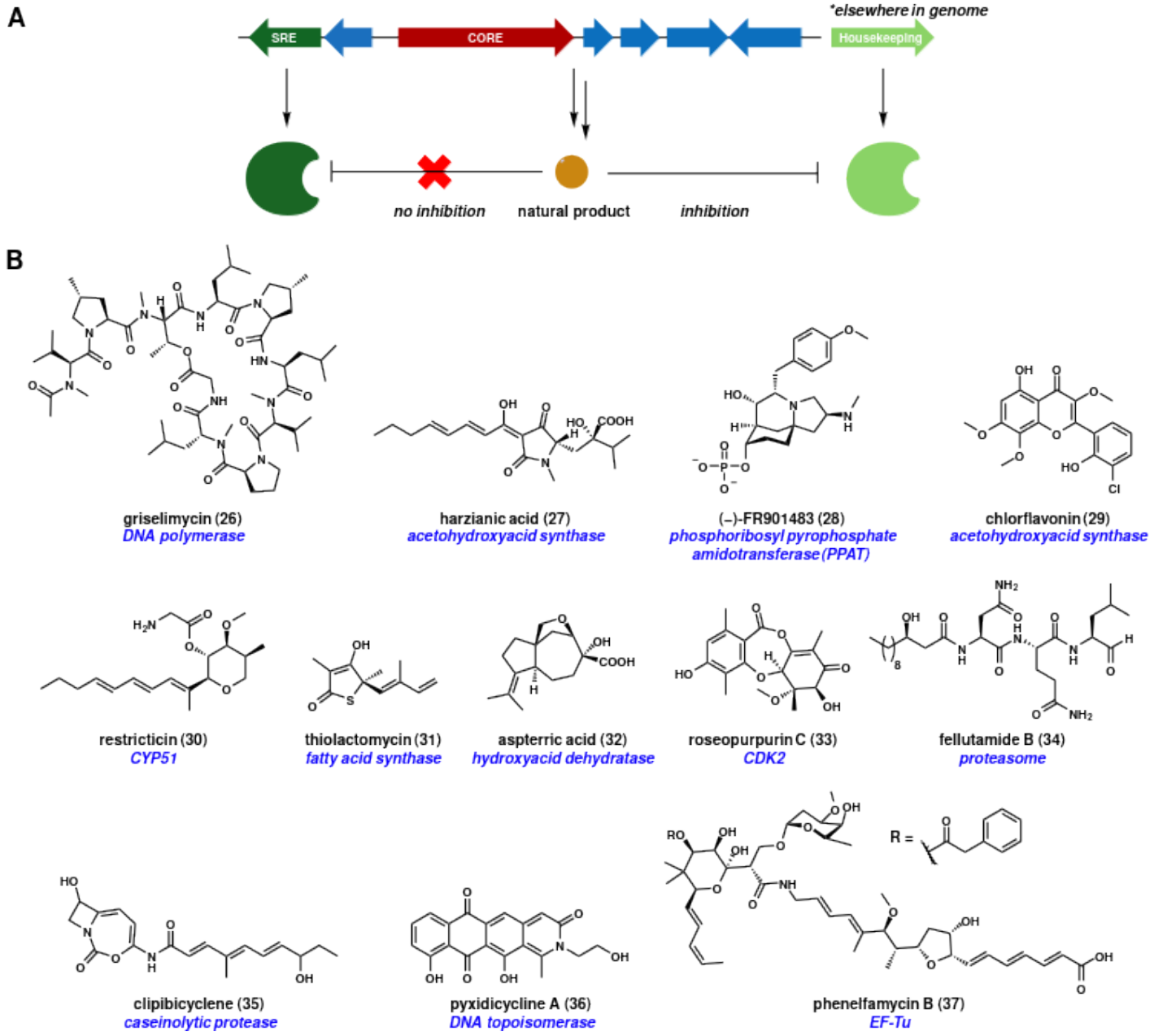


Figure 9. BGCs encode self-resistance enzymes (SREs) that are insensitive to the bioactive natural product. (A) Scheme of SRE colocalization and function; (B) Recent examples of natural products with SREs colocalized in the respective BGCs. The target/SRE functions are shown in blue.

Given that searching for the colocalization of two genes in a genome (SRE and core genes, for example) requires a bioinformatics approach, some of the papers that reported SRE-guided studies have written in-house codes to aid their searches. There are also several web-based tools that have been published and are dedicated to SRE-guided genome mining in both bacteria and fungi:

FRIGG: "**F**ungal **R**esistance **G**ene-directed **G**enome **M**ining" is an early example of a tool developed to explore the SRE-based approach.[87] The algorithm checks for copies of each gene (biosynthetic and proposed self-resistance) to help filter for potential BGCs. Further filtering occurs via the presence of a homolog of the SRE in other organisms. A final filtering step looks for only a single homolog in the genome outside of the potential BGC.

Rg3m: "**R**esistance **G**ene-**G**uided **G**enome **M**iner" is a high-throughput computational script to search for potential BGCs.[88] By specifying a proposed SRE domain and biosynthetic genes of interest, the program can mine through thousands of genomes for hypothetical BGCs. Based on the multiple occurrences of homologs to the target SRE in the genome and close proximity of one of these homologs to a core biosynthetic gene, the script outputs potential BGCs of interest. A positive control for this script was performed using a PKS as the core enzyme and an HMGR as the SRE, which identified the lovastatin (**2**) BGC. Comparing the hits suggested new natural products that target the HMGR and are structurally distinct from **2** could be discovered.

FunARTS: "**Fun**gal bio**A**ctive compound **R**esistant **T**arget **S**eeker" begins with the use of fungiSMASH to mine through genomes for BGCs.[89] Then, using prebuilt HMM profiles, housekeeping genes and potential duplicates within the genome are identified. Hits are determined based on the presence of duplicates to the housekeeping genes. The workflow also has the capacity to analyze the mined BGC and organize them into gene cluster families for further comparative analysis via BiG-SLICE.[90]

2.1.2.3 Similarity Clustering and Dereplication

The aforementioned genome mining tools have enabled the identification and annotation of BGCs, as well as the prediction of structures and bioactivities of the associated natural products. With interests and activities in genomics-driven natural product discovery growing rapidly, managing the database of both unassigned and assigned BGCs is a major challenge. Has a BGC of interest, or one that is slightly different in gene organization, been studied? Is there a known

natural product associated with a BGC? Just as effective compound dereplication is key to classic natural product discovery, keeping an accurate and up to date inventory of BGCs and associated natural products is of critical importance.

The solution to this is a public and updated database that accounts for all reported BGCs. MIBiG (**M**inimum **I**nformation about a **Bi**osynthetic **G**ene cluster) is a robust database that creates a standard set of details associated with each BGC.[91] General information includes biosynthetic class, key publications, sequences, genomic loci, functional annotations for the genes in the BGC, known natural products and associated molecular weights, structures, and bioactivities. Data specific to the biosynthetic class are noted along with functions, location and conserved domains (CDs) of each gene. The 3.0 update encompassed the advances in the biosynthesis field and improved the accuracy in annotated entries.[92] Recently, it was updated to version 4.0 to incorporate recently characterized BGCs, to curate the database via scientific community efforts, and to define standards for broad applications.[93] By setting a common format across BGCs from all kingdoms, this database makes it accessible to build new tools to dereplicate BGCs. Along the same lines, proGenomes3 creates a database of almost one million genomes of prokaryotic species and their associated BGCs.[94] Predicted functions and components of BGCs are annotated and stored to aid in transcriptomics, proteomic, and metabolomic data analyses.

With a centralized database for BGCs, efforts to develop tools to analyze and prioritize BGCs of interest became the next focus of computational analysis. Several recently developed tools further dereplicate BGCs based on clustering and subclustering. These include:

Big-SCAPE: "**Bi**osynthetic **G**ene **S**imilarity **C**lustering and **P**rospective **E**ngine" is one of the tools that has incorporated antiSMASH into its workflow. BiG-SCAPE uses the BGCs mined from antiSMASH as inputs and groups them into gene cluster families (GCFs) to understand the relationship between different BGCs.[95] Once all domains present in each BGC are identified, a correlation score is calculated based on the number of shared domain types, sequence identity between similar domains, and repetitive patterns of domain in the BGC (i.e. domains being adjacent to each other). The BGCs are then grouped into GCFs to form a similarity network. Building upon the GCF context, BiG-FAM and BiG-SliCE are newer tools to this group. BiG-SLICE (**Bi**osynthetic **G**enes **S**uper-**LI**near **C**lustering **E**ngine) was developed to increase the scalability of GCF clustering.[90] By changing the way how BGCs are compared amongst each other, this tool can reduce processing time. However, the group notes that omission of the typical pairwise sequence similarity algorithm results in

lower sensitivity of similarity alignments and decreased ability to detect point mutations in genes amongst BGCs in a GCF. BiG-FAM (**Bi**osynthetic **G**ene **C**luster **F**amilies Database) correlates the GCFs back to a database and creates an interactive tool that links BGCs to predicted natural product scaffolds, genomes, and species of origin.[96] As of June of 2024, Big-FAM had 1,225,071 BGCs in its database. Data from the BiG-FAM 2021 analysis showed the distribution of over a million BGCs across four major natural product classes: polyketides, nonribosomal peptides, RiPPs, terpenes (totaling 90%), and others (8.8%).

CORASON: "**COR**e **A**nalysis of **S**yntenic **O**rthologues to prioritize **N**atural product gene clusters" is a tool developed in parallel to Big-SCAPE to prioritize BGCs based on phylogenetic analysis of clustered BGCs.[95] The workflow takes a query gene from a BGC of interest and searches for other homologues of the query gene. BGCs containing homologues that resemble the query gene are then analyzed to create a cluster variation database (CVD). This analysis can be further visualized via a phylogenetic tree for comparisons amongst the GCFs formed via Big-SCAPE.

2.1.2.4 Machine Learning for BGC Prediction

The previous sections focused heavily on the use of the stochastic HMM to predict BGCs. With advances in the understanding of the fundamentals of machine learning (ML) models, ML has been recently applied to the bioinformatics field to predict BGCs. A recent review by Zhao and coworkers extensively discussed ML and artificial intelligence guided genome mining.[97] Here, three recent applications developed for BGC prediction are highlighted.

GECCO: "**GE**ne **C**luster prediction with **CO**nditional random fields" is a new tool in development that implements ML-like approaches to assist in genome mining.[98] Conditional random fields (CRFs) are used for natural language processing (NLP) and can be applied for parsing through genomes for new biosynthetic genes. The workflow starts similarly as many of the previously mentioned tools, with ORF prediction from genomic data, and the prediction of the ORFs function based on pHMMs. The next step is where ML is applied: the model is trained using known BGCs to recognize regions that have been segmented to be classified as BGCs and non-BGCs using CRFs. Finally, the predicted BGCs are clustered into common classes using a random forest classifier as an alternative to decision trees.

DeepBGC: Similar to GECCO, DeepBGC performs an initial analysis using ORF prediction and prediction of the domains using pHMMs. [99] Deep learning was implemented using BiLSTM (Bidirectional Long Short-Term Memory), a

processing model for NLP, that learns from forward and backwards context. Unlike previously mentioned tools, there are no rules preset to define a natural product class; this gives the tool potential to predict BGCs that are not confined to known classes as discussed in Section 1.2.

BiGCARP: "**Bi**osynthetic **G**ene **C**onvolutional **A**utoencoding **R**epresentations of **P**roteins" is another tool that uses deep self-supervised ML to mine for BGCs.[100] Compared to previous examples, this masked language model is trained on a curated set of BGCs to effectively predict BGCs independently of a curated pHHM for domain annotation.

2.1.2.5 Prediction of Protein Functions

Most of the tools described above rely on homology-based prediction of protein functions to establish GCFs, boundaries of BGCs, and possible natural product structures. Most proteins in BGCs can now be assigned with putative functions based on characterized enzyme sequences deposited in various databases. There are, however, many instances in which function of one or more proteins in a BGC cannot be predicted. These are annotated as hypothetical proteins (HPs) or domains of unknown function (DUFs). Moreover, predictions of protein function based on sequence similarity to known enzymes can be inaccurate. This is not surprising, since new functions can be evolved from well-established enzyme families with a small number of mutations.[101] Proteins with no predicted function or with an incorrectly predicted function can hamper or mislead biosynthetic investigations. On the other hand, these proteins often catalyze new reactions, including those new to the biological realm, or catalyze known reactions with novel mechanisms, and hence are of significant interest to the field. To illustrate this important aspect of genome mining, some related examples are shown in Figure 10. These include:

i) discovery of new terpene cyclases (TCs) in the biosynthesis of xenovulene A (**38**) (AsR6 that synthesizes humulene **39**)[102] and phomactin B4 (**40**) (PhmA that synthesizes phomactatriene **41**).[103] Both AsR6 and PhmA were initially annotated as HPs with no recognizable sequence motifs, including those that are well-conserved among characterized TCs;
ii) identification of new *N-N* forming enzymes in the biosynthetic pathways of kutzneride 2 (**42**) (KtzT)[104] and valanimycin (**43**) (VlmO).[105] KtzT was initially annotated as a transcriptional regulator but was shown to be a heme-containing protein that converts *N*-hydoxy-L-ornithine (**44**) to L-piperazic acid (**45**). VlmO was annotated as a

membrane-bound protein with unknown function, and was shown to synthesize hydrazine **46** from O-acylhydroxylamine precursor **47**;

iii) Lichman et al. identified a gene encoding a major latex protein-like (MLPL) protein in the genome of plants that synthesize monoterpene indole alkaloids (MIAs).[106] The protein product turned out to be the missing enzyme in the stereoselective cyclization of monoterpene precursor **48** into *cis-trans* nepetalactol (**49**), a precursor to strictosidine (**50**). In the absence of this protein, MIA heterologous reconstitution efforts were inefficient due to shunted products formed from nonselective cyclization;

iv) An annotated HP in a BGC with no core enzyme was shown to be an arginine cyclodipeptide synthase (RCDPS) that biosynthesizes L-Arg-L-Tyr cyclodipeptide (**51**) that is the starting point to the more elaborate pseudopeptide NK13650 C (**52**).[36] With minimal sequence homology to bacterial aminoacyl-tRNA dependent cyclodipeptide synthases (CDPSs), RCDPS homologs are widely found in fungi and can be considered as a new class of core enzymes.

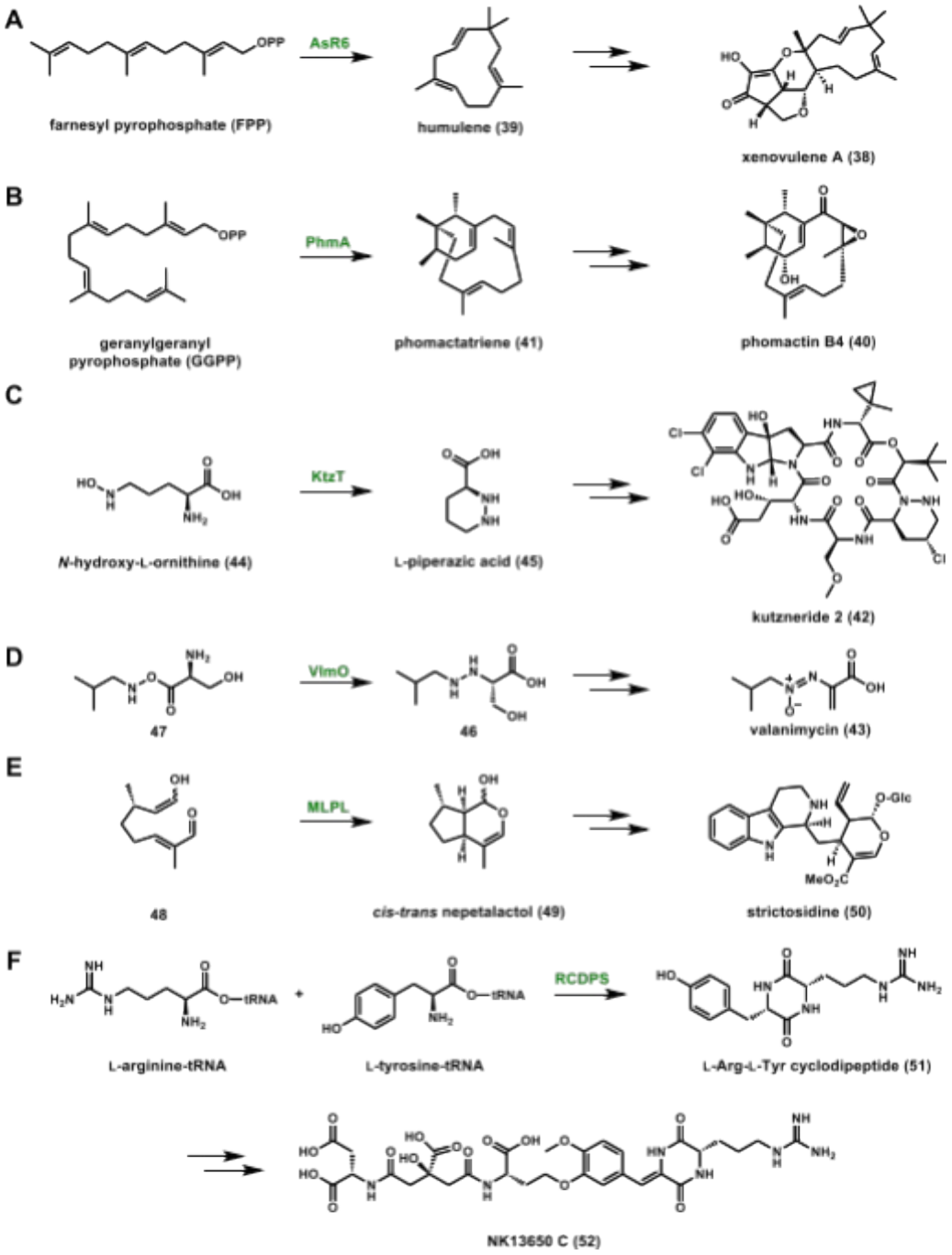


Figure 10: Examples of reactions catalyzed by enzymes that are initially annotated as hypothetical proteins or incorrect functions based on sequence alignments.

These examples demonstrate the importance of being able to correctly annotate protein functions in BGCs. ML is playing a powerful role in improving the functional prediction of proteins. No tool has changed protein structural and function prediction more than the three versions of AlphaFold.[41,107,108] AlphaFold uses a neural network-based model to predict the structure of any protein based on multiple sequence alignment and pairwise representation that has gone

through a trained neural network block and a structural prediction block.[41] This tool dramatically empowers many research labs as it can rapidly and accurately predict three-dimensional structures from amino acid sequences. Recent developments in Alphafold3 have allowed predictions for protein-protein interactions and small molecule binding modes.[107]

With databases of AlphaFold predicted structures now available, tools such as DALI and DeepFRIcan be employed to predict protein functions.[109,110] For instance, a predicted structure can be searched amongst the Alphafold Database using DALI. Annotated functions of the highest matches to the database can then be used to propose the potential function of the queried structure. This workflow becomes valuable for functional assignment in cases where predictions cannot be made using pHMM and other sequence alignment-like algorithms. Alternatively, the DeepFRI algorithm also processes sequences and protein structures to predict function based on ML. A pre-trained language model extracts key residue features from Protein Data Bank sequences while a parallel process creates a contact map formed by the pairwise proximity of each residue in protein. The coupling of these two datasets goes through a graph convolutional network to predict function from protein structure.

ProteInfer is an approach that uses deep neural networks to predict protein functions.[111] The goal of this new tool is to facilitate the use of uncurated sequence data to support model training. Many earlier models (and even the pHMM alignment algorithm) rely on a curated set of confirmed sequences associated with specific enzymatic functions. Since functional annotations of domains are only subject to the required length of the defined domain, full length proteins could be mislabeled by these previous ML tools. Using a deep dilated convolution network, ProteInfer maps a full-length sequence to protein function via sequentially dilated segmentation of proteins to look for subsequences that have higher hits. Another ML tool that predicts enzyme function is **C**ontrastive **L**earning-enabled **E**nzyme **AN**notation (CLEAN).[112] The algorithm uses contrastive learning, a self-supervised ML technique tasked with maximizing the distance with negative instances and minimizing the distance with positive instances. Positive or negative instances are sequences (and associated enzyme commission number) that are similar or dissimilar to the reference anchor sequences, respectively. This creates a system that can indicate and label an input based on the calculated distance to anchor sequences trained to the model. When tested on an annotated dataset, the algorithm outperforms previously reported tools. Experimental validation demonstrated that previously mislabeled halogenases can be predicted properly with the application of CLEAN.

All of these tools, however, still cannot predict new functions of an enzyme evolved from well-characterized folds. In these cases, biochemical assays are ultimately needed to elucidate enzyme functions in a biosynthetic pathway.

Once a function of the protein has been established through either prediction, genetics, heterologous expression and/or biochemical assays, a popular tool that can be used to identify related enzymes in other BGCs is the sequence similarity network (SSN).[113] This tool performs an alignment of all homologous sequences from the genome database against each other to create a network of relationships.[114] Visualization of an SSN can be tuned by creating a higher requirement for what it means to be related (edges), depending on sequence similarity and alignment scores. The tool can also cluster very similar sequences into single nodes. This provides users with the ability to analyze SSNs for new information about the enzyme family, including clusters, singletons, etc.[115] This has become a heavily used tool in genome mining in recent years.

A recent example of bacterial genome mining demonstrated the use of SSNs in the overall workflow.[116] Tripyrrole natural products such as prodiginines or tambjamines are synthesized from the condensation between 4-methoxy-2,2′-bipyrrole-5-carboxaldehyde (MBC, **53**) and a substituted monopyrrole such as methylamylpyrrole (MAP, **54**) or fatty alkylamines, respectively (Figure 11).[117] The enzyme responsible for the coupling step in prodigiosin (**55**) biosynthesis is PigC, a condensation enzyme that is proposed to generate diversity in this family of compounds. By constructing a SSN based on PigC followed by fine tuning the alignment score threshold, subclustering of PigC homologs based on species of origin were revealed. Using this method, the tambjamine BE-18591 (**56**) BGC was discovered from *Streptomyces albus* along with the BGCs of undecylprodigiosin (**57**) and tambjamine YP1 (**58**).

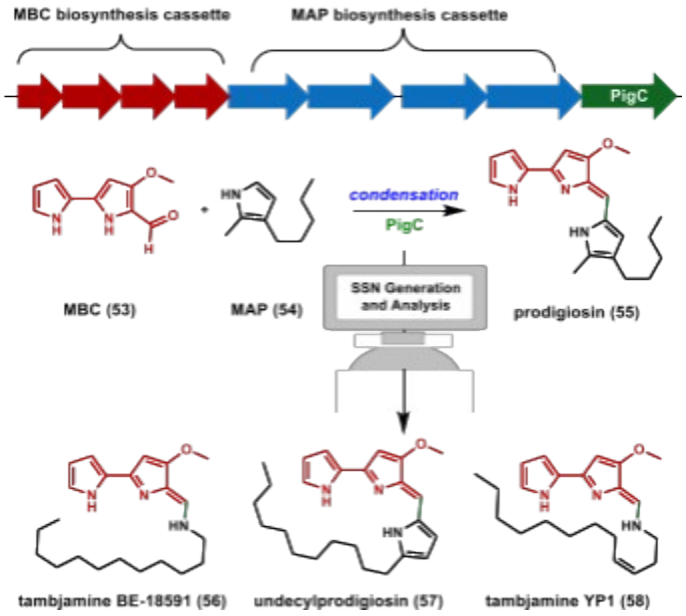


Figure 11. BGCs of tripyrrole natural products contain a condensation enzyme such as PigC that condense MBC and MAP to afford prodigiosin. By using the condensation enzyme PigC to build an SSN, new BGCs that condense MBC to different monopyrroles were discovered.

2.2 Synthetic Biology Tools for Mining Silent BGCs in Native Hosts

Given that more than 90% of the BGCs identified by the bioinformatics tools discussed above are silent under laboratory conditions, different genetic and nongenetic approaches have been applied towards activating the transcription of these BGCs. Nongenetic methods, which can be collectively termed OSMAC (one strain many compounds),[118] including growing strains on different media, coculturing with different strains, adding epigenetic inhibitors, etc., will not be discussed here and readers can refer to recent comprehensive reviews.[119,120] Here, we will present selected examples that involve genetic manipulation of BGC activation in the native hosts. As one would expect, these strategies center around artificial control of BGC transcription, either in a pathway-specific or a global manner. Many classic examples and reviews of promoter replacement or overexpression are present in the literature.[121] Heterologous expression strategies will be discussed in detail in Section 3.

2.2.1 Recent Examples in Bacteria

Genetics studies have previously identified many proteins that can lead to changes in expression levels of BGCs when overexpressed in *Streptomyces*. To expand the natural product metabolome in Actinomycetes, Wong and coworkers

used a *Streptomyces* phage-derived phiC31 integrase to introduce the following activator genes into 54 different actinobacterial strains: morphological and sporulation receptors cyclic AMP receptor protein (Crp), A-factor dependent protein A (AdpA), antibiotic regulatory protein (RedD), sporulation and antibiotics-related gene A protein (SarA), and fatty acyl CoA synthase (FAS).[122] Using such combinatorial genome integration together with OSMAC, the authors were able to obtain over 2000 extracts and discovered a significantly expanded chemical space of natural products, including tetramic acid derivatives BE-54476-A and BE-54476-B which displayed antibacterial activity.

Quorum sensing is a bacterial signaling system pertaining to cell density that communicates with its environment and the surrounding cells, and has been shown to control expression of BGCs. Rued and coworkers demonstrated this by identifying a unique short hydrophobic peptide (SHP)/Rgg system responsible for quorum sensing in *Streptococcus mutans*.[123] Gene activation/deletion experiments revealed a direct correlation between the quorum sensing system and biosynthesis of the RiPP tryglysin B. Another strategy to activate BGCs is by expressing a chimeric LuxR from the well-studied bacterial Lux operon (Figure 12).[124] Naturally, two LuxRs can bind to signaling molecules involved in quorum sensing to form a dimer and act as an activator. By engineering a LuxR dimer to be expressed in bacteria, silent BGCs can be activated when expressed under the bidirectional P_{Lux} promoter. This strategy was used to activate the production of mupirocin (**59**) and phenazine-1-carboxylic acid (**60**) in *Pseudomonas sp.* QS1027 with an engineered chimeric activator known to be involved in quorum sensing.

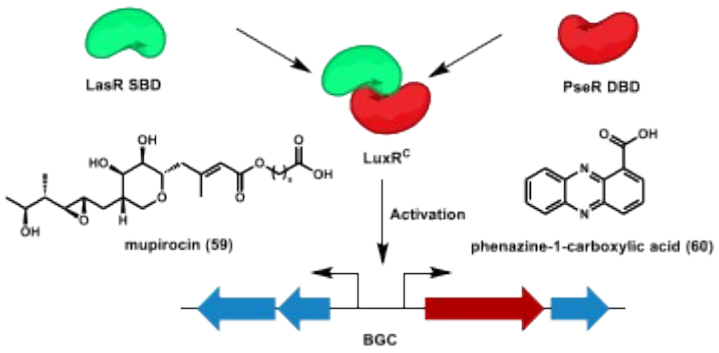


Figure 12. Chimeric LuxR is engineered as an activator in *Pseudomonas* to upregulate BGC expression.

In addition to quorum sensing molecules, expression of bacterial BGCs is affected in the presence of other natural products, a likely outcome of the "chemical warfare" between different species in the native environment. The precise chemical signals, however, are difficult to decipher and a high-throughput approach is required. Towards this end,

Seyedsayamdost developed a strategy called **Hi**gh-**T**hroughput **E**licitor **S**creening (HiTES) that screens for the upregulation of BGCs of interest in the presence of 640 potential small molecule elicitors or activators, including antibacterial compounds at sub-lethal concentrations.[125] In the initial study, a *lacZ* reporter was fused to a downstream biosynthetic gene in the BGC of interest which enabled the visual screening of BGC expression instead of untargeted metabolomics. Application of this tool was demonstrated on the known malleilactone BGC that is cryptic in *Burkholderia thailandensis.*[126] A selected group of activators upregulated the expression of BGC, which led to the isolation of a new analog malleilactone B (**61**). Subsequently, enhanced green fluorescent protein (eGFP) was used as a reporter due to the background expression of *lacZ* found in some *Streptomyces* species.[127] This led to the discovery of new natural products including the antifungal acyl-surugamide A (**62**). HiTES was most recently applied by Seyedsayamdost et al. to *Burkholderia plantarii* and *Burkholderia gladioli* cultures grown on agar plates rather than in liquid culture, leading to the discovery of burkethyl A (**63**) and B.[128]

Applications of HiTES were limited by the need to genetically modify the BGC of interest with the fused GFP reporter gene. To avoid this, HiTES coupled with imaging mass spectrometry (IMS) using laser-ablation-coupled electrospray ionization MS (LAESI–MS) was used to decouple from species specificity and eliminate the need for a reporter gene, while products from multiple BGCs can be monitored.[129] When applied to *Amycolatopsis keratiniphila*, the HiTES-IMS approach was able to discover novel glycopeptides including keratinicyclin A (**64**) (Figure 13). Keratinicyclin was shown to inhibit *Clostridioides difficile* by binding to cell wall teichoic acid, a target that is distinct from known glycopeptides such as vancomycin (**3**).[130] Another variation of HiTES is the coupling with fluorescence resonance energy transfer (FRET), which led to activation of silent BGCs in *Streptomyces clavuligerus.*[131] The authors identified potential steroid inducers and nine new compounds including clavulyne A (**65**) with DNA-cleaving and antibacterial activity.

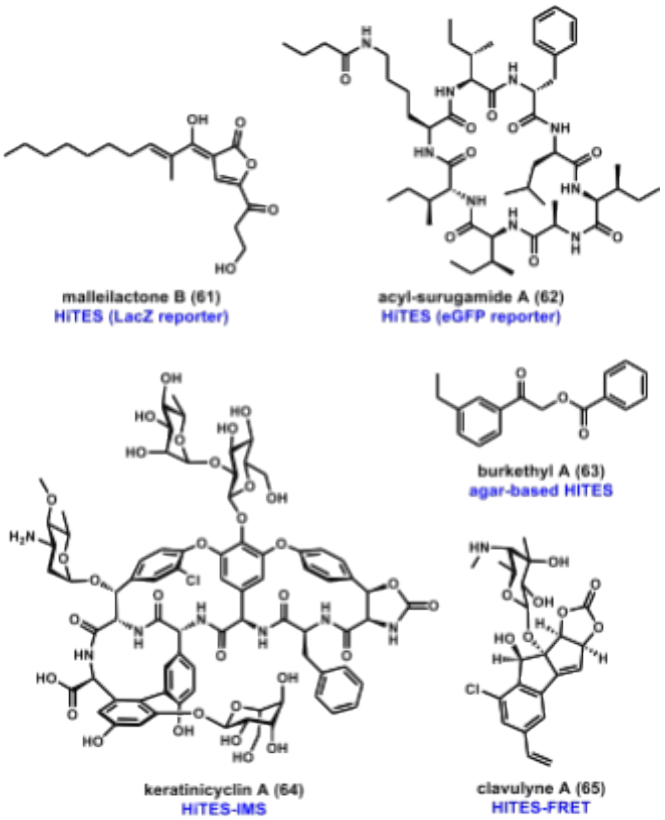


Figure 13. New natural products discovered from HiTES.

Ribosome engineering is another approach for activating cryptic BGCs. An early study showed that the spontaneous mutation of ribosomal protein S12 activated the production of actinorhodin in *Streptomyces lividans*.[132] Further studies proposed these mutations to affect binding of the signaling nucleotide, ppGpp, to RNA polymerase, which aids in directing RNA polymerase to specific promoters.[133] This was developed into a tool to activate silent gene clusters by mutating ribosomes and RNA polymerases under selection pressures, affording changes in secondary metabolite profiles. For example, rifampicin is a known inhibitor of RNA polymerase—when Huang and coworkers generated a library of strains with mutated RNA polymerases resistant to rifampicin, some of the rifampicin-resistant mutants activated the bohemamine BGC in *Streptomyces* sp. CB02009. Based on this ribosome engineering approach, Fan and coworkers reported Transcription–Translation in One (TTO), a modified approach to activate cryptic BGCs in *Streptomyces* strains.[134] They took previously reported mutation of RpsL (ribosomal protein S12) and RpoB (RNA polymerase) known for being beneficial to natural product activation and inserted them into a plasmid system to overexpress these mutated genes in *Streptomyces* species. This unlocked the production of piloquinone and homopiloquinone and pinpointed the associated

BGC.[132] Huang and coworkers have authored a comprehensive review on the application of ribosome engineering towards discovery and titer improvement in *Streptomyces*.[135]

Promoter engineering is a valuable tool to increase transcription levels of target genes. Inducible promoter systems can also be added to provide tunability of microbial cell factories, especially when the metabolite is toxic to the producing host. Combining random mutagenesis techniques (similar to that used in ribosome engineering) and a promoter-reporter system, researchers have developed a method to detect activated cryptic BGCs using reporter-guided mutant selection.[136] The reporter gene is placed under the same promoter as the gene cluster. After mutagenesis and selection based on the reporter gene, metabolite analysis can reveal new natural products. Zhang and coworkers reported a method that combines promoter engineering with recombinases.[137] Using a RED recombinase from *Burkholderiales strain DSM 7029*, the study presented a method for inserting functional promoters to uncover new RiPP natural products.

2.2.2 Recent Examples in Fungi

Filamentous fungi are multicellular species that have complex life cycles commonly involving the generation of asexual spores.[138] A majority of filamentous fungi fall under the phylum *Ascomycota*, which includes many genera such as *Aspergillus, Penicillium, Fusarium, Alternaria*, etc. These microorganisms are ubiquitous in the environment and play critical roles in host-pathogen interactions, biomass degradation, and various biotechnological applications. Filamentous fungi have been noted as prolific producers of natural products, exemplified by the discovery of penicillin from *Penicillium* molds by Sir Alexander Fleming, which ushered in the era of antibiotics.[139] Many important natural products have been identified and isolated from fungi, including the pharmaceuticals lovastatin (**2**) and cyclosporine, and the mycotoxins aflatoxin and ochratoxin. Biosynthetic investigations and genome sequencing have revealed important parallels to bacterial natural product biosynthesis. In fungal genomes, biosynthetic genes are indeed clustered in BGCs, although each gene is monocistronic and transcription is controlled via individual promoters.

Expression of dedicated or pathway-specific transcription factors is in turn controlled pleiotropically by upstream regulators of which the activities are influenced by temperature, pH, carbon and nitrogen sources, metal concentrations, and other environmental stimuli.[140] The exact combinations of conditions and stimuli that can lead to pathway-specific transcription factor overexpression in different BGCs are unknown and are difficult to replicate under laboratory conditions. To bypass this challenge in recapitulating the native conditions, direct overexpression of the regulators has

proven to be effective. Notably, Hertweck and coworkers overexpressed ApdR, a Gal4-like transcription factor in the *apd* BGC, under an inducible promoter in *Aspergillus nidulans* to overproduce polyketide-amino acid hybrids aspyridones A (**66**) and B (**67**) (Figure 14).[141] Using the same strategy, discovery of a library of new PKS-NRPS derived metabolites such as oxalemides A (**68**), C (**69**) and J (**70**) was accomplished by the overexpression of PoxB, a transcription factor found in the *pox* BGC in *Penicillium oxalicum*.[142] More recently, overexpression of RosH, a pathway-specific transcription factor, led to the rediscovery of the known compound roseopurpurin C (**33**), a CDK2 inhibitor.[82] Discovery of meroterpenoids berkeleyacetal D (**71**) and 11-epi-berkeleyacetal C (**72**) was accomplished by the overexpression of BerA, a transcription factor in the *ber* BGC.[143] The *ber* BGC was initially identified via genome mining of the *Neosartorya glabra* strain using meroterpenoid biosynthetic genes as beacons.

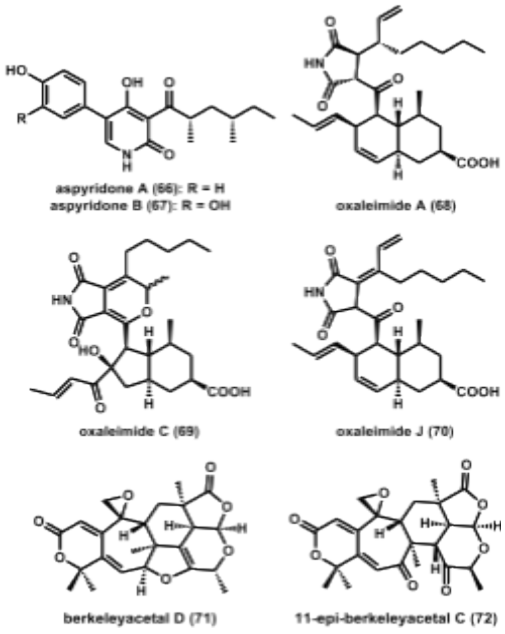


Figure 14. Examples of fungal natural products discovered via overexpression of BGC-specific transcription factors.

In addition to pathway specific transcription factors, fungal BGC expression is also controlled by global transcription regulators. The Keller lab discovered a global regulator, LaeA, that controls the expression of penicillin and lovastatin BGCs in fungi.[144] Inactivation of LaeA in *A. flavipes* led to overexpression of a NRPS BGC responsible for synthesizing two new piperazines: flavipamides A (**73**) and B (**74**) (Figure 15).[145] Yin and coworkers recently identified a new global regulator that consists of RNA-binding protein CsdA and regulator RsdA.[146] They demonstrated via transcriptomics that when the genes for the CsdA/RsdA complex are removed, global activation and deactivation of various BGCs can be

observed. Since the CsdA/RsdA complex is present in many fungal genera, this provides a new target to activate BGCs that were unaffected by other approaches.

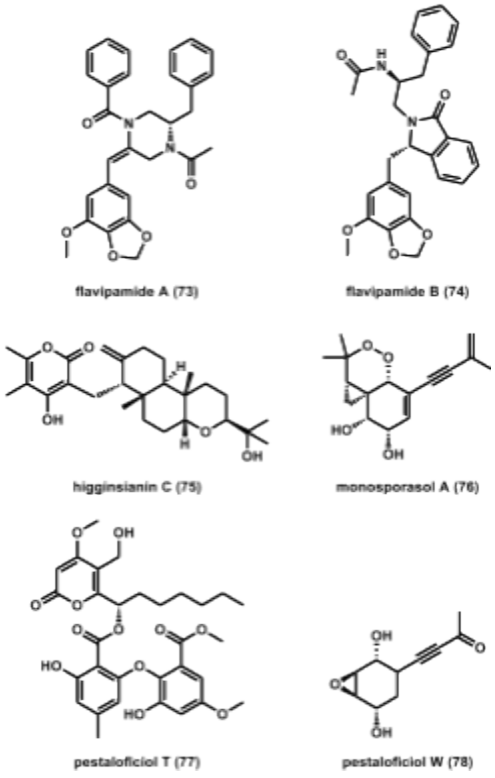


Figure 15. Examples of fungal natural products discovered from genetic manipulation of native hosts.

Epigenetics affects the ability of transcription machinery to physically access and transcribe different parts of the chromosome. For example, histone deacetylases (HDACs) are responsible for the removal of acetyl groups added to histones by histone acetyltransferases (HATs), rendering the chromatin less accessible to transcription machinery. Altering chromatin accessibility in fungi can lead to activation of silent BGCs and isolation of new natural products.[147] Keller and coworkers showed that deleting *cclA*, which is involved in histone methylation, activated the biosynthesis of multiple BGCs in *A. nidulans*.[148] By removing the same *cclA* gene in *Colletotrichum higginsianu,* new polyketide natural products including higginsianin C (**75**) were isolated (Figure 15).[149] Hu and coworkers showed that by removing the gene encoding for a HDAC in *Calcarisporium arbuscula*, 52 of 68 core biosynthetic genes were activated, leading to isolation of 10 new compounds including the depsipeptide arbumycin.[150] The same strain was used to uncover various diterpenoids.[151,152] Yin and coworkers demonstrated that by removing one of the gene encoding for HAT in *Metarhizium robertsii*, 11 new natural products (isocoumarin derivatives and NRPSs) were produced and characterized.[153] More recently, Zou and coworkers

overexpressed HAT in *Monosporascus eutypoides,* leading to the discovery of the cyclopropane-containing monosporasol A (**76**).[154] Combining the strategies, Yin and coworkers deleted PfCcl (similar to CclA) and PfHdaA (an HDAC) in *Pestalotiopsis fici*, which enabled the isolation of 15 novel polyketides including pestaloficiols T (**77**) and W (**78**).[155]

2.2.3 CRISPR-based Approaches

CRISPR-Cas based gene editing and gene targeting tools have had a major impact in natural product discovery.[156] Notably, CRISPR interference (CRISPRi) and CRISPR activation (CRISPRa) tools offer precise repression and activation of BGCs in the native hosts using simple-to-design guide RNA sequences, respectively. [157–159] For example in bacteria: i) Chappell and coworkers applied CRISPRi and CRISPRa tools to investigate various components of the jadomycin B BGC in *Streptomyces venezuelae*, and revealed a complicated gene regulatory network associated with the pathway in the native host;[156] ii) Lu and coworkers reported the FnCRISPR-Cpf1 system to regulate transcription factor expression.[160] This tool utilizes a multiplex CRISPRi system for multiple gene interference with applications for new metabolite discovery; iii) Bode and coworkers used the CRISPR-Cpf1 system to activate and refactor a safracin BGC to produce a semisynthetic precursor to ET-743 in *Xenorhabdus sp.* TS4;[161] iv) Yoshikuni and coworkers combined **C**hassis-independent **R**ecombinase-**A**ssisted **G**enome **E**ngineering (CRAGE) and CRISPR to unlock silent BGCs in *Photorhabdus luminescens*.[162] CRAGE is a tool that can integrate entire BGCs from genomes into native organism's chromosomes using the Cre-Lox recombination system.[163] When coupled with CRISPR, the BGCs can be activated or knocked out, without performing promoter replacements; and v) Zhao and coworkers demonstrated the use of CRISPR for activation of BGCs by insertion of constitutive promoters in less-studied *Streptomyces* hosts. When this workflow was applied to *Streptomyces viridochromogenes* and *Streptomyces roseosporus*, activation of various BGC resulted in production and isolation of polyketides such as **79** and **80** (Figure 16).[164] More recently, Fu and coworkers developed **A**dvanced **C**as9-media**T**ed **I**n vivo **MO**bilization and mul**T**iplication of BGCs (ACTIMOT), a method to dramatically elevate expression of BGCs by transferring the BGCs from the genome to multiple-copy plasmids. This approach led to the production of mobilipetin analogs in the native host *Streptomyces armeniacus*.[165]

CRISPR systems have also been applied to genome mining in filamentous fungi: i) Ji and coworkers demonstrated the use of CRISPR/Cas9 for metabolic engineering of *Fusarium fujikuroi*.[166] By mutating various genes involved in the biosynthesis of gibberellic acids, metabolic flux through the pathway was improved and new congeners such as gibberellic acid 7 (**81**) were identified; ii) a study investigated the utility of CRISPR-Cas9 gene editing in *Nodulisporium sp*. (No. 65-12-7-1)

for gene disruption through simultaneous transformation of transcriptional gRNA and the gene cassette containing marker DNA.[167] This allowed for rapid gene disruption via Cas9-mediated double strand breaks. This targeted gene disruption system was tested with additional fungal species, including *Sporormiella minima* (No. 40-1-4-1) to demonstrate its versatility; iii) gene disruption in *Trichoderma reesei* and other *Trichoderma* spp. based on a CRISPR-Cas9 was achieved.[168] This system consists of placing gene encoding Cas9 and gRNA under the control of U6 snRNA promoters, leading to gene disruption through frameshift mutation; iv) deletion of the citrinin BGC by CRISPR-Cas9 from *Monascus purpureus*, a valuable fungus for production of food dye Monascus Red;[169] v) a CRISPR-Cas9 system was developed for *Aspergillus wentii* to activate silent BGCs via the deletion of *mcrA* transcriptional regulator gene;.[170] vi) Chooi and coworkers reported a CRISPR-mediated transcriptional activation using the Cas12a-VPR system in *A. nidulans*.[171] Targeting of the NRPS in *mic* BGC led to discovery of intermediates in microperfuranone biosynthesis; vii) Driessen and coworkers reported the use of dCas9-VPR CRISPRa to create a CRISPR-based gene activation tool in *Penicillium* strains.[172] By targeting the *macR* transcription factor, the macrophorin BGC was activated; and viii) Gao and coworkers scaled up the application of CRISPR genome editing and developed a multiplex base-editing tool for application in fungi.[173] Using the tool to simultaneously inactivate three epigenetic regulators responsible for histone modifications in *A. nidulans*, nine new compounds were produced, including andicichorine A (**82**).

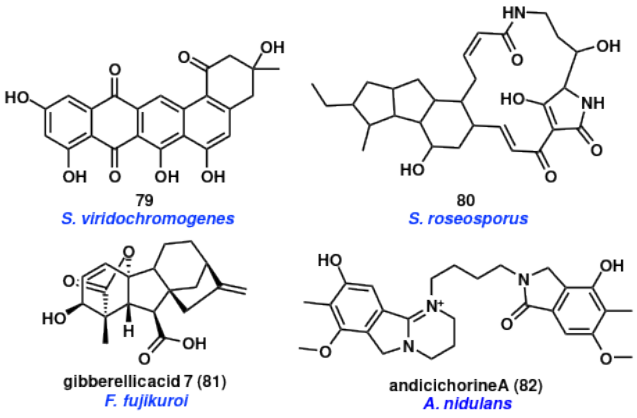


Figure 16. Examples of natural products discovered using CRISPR-based tools.

2.3 Analytical Tools for Dereplication

A key requirement in natural product discovery is chemical structure identification, which is often the bottleneck step that is technically the most challenging. The continuous development of analytical tools has enabled identification and

characterization of natural products from even sub-milligram quantities. While the previous sections discuss the various developments in searching for and "unlocking" new natural products starting from BGC sequences, this section focuses on advances in detection of new metabolites, analysis of large existing metabolomics datasets, as well as prediction of chemical (sub)structures from these datasets, with a particular emphasis on methods based on mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR).

2.3.1 Mass Spectrometry-based Approaches

MS/MS yields mass and precursor ion fragmentation data, providing researchers with key insights into chemical structures of natural products. Untargeted LC-MS/MS is a common starting point for natural product discovery and metabolomic studies alike. Extracts are analyzed via LC-MS/MS, producing large datasets containing peaks with mass-to-charge ratios (*m/z*), retention times, and characteristic fragmentation patterns. Raw spectra can be systematically deconvoluted using annotated reference spectra of known chemical structures, enabling the identification of potentially novel compounds. This process can be laborious and is dependent on the availability of existing reference MS/MS spectra and other data for structural assignment.

The development of the **G**lobal **N**atural **P**roducts **S**ocial Molecular Networking (GNPS) provided a comprehensive solution to address this problem.[174] GNPS provides an open-access infrastructure to share MS/MS data from targeted and untargeted experiments. Importantly, this data can be either processed or unprocessed, meaning that researchers are encouraged to upload all of their spectral data to make them globally accessible. Since its release in 2016, hundreds of millions of crowd-sourced MS/MS spectra have been collected in online repositories such as the GNPS-MassIVE database. Updates to GNPS, such as the GNPS Dashboard, which is a collaborative web-based MS/MS processing and analysis tool, continue to expand the accessibility of MS/MS data processing and analysis.[175]

However, with the development of GNPS and related tools, a new bottleneck has emerged: manually solving and assessing the novelty of spectra cannot keep up with the ever-increasing growth of spectral databases. As such, recent analytical developments have primarily been computational, with advances in ML algorithms, neural networks and other cutting-edge computational tools enabling large unannotated MS/MS datasets to be analyzed, processed, and utilized for rapid spectral annotation and structural prediction. These advances have led to powerful data analysis software packages, including i) MetaboAnalystR, which provides a comprehensive LC-MS/MS global metabolomics workflow, from the

processing of raw spectra to the functional interpretation of identified compounds.[176] The open-source R-based tool contains features that enable facile spectral annotation and compound identification; and ii) AntDAS-DDA, which is a new untargeted metabolomics tool that enables MS/MS data acquired through data-dependent acquisition mode in UHPLC-HRMS (ultra-high-performance liquid chromatography-high resolution mass spectrometry) to be readily processed. The platform incorporates various algorithms, such as feature extraction and fragment ion identification, which enable the rapid processing of UHPLC-HRMS untargeted metabolomics data.[177] Key computational advances will be briefly discussed below. For additional background, the review by Tasdemir and coworkers is suggested.[178]

2.3.1.1 Database Search Tools

Development of advanced database search tools is one solution to tackling the challenge of transforming large unannotated databases into a more "usable" form. Because natural products belonging to a certain class contain conserved core structures (i.e. the indole core in alkaloids such as serotonin and lysergic acid), related compounds have a common fragmentation "fingerprint." Therefore, being able to search databases for specific fingerprints allows researchers to approach the task of annotating an unknown spectrum with an understanding of the core NP structure or structural motif.

MassQL: "**Mass** Spec **Q**uery **L**anguage" is a comprehensive open-source database search tool. Specific search parameters, including *m/z* of fragments and mass tolerance, can be used to search publicly available MS/MS datasets.[179] MassQL is both scalable and customizable, allowing users to search raw (i.e. unannotated) data. The utility and flexibility of MassQL has been demonstrated by Reher and coworkers to identify novel endolides, a type of cyclic tetrapeptide containing an unusual *N*-methyl-3-(3-furyl)-alanine fragment (**83**), from *Stachylidium bicolor* (Figure 17).[180] By inputting endolide-specific fragmentation patterns into their search query, two new endolides, endolide E (**84**) and F (**85**) were identified from available MS/MS data. In another study, Castro-Gamboa and coworkers incorporated MassQL into a workflow seeking to identify new depsipeptides from the fungus *Fusarium oxysporum*.[180] A characteristic hexadepsipeptide fragmentation pattern from known annotations was inputted into their MassQL search query. This enabled the identification of 18 analogs of the mycotoxin beauvericin (**86**) from available MS/MS data, including 4 analogs that were not identified through other computational methods.

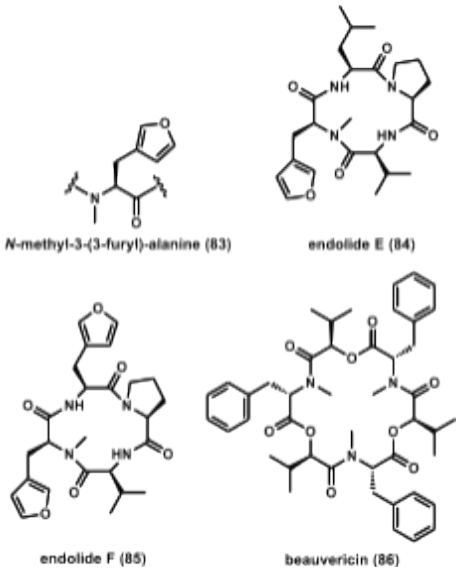


Figure 17. New endolides and beauvericin identified using MassQL-integrated approaches.

MASST and MASST+: "MASs **S**pectrometry **T**ool" developed by Dorrestein and coworkers enables searching of many publicly available spectral libraries for a query MS/MS spectrum or dataset.[181] Users can input *m/z* and intensity of the peaks in an MS/MS spectrum into MASST to identify any spectral matches. This capability is crucial for annotating convoluted spectra from untargeted studies, as already reported (and annotated) spectra in a given dataset can be readily identified. However, the process of finding a spectral match is computationally intensive, as the query spectrum must be systematically matched to every single spectrum in the repository. The recent development of MASST+ by Dorrestein and Mohimani, improved the scalability of MASST using a dot-product algorithm.[182] MASST+ only examines a small proportion of spectra in the indexing table that contain peaks similar to a given peak in the query spectrum. By narrowing the focus of the search, MASST+ boasts processing speeds up to 2-3 orders of magnitude greater than other database search tools.

Other MASST-related search tools have recently been developed. For instance, microbeMASST enables the taxonomic origin of an MS/MS spectrum to be identified. Researchers can input a query spectrum of unknown taxonomic origin into microbeMASST, which searches for a match in a curated MS/MS database of >60,000 microbial monocultures.[183] This matching does not require a query spectrum to be annotated, allowing the taxonomic origin of an MS/MS spectrum to be directly identified from untargeted metabolomics experiments. Another MASST-based tool, foodMASST, provides metabolomic analysis of foods and beverages that are consumed in our everyday diets.[184]

MS2Query: When researchers propose a putative structure from MS/MS, it is common to search for this structure in existing spectral databases for dereplication purposes.[185] However, it is frequent that searching for an exact structural match does not yield a hit. MS2Query aims to overcome this problem through ML, allowing for analogs of the inputted chemical structure to be identified. MS2Query incorporates two key algorithms, Spec2Vec and MS2DeepScore. These algorithms are similar and work under the premise that structurally similar molecules will have similar MS/MS spectra as an alternative to attempting to find direct hits.[185–187]

2.3.1.2 Annotation Tools

Metabolite annotation is the process by which metabolites are assigned to a given peak in an MS/MS spectrum. Natural product discovery workflows typically begin with untargeted LC-MS/MS analysis, which generate large MS/MS datasets. These spectra can contain hundreds to thousands of unique peaks, with each peak corresponding to a unique fragment. The process of annotating these spectra, which requires assigning a compound to each peak, is challenging and time consuming and relies on various independent pieces of data, including fragmentation patterns, chromatographic retention times and reported *m/z* in online databases.[188] Typically, spectral annotation begins by identifying the similarities between the unannotated spectrum and existing annotated spectra. The previously discussed database search tools are therefore of great utility in this stage of the annotation process. Some of annotation tools developed in the field include metID, IpaPy2, MetaboAnnotatoR, Knowledge-Guided Multilayer Network (KGMN) and MetEx.[189–193]

"**K**nowledge-**G**uided **M**ultilayer **N**etwork" (KGMN) is an annotation approach that combines three distinct networks to facilitate accurate metabolite annotation.[192] The first network is a knowledge-based metabolic reaction network, which relates unknown metabolites to existing biochemical reactions and data. Second, the knowledge-guided MS/MS similarity network identifies similarities in the MS/MS spectra of different metabolites. Lastly, the global peak correlation network enables KGMN to accurately assign MS/MS peaks by identifying common adduct peaks. KGMN was validated in the successful annotation of metabolites found in an *in vitro* enzymatic reaction mixture (>80% accuracy compared to *in silico*, i.e. traditional computational, methods). **M**etabolic ***in silico*** **N**etwork **E**xpansion (MINE 2.0) also utilizes biochemical data to enable compound annotation.[194]

Since large MS/MS databases are not completely comprehensive, automated annotation tools can leave a handful of spectra in an extract unannotated. The "**CO**nfidence of **S**mall **M**olecule **I**dentifications" (COSMIC) workflow enables

annotation of unreported natural products, by using CSI:FingerID structure search tool to assign structures to spectra.[195,196] Annotations are subsequently assigned a confidence score, allowing the accuracy of annotations to be assessed. Other tools such as "**C**omposite **S**pectra **A**nalysis" (IDSL.CSA) enable annotation of MS spectra without requiring MS/MS data.[197] IDSL.CSA processes multiple MS datasets to generate composite spectra, enabling a metabolite to be mapped to a certain peak based on its *m/z* and retention time. Then, MS repositories can be searched to characterize the peak from the composite spectrum.

2.3.1.3 Molecular Networking Applications

The development of GNPS highlighted the utility of molecular networks/networking (MN) in natural product discovery. MNs use spectral similarity as a proxy for molecular similarity, allowing researchers to graphically visualize the similarity between spectra. This enables large MS/MS datasets from untargeted metabolomic studies to be analyzed in a high-throughput manner. While the construction of MNs was previously limited to computational chemists, user-friendly protocols that describe how to create MNs using GNPS in detail are now publicly available.[198] Tools such as "**F**eature-**B**ased **M**olecular **N**etworking" (FBMN) and the nearest neighbor suspect spectral library are recent examples of the application of MN to draw relationships between unannotated spectra and existing annotated spectra, and to propose analog structures.[199]

Similarly, the "**S**tructural similarity **N**etwork **A**nnotation **P**latform for **M**ass **S**pectrometry" (SNAP-MS) relates unknown MS/MS spectra to the corresponding natural product class based on common features such as fragmentation.[200] These groups are then compared to the chemical similarity grouping in the Natural Product Atlas, allowing for natural product classes to be distinguished and annotated without requiring reference spectra. The "**S**tructure-**G**uided **M**olecular **N**etworking **S**trategy" (SGMNS) utilizes MNs to annotate MS/MS spectra by using a global connectivity molecular network to assign structurally characterized compounds in existing databases to a chemical fingerprint.[201] Other tools such as the "**M**ass-**INT**erpreter" (IDSL_MINT) enables molecular fingerprints to be assigned from MS/MS spectra, as opposed to chemical structures.[202]

2.3.1.4 Machine Learning Tools

ML is playing an increasingly important role in the annotation and dereplication of natural products from dense MS/MS datasets. The flurry of activities has led to excellent recent reviews by Medema[170] and Plisson,[171] which provide

comprehensive coverage of applications of ML and artificial intelligence to natural product discovery, with detailed subsections on the role of ML in MS-based analytical workflows. Here, only a few examples will be briefly discussed.

ML allows prediction of molecular structures from MS/MS data. SIRIUS 4 enables the identification of molecular formulas with high accuracy and the prediction of metabolite from MS/MS data with reasonable (~70%) accuracy.[203] The metabolite prediction capabilities of SIRIUS 4 are due to the integration of CSI:FingerID, an older ML tool that trains support vector machines (SVMs) on molecular properties associated with molecular fingerprints. After predicting possible molecular formulas from the characteristic isotope patterns in the MS/MS spectrum, SIRIUS 4 uses CSI:FingerID to identify chemical structures that coincide with the predicted molecular formulas and MS/MS peaks. The workflow continues to be improved, with SIRIUS 5 containing new features such as lipid annotation. Other tools such as MSNovelist[204] have been developed using alternative ML models, such as recurrent neural networks, to predict chemical structures from MS/MS spectra.

ML has also been used to predict MS/MS spectra from chemical structure. For example, 3DMolMS uses a deep neural network to predict MS/MS spectra from 3D conformations of molecules.[205] 3DMolMS is advantageous over existing methods for MS/MS prediction, as these previous methods use 2D structures of molecules to make their predictions, neglecting important information contained in the 3D-structure.

Specifically designed for RiPPs, HypoRiPPAtlas combines BGC data with query MS/MS data to predict hypothetical RiPPs (Figure 18).[206] HypoRiPPAtlas integrates seq2ripp, a ML tool that is able to predict RiPP structures from putative BGCs including the sites at which PTMs can occur. This collection of structures is then individually predicted with an MS/MS spectrum, which can be searched in databases for spectral matches.

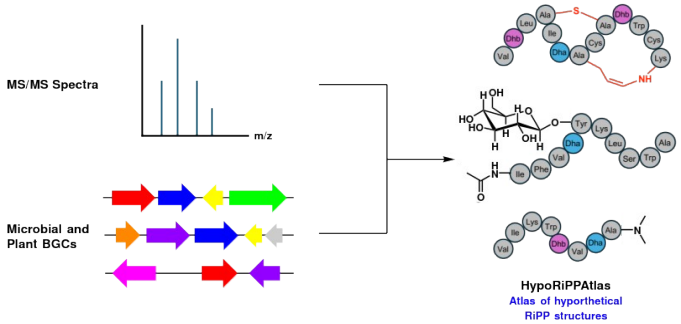


Figure 18. HypoRiPPAtlas predicts hypothetical RiPPs structures from a query spectrum, enabling matches to be identified from entries in a spectral database.

2.3.2 Nuclear Magnetic Resonance Spectroscopy (NMR)-based Approaches

Nuclear Magnetic Resonance (NMR) is an essential analytical method that allows for the complete structural elucidation of complex natural products. Algorithms to annotate and dereplicate NMR spectra are desired to expedite natural product discovery. However, impurities, solvent effects, and difficulties in accurately predicting chemical shifts have made designing such algorithms more challenging than those described for MS datasets.

2.3.2.1 One-dimensional NMR Tools

NMRFinder is a tool for 1D ^{1}H-NMR metabolite annotation that matches 1D ^{1}H-NMR samples with a reference library, and was successful in annotating 71.8% of metabolites in a tested dataset.[207] MixONat is a tool that dereplicates 1D ^{13}C-NMR spectra.[208] Using 1D ^{13}C-NMR, DEPT-135, and DEPT-90 data and molecular weight filtering, MixONat enables structurally similar natural products, including stereoisomers, to be distinguished. One limitation is the requirement for external ^{13}C chemical shift data, either predicted or from databases. Other 1D ^{13}C-NMR-based predictive models have been developed. One example is a model based on XGBoost (Extreme Gradient Boosting), a ML algorithm capable of using structured data to make accurate predictions.[209] Performance values of ~80% were observed for most classes of compounds. These values quantify the effectiveness of the selected classifier algorithms, and values of 80% indicate that the prediction model performed well for natural products outside of the dataset. "**C**onditional **M**olecular **G**eneration **net**" (CMGnet) uses various types of data, including ^{13}C-NMR spectra, molecular formulae, and molecular fragments as inputs, enabling molecular structures to be predicted.[210] CMGnet utilizes large-scale pretraining, a ML technique that involves optimizing model parameters prior to data collection, allowing it to have a successful rate of 94.7% for the top 10 structural predictions.

2.3.2.2 Multidimensional NMR Tools

The development of "**S**mall **M**olecule **A**ccurate **R**ecognition **T**echnology" (SMART) was a breakthrough in NMR-based natural product dereplication tool.[211] SMART was designed by combining 2D NMR methods and deep convolutional neural networks (CNN), trained using a database containing 2054 heteronuclear single quantum coherence (HSQC)

spectra. The resulting SMART tool enables accurate dereplication, allowing newly isolated natural products and known analogs to be rapidly characterized. In a follow-up study, SMART 2.0, was used in conjunction with MS/MS based molecular networking (MN) to analyze a *Symploca* extract. This led to the rapid discovery of the macrolide symplocolide A (**87**) and the reannotation of notable known natural products: cytotoxic swinholide A (**88**) and samholide A (**89**) (Figure 19).[212]

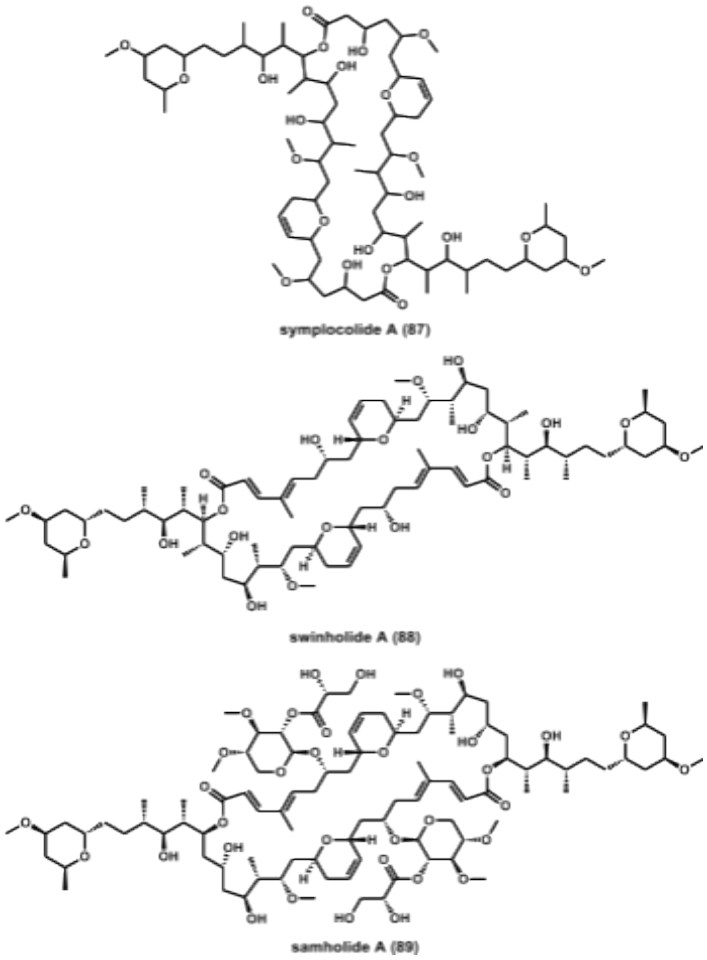


Figure 19. Cyclic peptides identified from an environmental *Symploca* extract using SMART 2.0.

More recent examples of 2D ^{1}H-^{13}C HSQC-based tools are DeepSAT and SMART-Miner. DeepSAT enables both structure annotation, as well as scaffold prediction.[213] DeepSAT extracts ^{1}H-^{13}C HSQC features and outputs known compound with similar structural motifs and/or structures, expediting NMR-based workflows. SMART-Miner uses a CNN to overcome the challenges of different peak intensities of various metabolites in a complex mixture and is able to accurately identify peaks in the ^{1}H-^{13}C HSQC spectra of complex mixtures.[214] Another example is the pseudo-Siamese convolutional

neural network (pSCNN) DeepMID, which is able to accurately identify compounds in mixtures from NMR data. DeepMID has the additional advantage of correcting for variations in chemical shift between different NMR spectra.[215]

2.3.2.3 Chemical Shift Predictions

Chemical shift variability is another challenge in NMR-based discovery tools, limiting the degree to which such tools can be automated. Recently, a linear regression model based on chemical homology was constructed from 940 serum samples, allowing for the prediction of ^{1}H-NMR chemical shifts for 15 serum and plasma metabolites.[216] However, the accuracy of the model must be improved to further automate chemical shift prediction. "**G**lobal **I**ntensity-guided **P**eak **M**atching and **A**lignment" (GIPMA) expands upon the linear regression model, enabling peak matching and alignment for 2D ^{1}H-^{13}C HSQC data. GIPMA was validated using 2D ^{1}H-^{13}C HSQC spectra for 87 *Dendrobium* extracts.[217]

ML tools have also been applied to predicting chemical shifts. Using a graph neural network (GNN), a neural network that operates on data that can be represented by a graph, 3D molecular structures are inputted and ^{1}H and ^{13}C chemical shifts are outputted.[218] Through a technique called transfer learning, a process by which a model is repurposed for another task as a means of optimizing it, the model is optimized using experimental datasets, resulting in the optimized model performing with similar accuracy as other cutting-edge prediction tools. Remarkably, the online model makes predictions in 1/6000 of the CPU time compared to other tools, underscoring its scalability.

2.3.3 Integrated Approaches for Natural Product Discovery

As the previous two subsections have highlighted, developments in MS-based and NMR-based discovery workflows have enabled both approaches to be used in isolation to discover new natural products with much-improved efficiency. Combining the tools used in annotation from MS and NMR data can produce more accurate predictions. For example, Data-Fusion-based Discovery (DAFdiscovery) is a user-friendly pipeline that enables MS, NMR, and bioactivity data to be combined by employing STOCSY and SHY (statistical total correlation spectroscopy and Statistical HetereoSpectroscopy, respectively) calculations. While STOCSY serves to process and analyze NMR spectra, the SHY component of DAFdiscovery enables multispectral analysis through cross-correlation of spectral parameters (e.g. *m/z* and chemical shift).[219,220] By using DAFdiscovery to examine multispectral correlations in already annotated spectra, the annotation of unannotated experimental spectra can be expedited. The drawback of DAFdiscovery is that it is a manual

tool, limiting its applicability in high-throughput NP discovery workflows. The “Semi-automatic correlation analysis for reliable metabolite annotation" (SCORE-metabolite-ID) approach is a similar multidimensional mathematical correlation tool.[221] SCORE-metabolite-ID enables semi-automated correlation of MS spectra with ^{1}H-NMR spectra, with the third dimension being chromatographic retention times from LC data. The approach was successfully applied to a polar pine nut extract, enabling 40 natural products to be identified, including N^{α}-(2-hydroxy-2-carboxymethylsuccinyl)- L-arginine, a metabolite that was previously unreported in food samples.[222]

3. Heterologous Expression of Natural Product Pathways

3.1 Introduction

The *in-silico* prediction of biosynthetic genes and BGCs have unlocked the true biosynthetic potential of Nature far beyond the classic isolation approaches, which are limited by native strain production under natural or laboratory settings. The biggest impact synthetic biology has made on natural product biosynthesis is undoubtedly the heterologous expression of biosynthetic pathways in model organisms. In the last thirty years, nearly every class of natural products has been produced in a handful of engineered microbial or plants hosts. These advances are enabled by modern technologies in DNA sequencing, DNA synthesis, gene editing, strain engineering, and protein engineering. Heterologous expression has enabled i) detailed analyses of biosynthetic pathways and enzymes in standardized hosts; ii) the renewable production of valuable compounds in abundant titers compared to the minute amounts from natural hosts; iii) facile engineering of pathways that leads to generation of analogs and improved variants; and iv) discovery of new natural products from cryptic biosynthetic pathways, often without the need to ever obtain the original host.

3.1.1 Drawbacks of Native Strains

In this section, native strains represent the organisms in which the natural product is first reported or where the BGCs of interest are first sequenced from. While many advantages of working with native strains are obvious, including the high likelihood of producing the final, "functional" version of the natural product, there are also many disadvantages. The biggest obstacles in working with native producers are slow growth, unculturability, and genetic intractability, traits associated with many plants and microorganisms. In addition, natural products, especially those with potent biological activities, are often produced at very low quantities for isolation from the natural host.[223,224] The regulatory mechanisms of metabolite production in natural hosts are difficult to decipher and replicate under laboratory conditions. As a result, many BGCs remain "silent" with unknown functions despite approaches such as OSMAC and others highlighted in Section 2.2.[225,226]

3.1.2 Features of Heterologous Hosts

In this next section, we will cover four main classes of heterologous hosts: bacteria, Baker's yeast, filamentous fungi, and plant host systems (Table 2). A selected number of compounds that have successfully been expressed in

heterologous hosts are shown in Figure 20. These compounds span a range of natural product classes. *E. coli* has been used for the production of polyketides such as erythromycin (**1**) and terpenes such as taxadiene (**90**), a precursor to the drug paclitaxel (**7**).[227,228] Plant hosts such as the tobacco plant relative *Nicotiana benthamiana* has been used in the production of (-)-podophyllotoxin (**11**), a secondary metabolite of the mayapple plant and a precursor to the chemotherapeutic etoposide; and the alkaloid colchicine (**91**), an anti-inflammation agent used to treat conditions such as gout.[229,230] Baker's yeast (*Saccharomyces cerevisiae*) is covered as a separate section from fungi due to the sheer volume and variety of examples of its usage as a heterologous host. Prominent examples include the opioid hydrocodone (**92**), the cardiovascular disease treatment lovastatin (**2**), the tetrahydrocannabinol and cannabidiol precursor cannabigerolic acid (**93**), artemisinic acid (**94**), precursor to the anti-malarial drug artemisinin (**8**), and the cough suppressant noscapine (**95**).[26,231–236] Finally, model filamentous fungi have been utilized to heterologously produce diverse natural products from mold and mushroom, including β-lactam antibiotic penicillin G (**96**), antibiotic precursor pleuromutilin (**97**), and psilocybin (**98**), a bioactive hallucinogenic alkaloid from the mushroom *Psilocybe cubensis* that is a candidate for treating various neurological or psychological disorders.[237–239] These model microorganisms, especially *E. coli*, *S. cerevisiae*, and *Streptomyces sp.* are suitable for metabolic engineering of high titer, rate, and yield (TRY) strains suitable for industrial applications. Metabolic engineering campaigns are conducted in an iterative design, build, test, and learn (DBTL) fashion in which genetically modified strains are tested and assessed for optimal biosynthetic TRY. A recent review by Volk et al. describes practices in metabolic engineering and their applications.[240]

Class	Example Species	Advantages	Disadvantages	Notable Heterologous Products
Bacteria	*Escherichia coli* *Streptomyces lividans* *S. coelicolor* *S. albus* *Pseudomonas spp.*	Simple, single-celled, fast growth Genetic tractability Precedence as heterologous host *P. putida* is resistant to xenobiotics	Lack of cellular compartmentalization Lack of machinery to support post-translational enzyme modifications Building block pathways not always available Prone to bacteriophage attack	erythromycin (**1**) taxadiene (**90**) shinorine (**102**)
Yeasts	*Saccharomyces cerevisiae* *Yarrowia lipolytica* *Pichia pastoris*	Simple, single-celled, fast growth Genetic tractability Robust genetic recombination for large vectors construction Precedence as heterologous host Presence of organelles such as ER for expression of specific enzymes (*P. pastoris*) methanol utilization capability (*P. pastoris*) recombinant protein expression up to 30% of total protein content[241]	Intron splicing machinery is different from other eukaryotic hosts, requiring the use of cDNA Biosynthetic pathway can often crosstalk with endogenous metabolism, particularly redox and detoxification pathways	thebaine (**200**) and hydrocodone (**92**) lovastatin (**2**) noscapine (**95**) artemisinic acid (**94**) penicillin G (**96**) tropane alkaloids (**202**, **203**) cannabinoids (**221**, **222**)
Fungi	*Aspergillus nidulans* *A. oryzae* *A. niger* *Fusarium spp.* *Penicillium spp.*	Proper intron splicing for fungal biosynthetic pathways Well-characterized cellular compartmentalization Genes can be expressed from episomal vectors and from genome integration	More difficult to manipulate genetically compared to yeast Rich endogenous secondary metabolite background Biosynthetic pathway can often crosstalk with endogenous metabolism	lovastatin (**2**) pleuromutilin (**97**)
Plants	*Nicotiana benthamiana*	Inherent suitability for expressing complex plant enzymes Cytotoxicity is less of a concern Highly amenable to *Agrobacterium*-mediated transient expression	Longer growth times Difficult to scale compared to microorganisms Working directly with plant host and metabolism can be	paclitaxel (**7**) (-)-podophyllotoxin (**11**) colchicine (**91**)

Table 2. Overview of heterologous expression hosts, advantages/disadvantages thereof, and notable examples.

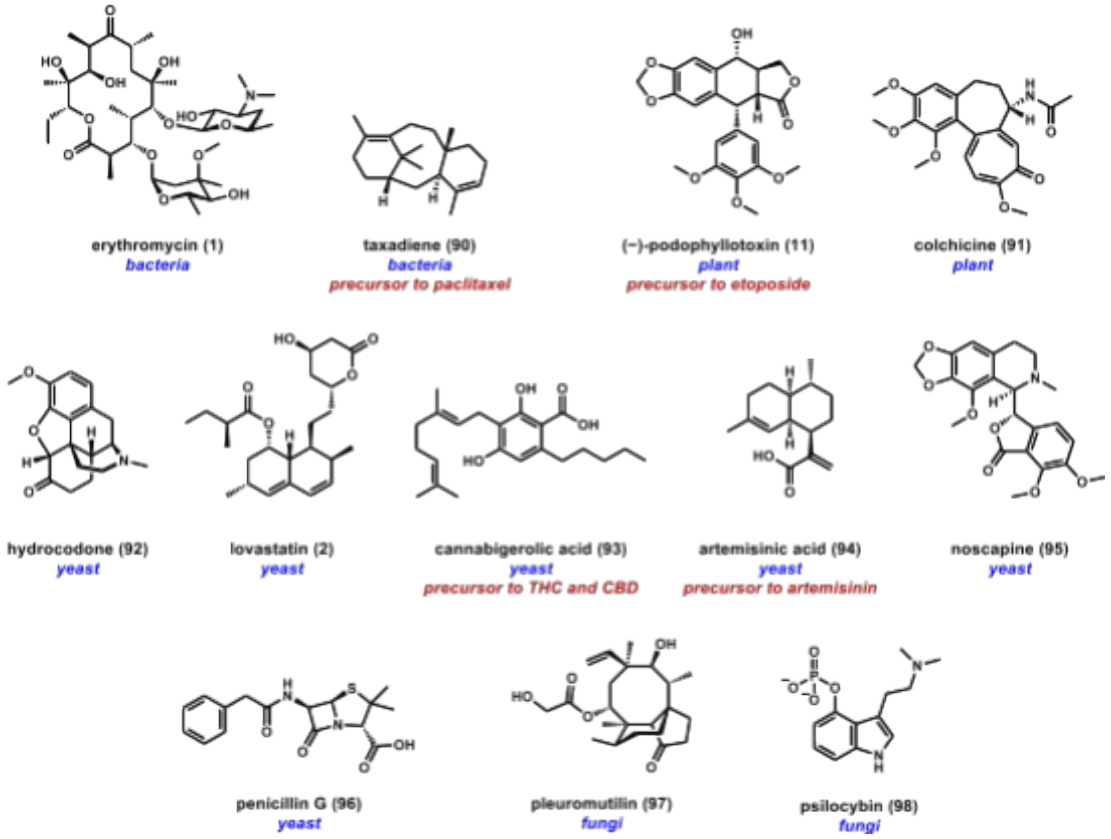


Figure 20. Selected examples of natural products that have been heterologously produced.

3.2 Heterologous Expression in Bacterial Heterologous Hosts

3.2.1 Commonly Used Bacterial Hosts

Several bacterial strains have served as workhorse heterologous hosts for BGC expression and engineering. Among them, the Gram-negative *E. coli* remains widely used due to its fast growth, routine culture conditions, and ease of genetic modifications. To make *E. coli* more suitable for expression of a wide variety of BGCs, numerous modifications have been made and adopted by the scientific community. For example, in the case of megasynthases such as PKSs or NRPSs, modification of thioester carrier protein from the *apo* state to the 4'-phosphopantetheinylated *holo* state is essential (Figure 21). To facilitate this, the phosphopantetheinyl transferase (PPTase) from *Bacillus subtilis,* Sfp, was genetically integrated into *E. coli* BL21(DE3) to afford the strain BAP1.[242] Additional modifications to the BAP1 strain include integration and

overexpression of the *prp* operon (and specifically the *prpE* gene that codes for a propionyl-CoA synthetase) to provide propionyl-CoA and methylmalonyl-CoA used by bacterial type I PKSs such as DEBS1-3 involved in erythromycin biosynthesis.[243] The *E. coli* Marionette strain has been developed to provide more tunability for expression of specific enzymes within a pathway during heterologous expression.[244] One glaring disadvantage of *E. coli* is the difficulty in expressing most fungal and plant-derived cytochrome P450 enzymes responsible for oxidative decoration of natural products. These P450s contain membrane-inserting *N*-terminal regions that anchor the enzymes to the endoplasmic reticulum (ER) in native hosts. P450 engineering, such as deletion or modification of the *N*-terminus, has been reported; [245,246] nevertheless, this bottleneck has limited *E. coli* as a host for expressing most eukaryotic pathways.

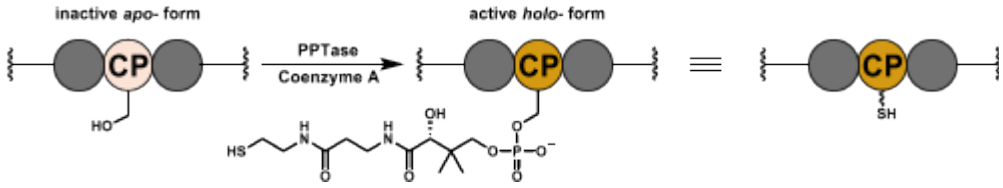


Figure 21. Activity of phosphopantetheinyl transferase (PPTase) to convert carrier protein (CP) from *apo*- to *holo*- form. A serine residue is phosphopanthetheinylated with coenzyme A in the presence of PPTase.

Several heterologous expression achievements using *E. coli* are shown in Figure 22. Using BAP1, Pfeifer and Khosla demonstrated expression of all three megasynthases (DEBS1, DEBS2 and DEBS3) from the erythromycin BGC found in *Saccharopolyspora erythraea* to produce 6-deoxyerythronolide B (**99**, 6-dEB), the macrocyclic aglycon of erythromycin (**1**).[227] Follow up studies with full expression of the remaining pathway enzymes led to production of fully decorated **1** from *E. coli.* [247] Stephanopoulos and coworkers expressed taxadiene synthase and optimized *E. coli* for production of the diterpene taxadiene (**90**), the precursor to paclitaxel (**7**).[228] Efforts to increase titer included optimizing the methylerythritol phosphate (MEP) pathway to produce the isoprenoid precursors IPP and DMAPP, which resulted in titers of **90** exceeding 1 g/L. Further expression of the first three enzymes in the paclitaxel biosynthetic pathway led to accumulation of taxadien-5α-ol (**100**) at 60 mg/L. The third example highlighted here is the biosynthesis of mycosporine glycine (**101**) and shinorine (**102**), and mycosporine-like amino acids (MAAs) from cyanobacteria in *E. coli* by Balskus and Walsh.[248] MAAs are key ingredients in the formulation of sunscreen and other skincare/cosmetic products.

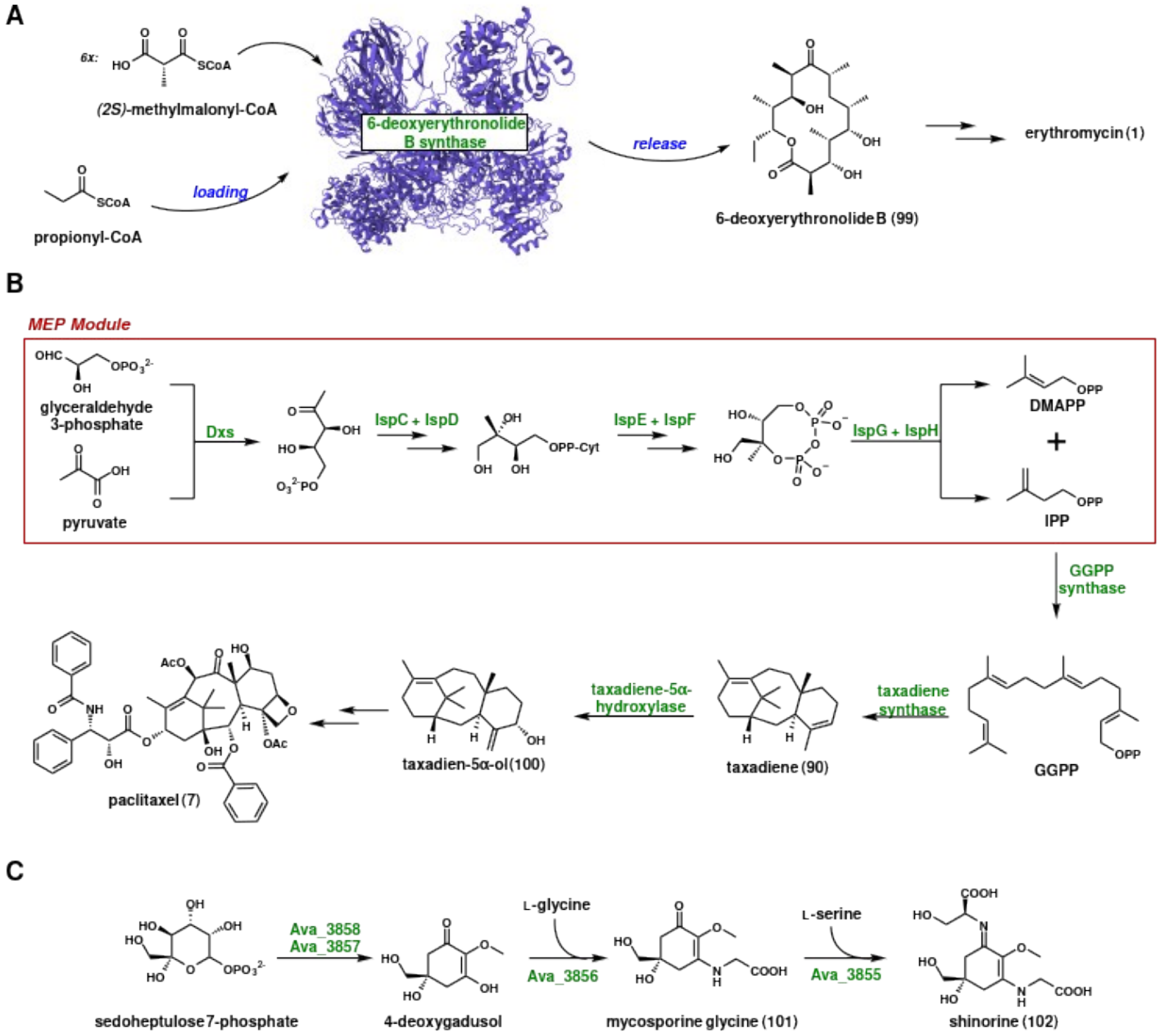
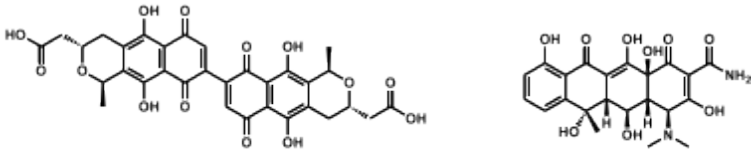


Figure 22. Examples of heterologous expression in *E. coli.* (A) **99** and **1** from the erythromycin pathway. (PDB: 7M7I); (B) taxadiene (**90**) and taxadien-5α-ol (**100**) which are precursors to paclitaxel (**7**); and (C) mycosporine-like amino acids such as shinorine (**102**).

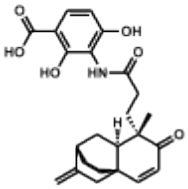
Given that a vast majority of natural products are produced by bacteria belonging to the Actinomycete genus, it is not surprising that numerous *Streptomyces* strains have been developed as heterologous hosts.[249] The two classic strains are *Streptomyces coelicolor* and *Streptomyces lividans*, followed by several other widely used strains such as *Streptomyces albus* and *Streptomyces avermitilis. Streptomyces* hosts are valued for their well-defined metabolic backgrounds and ease of

genetic manipulation. Both genome integration and episomal vectors can be used to express heterologous pathways. Extensive efforts have been invested into the optimization of *Streptomyces* heterologous hosts,[250,251] including the development of strong constitutive promoters.[252] *Streptomyces* hosts have a rich history of use in expressing natural product BGCs, with examples ranging across polyketides, NRPs, and others (Figure 23).[253–255] Sir David Hopwood first documented the heterologous expression of the actinorhodin (**103**) BGC in *S. parvulas.*[37] Khosla and Hopwood developed the *S. coelicolor* strain CH999 for the expression of both type I and type II polyketides, including the full expression of the 6-deoxyerythronolide B synthase (DEBS) for production of **99**.[256] One of the first uses of *S. lividans* for heterologous expression was the production of oxytetracycline (**104**) through expression of the *otc* BGC from *S. rimosus* using a cosmid-based strategy.[257] More recently, *S. lividans* has been employed for the production of platencin (**105**), the alkaloid guanipiperazine B (**106**), kinanthraquinone (**107**), the polyene JBIR-159 (**108**), the hybrid polyketide-RiPP goadvionin A4 (**109**), linaridin RiPPs such as grisemycin, and more.[258–263]

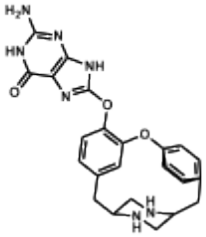


actinorhodin (103)
S. parvulas

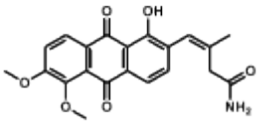
oxytetracycline (104)
S. lividans



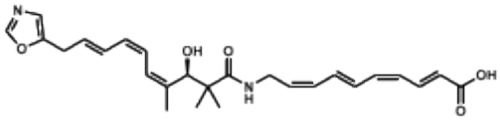
platencin (105)
S. lividans



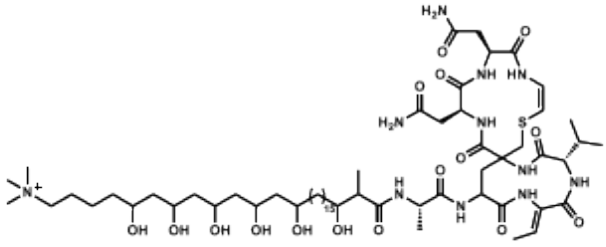
guanipiperazine B (106)
S. lividans



kinanthraquinone (107)
S. lividans



JBIR-159 (108)
S. lividans



goadvionin A4 (109)
S. lividans

Figure 23. Examples of natural products heterologously produced in *Streptomyces* hosts.

Other bacteria have also been investigated as heterologous hosts to produce a wide range of natural products (Figure 24). *Pseudomonas putida,* which has a similar safety rating as *E. coli,* is a popular Gram-negative host.[264,265] *P. putida* has been investigated as a strain for the production of bulk chemicals, as well as a diverse range of natural products including isoprenoids, polyketides, NRPs, rhamnolipids, and more.[264,266] The well-characterized primary metabolism and strong resistance to xenobiotics are key advantages of this strain.[266–268] *Bacillus subtilis* has been used as a heterologous expression host due to its natural competence, Generally-Regarded-As-Safe (GRAS) status, and well-defined genetics. *B. subtilis* contains a rich portfolio of endogenous secondary metabolites, indicating the bacterium has most of the necessary enzymatic machinery required for a heterologous host. A recent review by Put and coworkers describes heterologous expression of natural product BGCs in *B. subtilis*.[269]

Anabena (Nostoc) PCC7120 has been used as a model host for expressing BGCs from cyanobacteria because of its genetic tractability and ability to recognize promoters from other cyanobacterial strains.[270] Successful expression of BGCs for columbamide A (**110**),[271] cryptomaldamide (**111**),[272] tolypodiol,[273] etc., have been documented. Other cyanobacteria strains such as *Synechococcus elongatus* PCC 7942 have also been investigated as heterologous hosts, for example to produce the polybrominated compound 2OH-BDE68 (**112**).[274]

Corynebacterium glutamicum is a Gram-positive bacterium and is valued for its ability to natively produce methylmalonyl-CoA. This makes it an attractive option for the heterologous expression of polyketide BGCs.[275] *C. glutamicum* has also been investigated for the heterologous production of amino acid-derived natural products such as ergothioneine (**113**).[276] *Burkholderiales* strain DSM 7029, now known as *Schlegelella brevitalea* DSM 7029, has undergone gene reduction experiments in order to facilitate its use as a heterologous host.[277] Heterologous expression using *S. brevitalea* DSM 7029 led to the identification of the chitinimides including chitinimide A (**114**) from *Chitinimonas koreensis.* Furthermore, anaerobic hosts such as *Streptococcus mutans* UA159 and *Clostridium saccharoperbutylacetonicum* N1-4 have been studied for expression of BGCs identified from anaerobic microbes.[278–280] A recent review by the Cho group describes the development of non-model strains as microbial chassis for the production of natural products and commodity chemicals.[281]

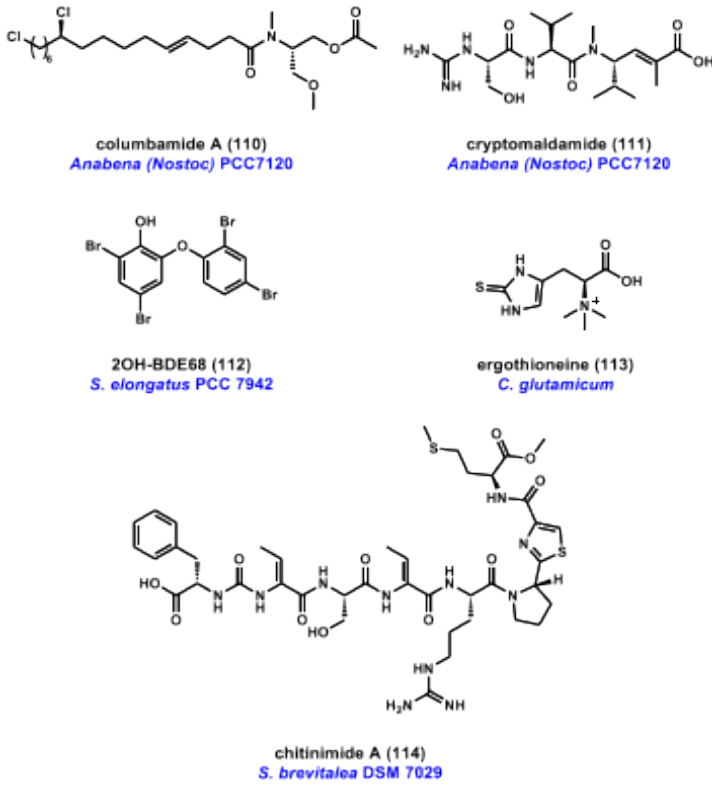


Figure 24. Examples of natural products heterologously produced in other bacterial hosts.

3.2.2 Cloning BGCs into Bacterial Heterologous Hosts

A prerequisite for heterologous expression in model bacteria is to transplant the BGCs from the original host into the host of choice. Bacterial BGCs, especially those involving multimodular PKSs and NRPSs, can exceed well over 100 kb in size. This poses significant challenges in molecular biology manipulations. These challenges are less pronounced for BGCs from fungi, as the sizes are much smaller with fewer genes. Many technologies to capture and express BGCs have emerged since the initial identification of the actinorhodin BGC. A few of them are highlighted in this section.

3.2.2.1 Cosmid Libraries

A traditional method of cloning involves the use of cosmids or bacterial artificial chromosome (BAC) libraries. Genomic DNA of native producing strains isolated from lab-grown samples or soil metagenomic DNA can be used to generate cosmid libraries, of which each cosmid can contain up to 40 kb of genomic DNA.[282] Transformants of these libraries into heterologous hosts can then be screened for the production of metabolites. Examples of recent cosmid-based heterologous expression include the lanthipeptide noursin (**115**) in *S. lividans* TK24,[283] karnamicin E_1 (**116**) in *Streptomyces albus* J1074[284] and the cyclopentane–β-lactone globilactone A (**117**) in *S. albus* J1074 (Figure 25).[285] Orphan NRPS-containing

BGCs from *Streptomyces* sp. Na03103 were heterologously expressed in *S. lividans* using cosmid libraries, which led to the discovery of two new cytotoxic cyclopeptides, ashimide A (**118**) and ashimide B (**119**).[286] Reconstitution of myxofacycline (**120**)[287] and vioprolide A (**121**)[288] biosynthesis were accomplished using a cosmid-based approach in *Myxococcus xanthus*.

Cosmid-based metagenomic libraries have been used for bioinformatics analysis to detect rare BGCs containing domains of interest. One such example is the **C**o-**O**ccurrence **N**etwor**K** **A**nalysis of **T**argeted **SEQ**uences (CONKAT-Seq), developed by Brady et al.[289] Cosmid libraries could be prepared from purified soil metagenomic DNA, followed by arraying *E. coli* transformants and identification of novel BGCs. Transformation-assisted recombination (TAR, see section 3.2.2.4) in yeast was used to assemble BGCs from across multiple cosmids. Subsequent heterologous expression in *S. albus* allowed for LC/MS analysis to identify new compounds.

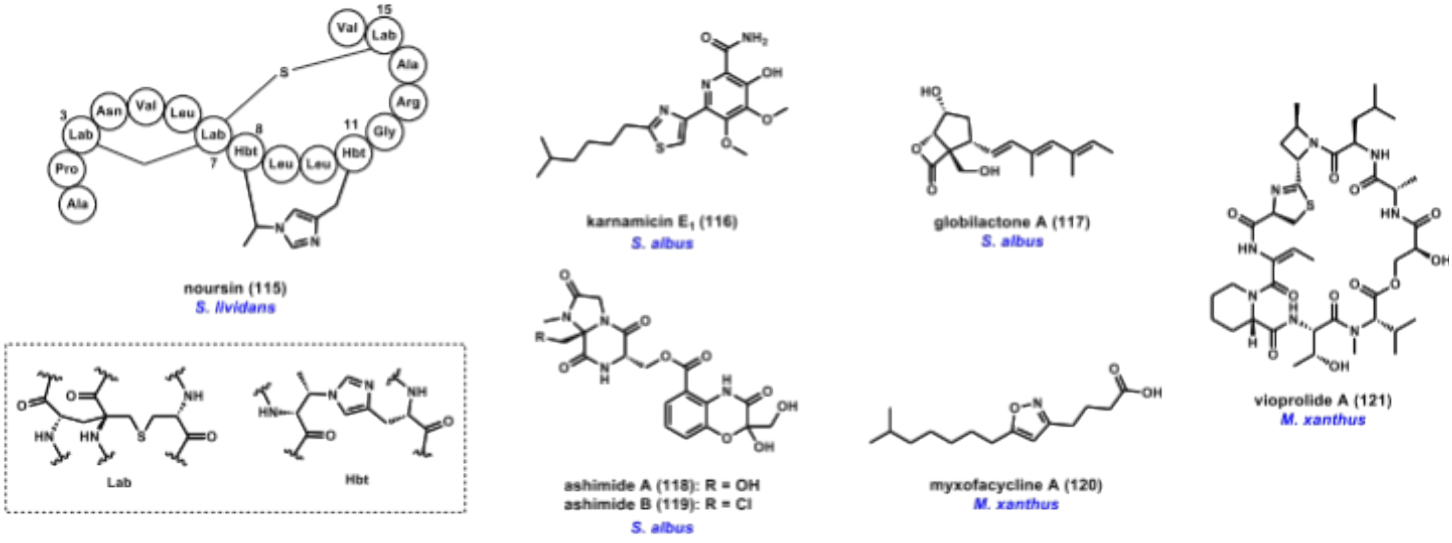


Figure 25. Examples of natural products heterologously expressed in bacterial hosts using cosmid libraries.

3.2.2.2 Bacterial Artificial Chromosomes (BACs)

Bacterial artificial chromosomes (BACs), which are based on F-plasmids, have been effective for cloning large BGCs of different natural product classes. BACs are preferred when the target BGCs exceed the capacity of cosmids (> 40 kb). BACs, while technically challenging to prepare, are particularly well-suited for expression of BGCs anchored by large, multimodular PKSs, as illustrated by the following examples (Figure 26): i) heterologous expression of quinolidomicin A_1 (**122**), which is a 60-membered macrocyclic compound from *Micromonospora* sp., in *S. lividans* at 0.1 mg/L.[290] The aglycone is biosynthesized by a type 1 PKS containing 34 modules encoded across 13 genes, while the entire BGC spans 215 kb; ii)

expression of the 77 kb *asm* BGC from *Streptomyces seoulensis* A01 in three *Streptomyces* strains: *S. lividans* SBT18, *S. coelicolor* M1146, and *S. albus* J1074.[291] From *S. lividans* and *S. coelicolor,* new metabolites ansaseomycins A (**123**) and B (**124**) were isolated; iii) cloning of the 211 kb BGC responsible for the biosynthesis of concanamycin (**125**) and expression in *S. avermitilis* SUKA32. The resulting strain produced **125** and two additional aromatic polyketides *ent*-gephyromycin and JBIR-157;[292] iv) expression of the 99 kb *lob* cluster from *Streptomyces pactum* in several *Streptomyces* hosts.[293] Upon heterologous expression in *S. lividans* TK64, *S. coelicolor* YF11, and *S. coelicolor* M1154, eleven lobophorin analogs including lobophorin H1 (**126**) were discovered. The glycosyltransferase LobG3 was identified as the key enzyme responsible for transfer of a nitrosugar, D-kijanose, to five of these eleven compounds, among which two exhibited cytotoxic activity against four human cancer cell lines; v) expression of the 50 kb *lon* BGC in *S. albus* to yield lorneic acid A (**127**);[294] vi) production of the diazo-containing compound kinamycin D (**128**) in *S. albus* J1074;[253] vii) production of oviedomycin (**129**) in the *S. coelicolor* heterologous host;[295] and viii) production of verticilatams, including verticilactam B (**130**) in *S. avermitilis* SUKA17.[296] The ferredoxin VtlF was shown to be responsible for the Diels-Alder reaction to form the substituted cyclohexene in **130**.[297]

BAC-based cloning techniques have been successfully used for heterologous expression of NRPS-anchored BGCs. The herbicidal thaxtomin A (**131**), part of a class of cyclic diketopiperazine-containing NRPs produced by *Streptomyces scabiei* and other *Streptomyces* strains, was heterologously produced in a *S. albus* J1074 host, at a titer ten times higher than that from the native host.[298] Using a BAC library, an NRPS-anchored BGC from *Streptomyces* sp. SCSIO 40020, was expressed in three *Streptomyces* hosts: *S. lividans* SBT18, *S. coelicolor* M1154, and *S. albus* J1074, and led to the production of eight thiazole-containing compounds and two quinoline derivatives, including grisechelin E (**132**).[299] Bioactivity screening showed **132** displays activity against the pathogenic bacterium *Mycobacterium tuberculosis*.

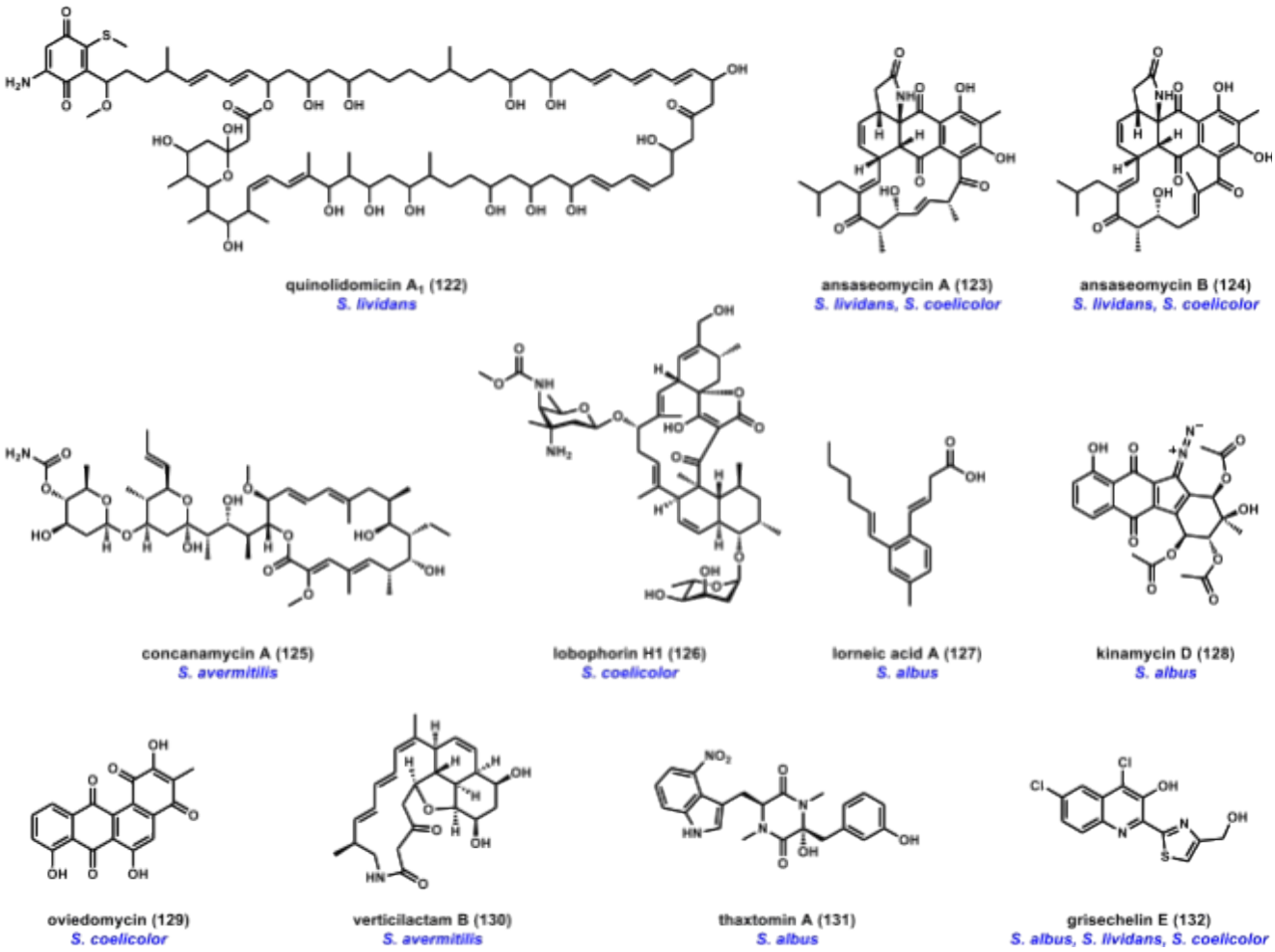


Figure 26. Examples of natural products produced in heterologous hosts using BACs.

3.2.2.3 Targeted Methods for Capturing and Cloning BGCs

Despite the continued use of cosmids and BACs in cloning BGCs, newer technologies have been developed to be more user-friendly. RNA-guided Cas9 nuclease has provided the ability to excise specific locations within bacterial genomes. The Zhu lab reported **C**as9-**A**ssisted **T**argeting of **CH**romosome segments (CATCH), a procedure in which target DNA is first cleaved *in vitro* from agarose-embedded, intact bacterial chromosomes.[300] The excised target DNA, up to 100 kb, can then be ligated through methods such as Gibson assembly into expression vectors. CATCH was used to capture and heterologously express the *cao* BGC that produces cacaoidin (**133**), a class V lanthipeptide from *Streptomyces cacaoi* CA-170360, in *S. albus* (Figure 27).[301] Additionally, using CATCH, a BGC that encodes two NRPSs was captured from *S. cacaoi* CA-170360.[302] Following transformation into *E. coli* and conjugation into *S. albus* J1074, new cyclic pentapeptides such as pentaminomycin A (**134**) were produced. CATCH was also used to heterologously express the 74.6 kb BGC spinosad BGC

from *Saccharopolyspora spinosa* ATCC49460 in *S. coelicolor* M1146 to produce spinosyn A (**135**).[303] The mechanism of sulfur incorporation into thioangucycline A (**136**) was identified through heterologous expression of the *tac* cluster from *Streptomyces* sp. CB00072 in *S. albus* using CATCH.[304]

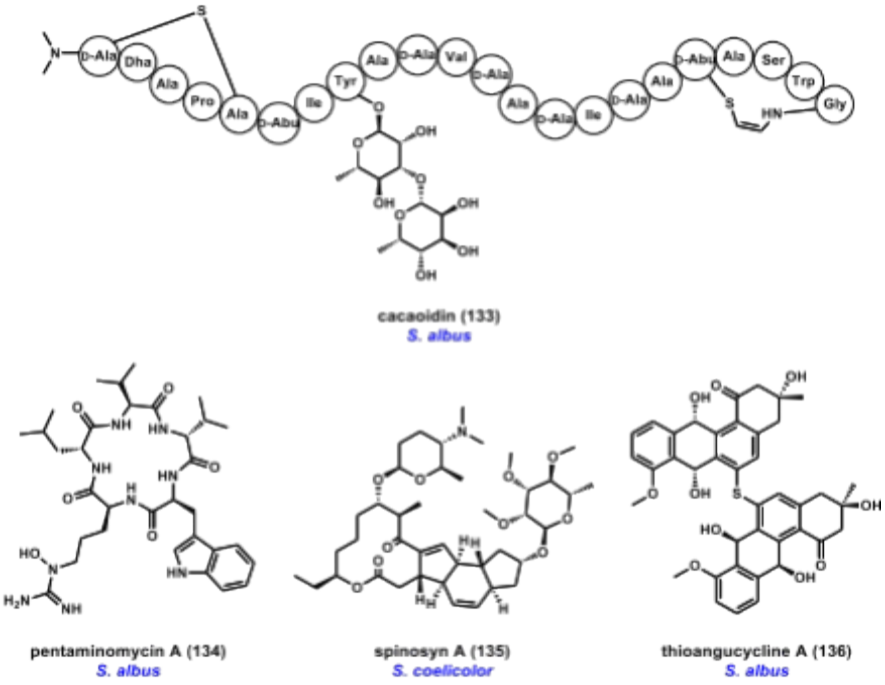


Figure 27. Examples of natural products produced in heterologous hosts after BGC capturing using CATCH.

A different cloning strategy uses Cas12a nuclease to target and excise BGCs of interest from genomic DNA. While Cas9 leaves blunt ends upon double strand breaks, Cas12a create 5'-overhangs, which can enable more facile DNA assembly in subsequent steps.[305] This strategy is termed **C**as12a-**A**ssisted **P**recise **T**argeted cloning **U**sing *in vivo* Cre-*lox* **RE**combination (CAPTURE) (Figure 28).[306] To overcome challenges in Gibson cloning as a result of high GC content of BGCs, CAPTURE relies on *in vivo* circularization of a linear vector created using T4 DNA polymerase, mediated by Cre recombinase. The resulting plasmid can then be transformed into the final heterologous host. Several natural products have been heterologously produced using this method. In the initial study, 47 BGCs of sizes up to 113 kb were cloned into either *S. avermitilis*, *S. lividans*, or *B. subtilis,* leading to the discovery of fifteen new natural products.[306] Recent applications of CAPTURE to RiPPs BGCs have also been successful, as demonstrated with that of lipoavitide (**137**).[307,308] A fragment of apratoxin A (**138**) was produced in a heterologous cyanobacterial host using a similar Cas12a-mediated method.[309] Other

methods such as CAT-FISHING, which allows for direct cloning and subsequent heterologous expression of BGCs in a *Streptomyces* chassis, have been recently described, leading to production of compounds such as marinolactam A (**139**).[310]

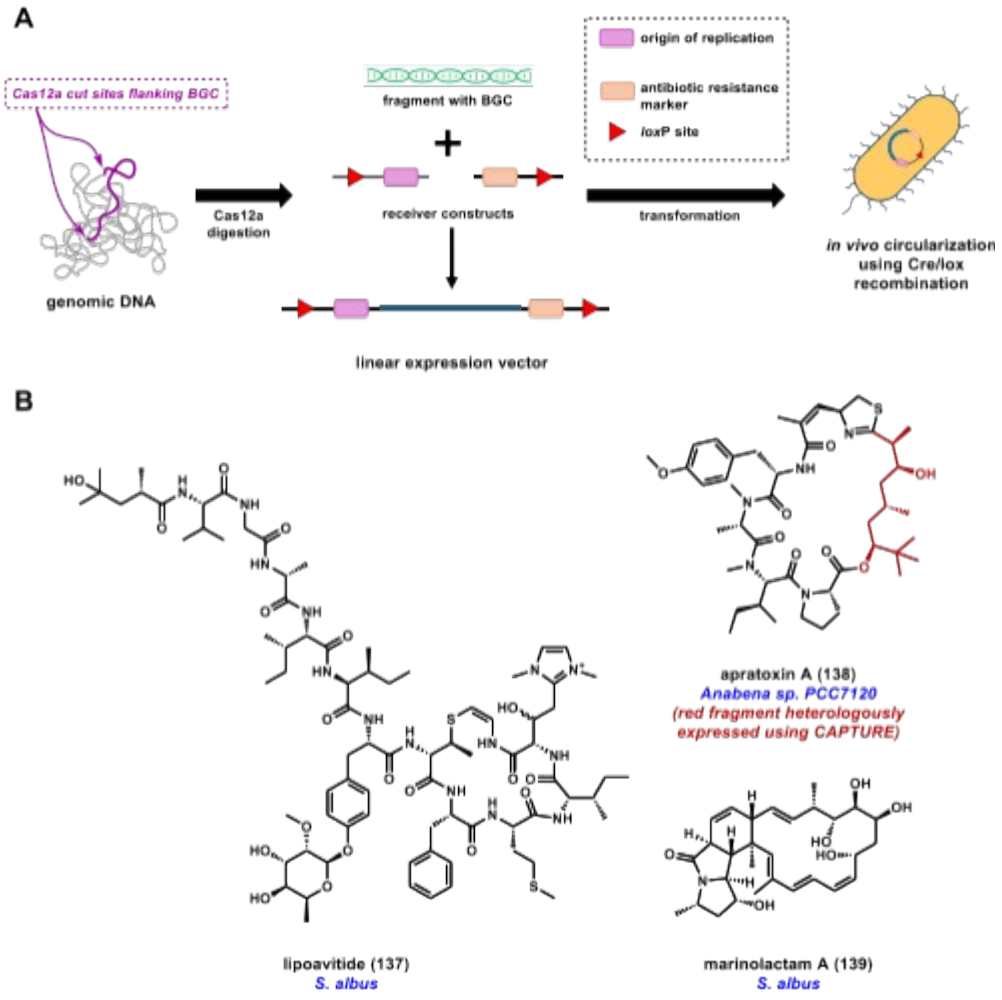


Figure 28. Examples of natural products produced in heterologous hosts after BGC capturing using CAPTURE. (A) Overview of CAPTURE workflow. (B) Compounds discussed in text.

3.2.2.4 Recombineering

Homologous recombination has been used extensively to create vectors for expressing BGCs. One such method is transformation assisted recombination (TAR).[311] Typically, this approach uses the high recombination rates seen in *S. cerevisiae* to selectively capture large DNA fragments, using a linearized TAR vector that has ends which are homologous to the termini of DNA to be captured (Figure 29).[312] For example, TAR has been used for the heterologous reconstitution of i) taromycin A (**140**) from *Saccharomonospora* sp. CNQ-490 in *S. coelicolor*; ii) cystobactamid 919-1 (**141**) in *Myxococcus xanthus* DK1622;[313] iii) pentabromopseudilin (**142**) from *Pseudoalteromonas* spp. in *E. coli;*[314] iv) thiolactomycin (**31**) from

Salinispora pacifica in *S. coelicolor*.[315] The study to elucidate thiolactomycin biosynthesis is also discussed in the context of SRE-guided genome mining in Section 2; v) alpiniamide A (**143**) from *Streptomyces* sp. IB2014/011-12 in *S. coelicolor*;[316] vi) caryoynencin (**144**) from *B. gladioli* and cepacin A (**145**) from *Burkholderia ambifaria* in various *Burkholderia* and *Paraburkholderia* hosts;[317] and vii) the RiPP kintamdin (**146**) from *Streptomyces* sp. RK44 in *Streptomyces coelicolor* M1152.[318]

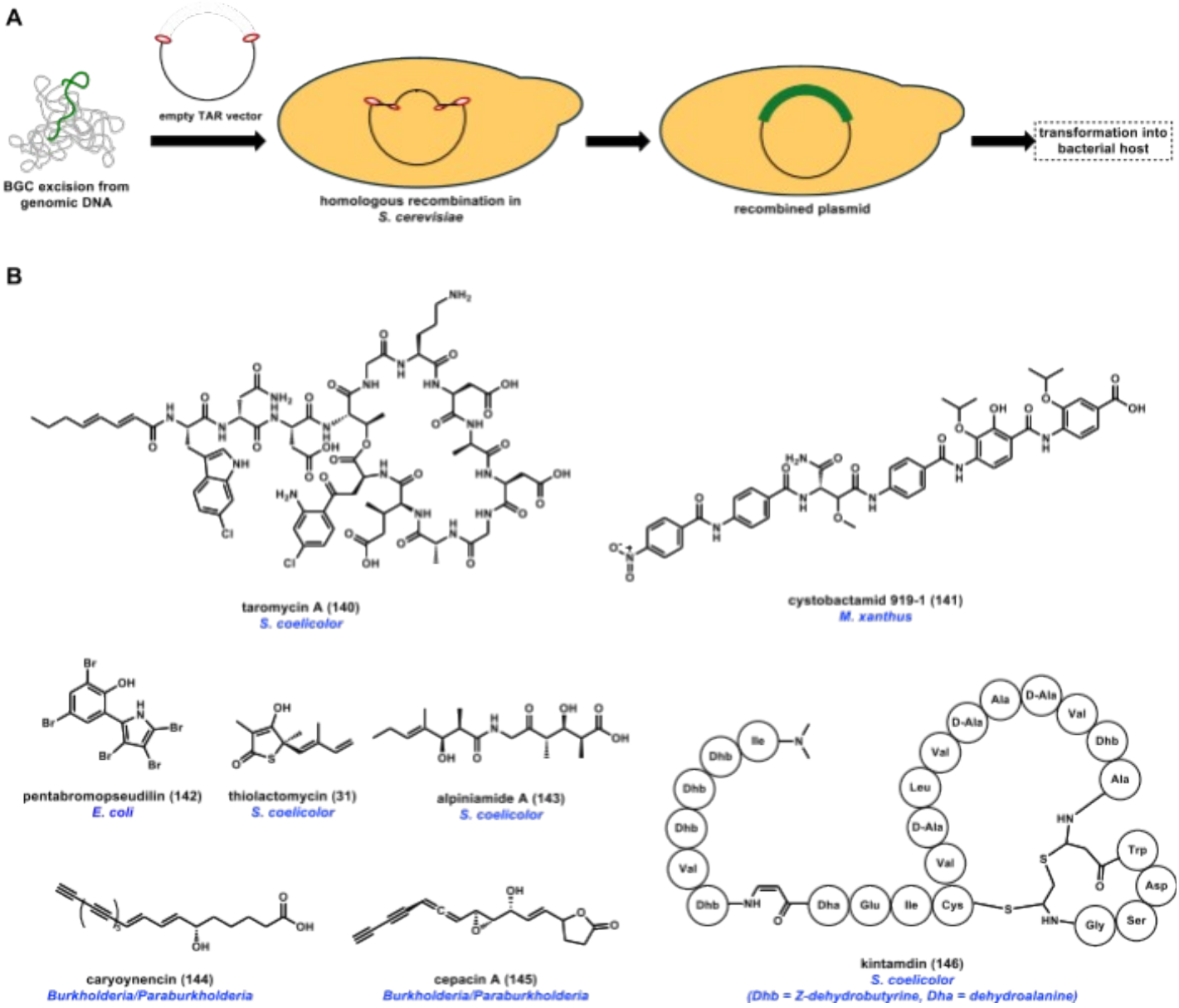


Figure 29. Examples of natural products produced in heterologous hosts after BGC capturing using TAR. (A) Overview of TAR workflow; (B) Compounds discussed in text.

Recombineering is a cloning strategy during which direct DNA insertion into targeted region is achieved through homologous recombination. Homologous recombination in *E. coli* is mediated by recombinases, namely RecA and the

RecBCD complex. These recombinases use short homology arms which can be designed to allow recombination between chromosomal DNA encoding a BGC and exogenous circular plasmid DNA. Recombineering uses phage-derived proteins, in addition to endogenous *E. coli* enzymes, to improve homologous recombination. Specifically, the exonucleases RecE/Redα and DNA-annealing proteins RecT/Redβ can be combined (RecE/T or Redα/β) to provide such functionality.[319,320] Recombinase-mediated homologous recombination can be categorized into linear-linear homologous recombination (LLHR) or linear-circular homologous recombination (LCHR). RecE/T recombination has been shown to favor LLHR.[319,321,322] Direct cloning of BGCs using RecE/T-mediated LLHR led to the discovery of luminmycin A (**147**) and luminmide A (**148**) (Figure 30).[319]

Advances in recombineering technologies for BGC cloning include ExoCET, which combines RecE/T recombineering with the *in vitro* exonuclease activity of T4 polymerase. The *in vivo* homologous recombination step can be a limiting factor for conventional RecE/T recombineering. ExoCET allows for *in vitro* annealing followed by transformation into *E. coli* as a method to improve efficiency of recombineering-based vector construction, especially for larger DNA fragments. Cloning efficiency can also be increased by targeted BGC excision using restriction enzymes or CRISPR-Cas9-targeted cutting instead of untargeted genomic DNA digestion.[323] An 80 kb cryptic PKS-containing BGC from *Streptomyces* sp. B59 was expressed in *S. albus* J1074 using ExoCET-mediated direct cloning methods, leading to the discovery of the polycyclic shuangdaolides including shuangdaolide A (**149**).[324] The technique was also used to heterologously express the spinosyn A (**135**) BGC in *S. albus* J1074, resulting in a 328-fold increase production of **135** over the native host.[325]

Other than in *Streptomyces,* the RecE/T system was also shown to be compatible for expression of BGCs > 100 kb in *P. putida,* as demonstrated in the heterologous production of violacein (**150**).[266] ExoCET was used to create heterologous expression vectors for the rhizomide and holrhizin BGCs for expression in the heterologous host *Schlegelella brevitalea* DSM 7029. Because the RecE exonuclease is known to facilitate LLHR,[319] a recombineering strategy was also used in this same study for domain swapping of two NRPS starter condensation domains (see Section 4.2.5.3). [326]

Cre/lox recombination was used to integrate a donor plasmid containing a BGC under the control of the T7 promoter into the chromosomes of several different hosts. This system is known as **C**hassis-independent **R**ecombinase-**A**ssisted **G**enome **E**ngineering (CRAGE).[163] Proof-of-concept work involved expression of PKS- and NRPS-anchored BGCs

from *P. luminescens* into 25 γ-proteobacteria species, yielding 22 compounds. A follow-up study employed CRISPR-mediated gene knockout to rapidly characterize the BGC in diverse heterologous hosts, including those lacking efficient CRISPR-based approaches.[326] Another study developed a computational method to redesign genes and regulatory elements for cross-species and cross-kingdom expression. Such work led to the discovery of the tyrocitabines, including tyrocitabine-626 (**151**) from *Lactobacillus iners*.[327] Other studies such as MetClo and SIGE have also shown promise to aid heterologous expression experiments in the future.[328,329]

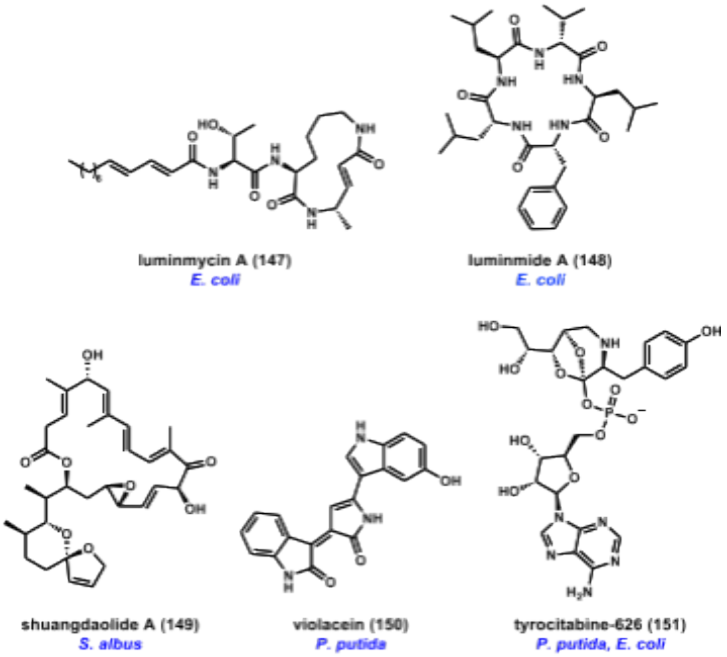


Figure 30. Examples of natural products heterologously expressed using recombineering-based strategies.

3.2.3 Selected Examples of Heterologous BGC Expression in Bacteria

3.2.3.1 Polyketide Biosynthesis

Heterologous expression in *E. coli* enabled the characterization of a nocardiosis-associated polyketide (**152**) (Figure 31), a glycolipid produced by *Nocardia* strains capable of infecting human hosts, along with *in vitro* characterization of the associated PKS (NOCAP synthase).[330,331] The biosynthesis of this compound includes tailoring of the nascent polyketide product with two sugar moieties, 3-α-epimycarose and 2-*O*-methyl-α-rhamnose, and involves "flipping" of a pathway intermediate across the cell membrane from the cytosol before completion of biosynthesis and release into the environment. A flavin monooxygenase from the pathway, NocapM, was shown to catalyze hydroxylation of a wide range of

aromatic substrates in a follow-up study.[332]. The arsenic-containing bisenarsan (**153**) produced by a type I PKS-anchored BGC was reconstituted in *S. albus* J1074.[333] The enzyme BsnM catalyzes the reaction between As(V) and 5-enolpyruvylshikimate-3-phosphate, an intermediate in the shikimate pathway, to form arsenoenolpyruvate. This molecule could then undergo an intramolecular rearrangement catalyzed by BsnN to form the C–As bond.

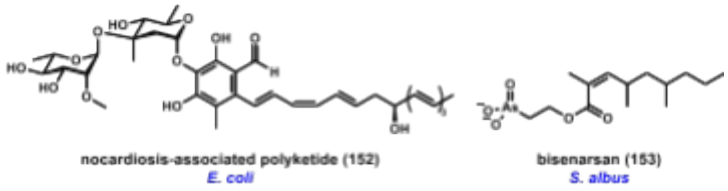


Figure 31. Examples of polyketides produced in bacterial heterologous hosts.

Enediyne natural products, which are polyketides possessing the 1,5-diyne-3-ene core, are prized for their exceptional DNA-damaging cytotoxicity. Thousands of enediyne-containing natural product BGCs have been recently predicted through bioinformatics methods.[334] Enediynes can be classified as either ten-membered, such as tiancimycin A (**154**), or nine-membered, such as C-1027 (**155**) (Figure 32). Cohen and Townsend identified a Type I PKS responsible for the formation of both the enediyne and anthraquinone moieties in dynemicin biosynthesis.[335] Bhardwaj, van Lanen, and coworkers identified polyene **156** as the PKS product and showed that two equivalents of **156** were used to form the enediyne and anthraquinone moieties in **154**.[336] Recent work in both native and heterologous hosts by Shen and coworkers led to the identification of a diiodotetrayne compound (**157**) that is the immediate downstream product of the polyene **156**.[334] Although the enzymatic basis for iodination remains mysterious, **157** was demonstrated to be a universal precursor for enediyne biosynthesis.

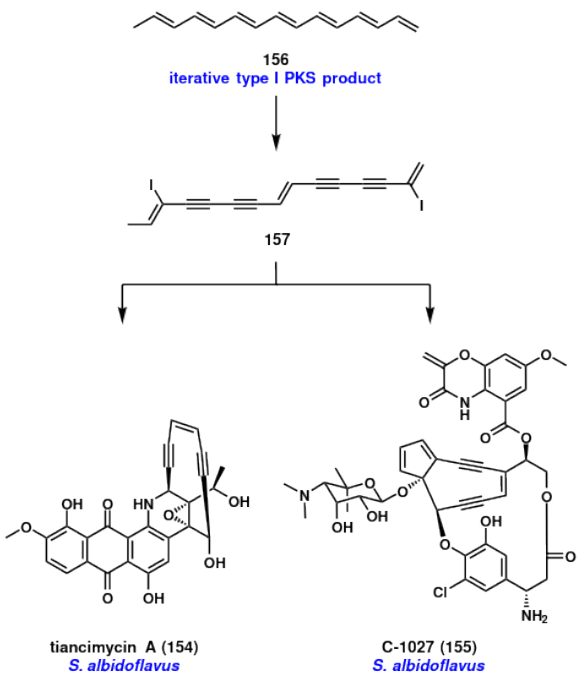


Figure 32. A diiodotetrayne **157** is a universal precursor to enediyne natural products such as **154** and **155**.

Supply of acyl-CoA precursors for polyketide biosynthesis, such as acetyl-CoA, malonyl-CoA, methylmalonyl-CoA, and propionyl-CoA, in heterologous hosts has been extensively optimized. Strategies involve overexpression of pathways leading to formation of acetyl-CoA, deletion of pathways that compete for acetyl-CoA, deletion of propionate catabolic pathways, or direct feeding of malonate followed by overexpressing malonyl-CoA.[337] One abundant source of intracellular acetyl-CoA is through the β-oxidation of fatty acids made from breakdown of triacylglycerols (TAGs) during stationary phase growth. A recent strategy, named **D**ynamic **D**egradation of **TAG** (ddTAG), was developed to channel carbon fluxes from TAG degradation to polyketide biosynthesis, guided by genome flux modeling (Figure 33).[338] The approach involves placing a fatty acyl-CoA synthetase gene under a cumate-inducible promoter, which enabled regulated expression of this key enzyme in lipid catabolism. The ddTAG approach was used to increase titer of spinosyn A (**135**) to 2 mg/L in *S. coelicolor* M1146, representing a 347-fold increase over the native producer.[303]

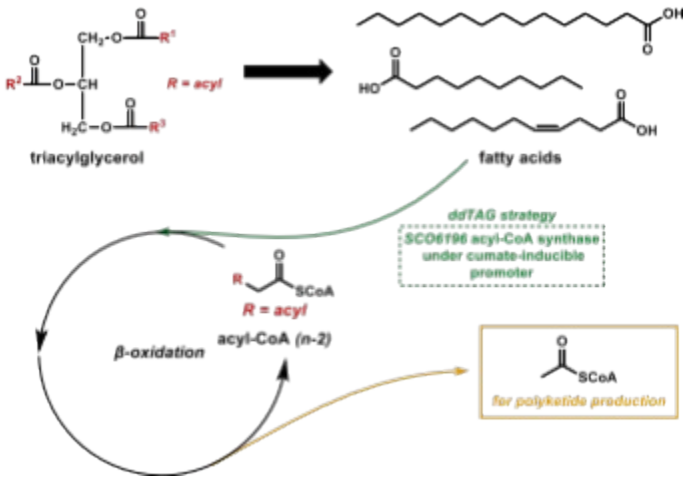


Figure 33. Increase polyketide precursor acetyl-CoA through controlled breakdown of triacylglycerols.

3.2.3.2 NRP Biosynthesis

Examples of heterologous biosynthesis of NRPs include production of chalkophomycin (**158**) which contains a *N*-hydroxypyrrole moiety, and the production of the siderophore sorangibactin A (**159**) (Figure 34).[339,340] Reconstitution of the biosynthesis of the diazo-containing azaserine (**160**), was also achieved by heterologous expression.[341] The NRPS AzaD is responsible for linking hydrazonoacetic acid or its oxidation product, which contains a diazo group, with L-serine. The biosynthesis of isonitrile-containing compounds such as **161** was investigated using heterologous expression.[342] Homologs to the BGC of **161** can be found in pathogenic mycobacteria, and the authors proposed that such isonitrile natural products play a role in metal transport during pathogenesis. Heterologous expression of the *xhp* BGC in *E. coli* was used to elucidate the biosynthetic pathway of pyrrolizwilline (**162**), which includes a two-module NRPS XhpA that catalyzes the condensation of L-serine and L-proline, and the Baeyer-Villiger monooxygenase XhpB which catalyzes ring contraction to yield a pyrrolizidine core.[343] XhpG-catalyzed deacetylation and XhpD-catalyzed reduction are followed by an addition of glyoxylic acid, a second addition of a pyrrolizidine intermediate, and XhpF-catalyzed condensation with L-Val to yield the final pentacyclic **162**.

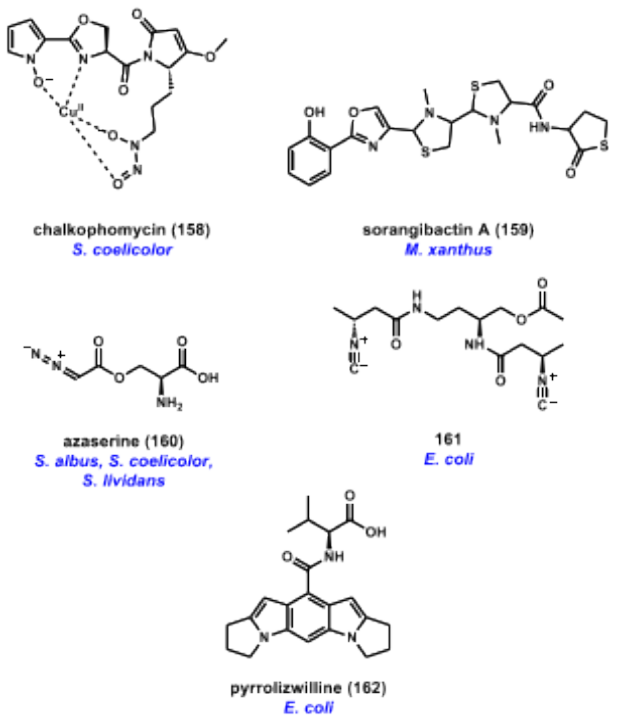


Figure 34. Examples of nonribosomal peptides produced in bacterial heterologous hosts.

3.2.3.3 RiPP Biosynthesis

Heterologous co-expression of a precursor peptide with combinations of modification enzymes in a model host allows for characterization of RiPP BGCs. Examples include imiditides such as amycolimiditide (**163**) and fuscimiditide (**164**), cloacaenodin, myxococin, thatisin, blauticin, mSmoA, SapT, and more (Figure 35).[344–356] Recent work by Zhao and coworkers developed a higher-throughput method in *E. coli* to identify RiPPs with antimicrobial properties.[357]

One class of RiPP PTM enzymes is the multinuclear iron oxygenase (MNIO) previously annotated to be part of the DUF692 family. RiPPs derived from the action of this enzyme family include methanobactin (**165**), ChrH, and pearlin.[358–360] Heterologous co-expression has been instrumental in elucidation of the role of MNIOs. One example is the characterization of the *mov* cluster from *Vibrio fluvialis*, which contains two MNIO enzymes.[361] The *mov* MNIOs were shown to be involved in installation of an oxazolone-thioamide moiety and cleavage of N-C_α of an Asn residue located near the C-terminus. *Burkholderia* sp. FERM BP-3421 has also been investigated as a heterologous host for a MNIO-containing RiPP BGC from *B. thailandensis* E264.[362]

A major class of RiPPs PTM enzymes that has been studied through heterologous expression is the radical SAM enzymes.[353] For example, two radical SAM enzymes in the *grc* cluster from *Streptococcus pneumoniae* catalyze epimerization of an L-threonine β-carbon and desaturation of a histidine residue during the biosynthesis of **166**.[353] A radical SAM enzyme ItfD from *Bacteroides thetaiotaomicron* was shown to be responsible for catalyzing C-C bond formation between a tryptophan and isoleucine residue.[352] Furthermore, in the biosynthesis of streptosactin, the radical SAM enzyme GggB was shown to catalyze C-S bond formation to yield two α-thioether moieties, creating two eight-membered rings in the process.[363] A single radical-SAM enzyme DarE is involved in the biosynthesis of darobactin (**167**) from *Photorhabdus khanii* DSM3369.[364–366] **167** has attracted significant attention due to its activity against Gram-negative pathogens through the inhibition of the BamA chaperone protein, which is a cell-surface outer-membrane protein that catalyzes folding of other proteins and their insertion into the outer cell membrane.[364] Heterologous expression in *E. coli* was key in sourcing **167**, as well as elucidating the key biosynthetic steps. Five genes from the *dar* BGC, *darA* – *darE*, were cloned into a plasmid and intergenic regions were removed. Transformation of this plasmid enabled expression of the *dar* cluster under arabinose-inducible control in *E. coli* BW25113.[364] Expression in *E. coli* led to shorter batch fermentation times over the native producer *Photorhabdus khanii* DSM3369. An improved titer was observed following use of the *E. coli* Rosetta strain and by integrating an additional copy of *darA* that encodes the precursor peptide.[367] Biochemical studies of DarE showed it is solely responsible for catalyzing the two intrapeptide crosslinking reactions that afford the unique structure of **167**.[365]

One important class of RiPP PTM enzyme is the YcaO family of enzymes, which catalyze the nucleophilic attack on backbone amide carbonyl to generate a phosphorylated intermediate.[368] Subsequent cyclization via release of the phosphate can yield structures such as oxazolines, thiazolines, or lactamidines. The Nair group recently characterized a YcaO family enzyme involved in the formation of a pyrimidone moiety from nucleophilic attack of the side-chain amide nitrogen from Asn25 on the backbone carbonyl of the previous amino acid, Ser24, in the RiPP precursor peptide. This enzyme, named SpyD, acts as part of a heteromeric complex with the RRE SpyC and a dehydrogenase SpyB.

Another interesting example of bacterial RiPP modification is the formation of a steroid-like heptophan moiety, as exemplified by the *nct* BGC from the cyanobacteria *Nostoc calcicola* FACHB-389.[369] A Trp residue in the peptide residue is modified by the addition of geranylgeranyl pyrophosphate, which cyclizes to form the heptophan moiety seen in the final

compound steromaze. The *nct* BGC was heterologously reconstituted in *E. coli* using the co-expression approach used for RiPP biosynthetic pathway elucidation.

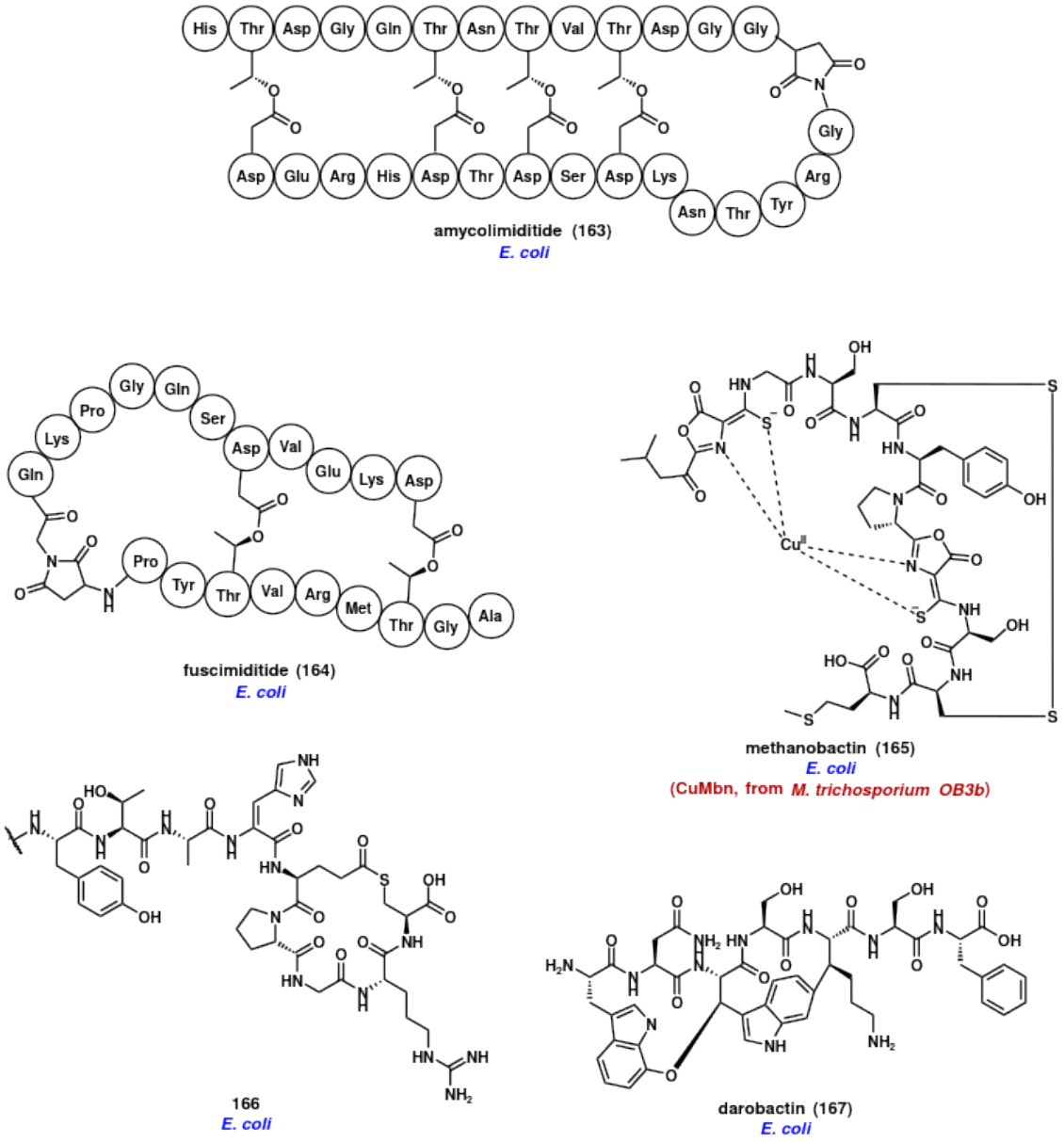


Figure 35. Examples of RiPPs produced in *E. coli* that led to elucidation of key PTM enzymes.

During the biosynthesis of some amino acid-derived metabolites or antimetabolites, a proteinogenic amino acid can be attached to the *C*-terminus of a RiPP-like carrier peptide via an intriguing PTM.[360] The amino acid is enzymatically transferred from an aminoacyl-tRNA to the carrier peptide and can undergo dramatic modifications (Figure 36). Release of the modified amino acid product repeats the catalytic cycle. Enzymes that catalyze this ATP- and tRNA-dependent modification have been coined **PE**ptide **A**minoacyl-t**R**NA **L**igases (PEARLs). The van der Donk lab studied the biosynthesis of

these compounds, including ammosamide C (**168**) and 3-thiaglutamate (**169**), using heterologous co-expression in *E. coli*.[360,370,371] A follow-up study tracked the maturation of cysteine into 3-thiahomoleucine (**170**) in another PEARL-anchored BGC.[372] PEARL BGCs are considered separate from RiPP BGCs because while the initial carrier peptide is ribosomally synthesized, the tRNA-dependent amino acid addition is non-ribosomal and the peptide product is only serving as a scaffold for repeated amino acid modifications.

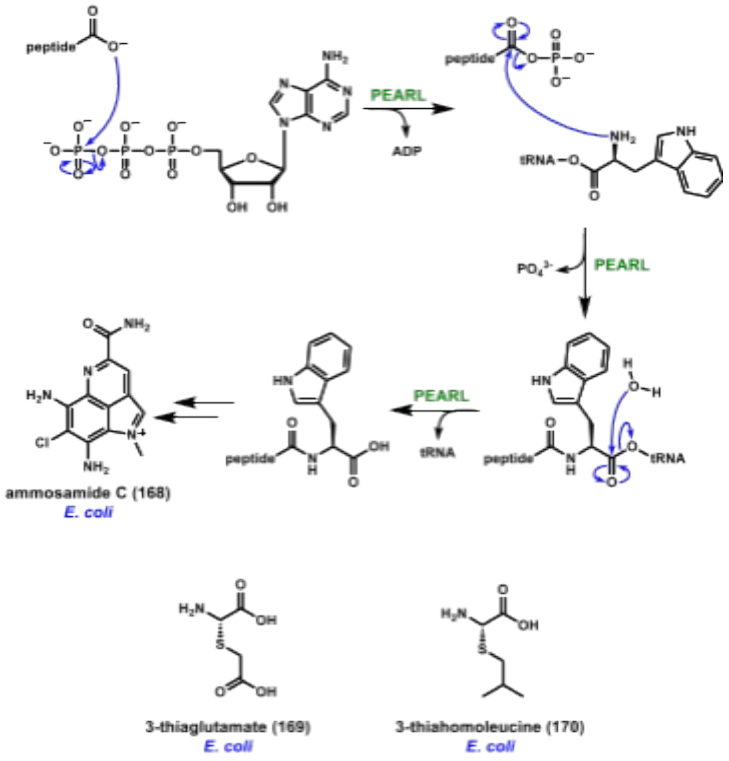


Figure 36. Biosynthesis of amino acid derived metabolites using PEARL enzymes.

3.2.3.4 Terpene Biosynthesis

The two main isoprenoid precursors, IPP and DMAPP, are produced in eukaryotes and prokaryotes by either the mevalonate (MEV, also see Figure 41) or methylerythritol phosphate (MEP) pathways, respectively. Because these pathways are related to primary metabolism and require cofactors such as NADH, NADPH, Coenzyme A, and more, many strategies have been developed towards overproduction of IPP and DMAPP in order to achieve higher levels of isoprenoid production, as summarized in a review by Stephanopoulos.[373] The Keasling lab demonstrated the MEV pathway can be introduced in *E. coli* to achieve high flux to IPP and DMAPP, and demonstrated the biosynthesis of the sesquiterpene amorphadiene, the hydrocarbon precursor to artemisinin (**8**).[374] Similarly, expression of exogenous MEV pathway genes in

E. coli directly led to a 5-fold increase in the production of (-)-patchoulol, a sesquiterpene with important uses in perfumes.[375] (*R*)-(+)-perillyl alcohol was produced in *E. coli* using a neryl pyrophosphate synthase and NADPH regeneration machinery to increase MEV pathway activity.[376] To improve production of the sesquiterpene α-santalene in *E. coli,* the MEV pathway was introduced heterologously and an optimal farnesyl pyrophosphate synthase (FPPS) was identified, leading to a 169-fold increase in titer of α-santalene.[377]

An emerging alternative method is to bypass the primary precursor supply pathways through overexpressing two kinases (isopentenyl phosphate kinases, IPKs) that can phosphorylate the simple five-carbon alcohols prenol and isoprenol into the corresponding pyrophosphates.[378,379] An IPP isomerase is typically coexpressed to control relative levels of IPP vs. DMAPP. The two kinases and isomerase can be expressed under an inducible promoter and a terpene BGC can be introduced into the heterologous host. This pathway is known as the **I**sopentenol **U**tilization **P**athway (IUP) or the alcohol-dependent hemiterpene pathway (Figure 37). Application of this approach towards terpene biosynthesis was demonstrated in the production of β-carotene (**171**), (*R*)-(-)-linalool (**172**), and lycopene (**173**) from *E. coli* with titer reaching at least 50 mg/L for each compound.[380] Geranate was produced at 764 mg/L in a 24-hour batch time from feeding 2 g/L (iso)prenol.[381] An engineered *E. coli* was transformed by the Dong group with three enzymes that can generate geranylgeranyl pyrophosphate (GGPP) starting with isoprenol and prenol through the IUP, along with taxadiene synthase, to yield taxadiene (**90**).[382] A P450 enzyme was then introduced to generate three oxidized taxanes. The substrate flexibility of IPKs has been explored to generate non-native alkyl diphosphates, with one study showed the IPK-mediated formation of 38 non-natural alkyl diphosphates *in vitro*.[383]

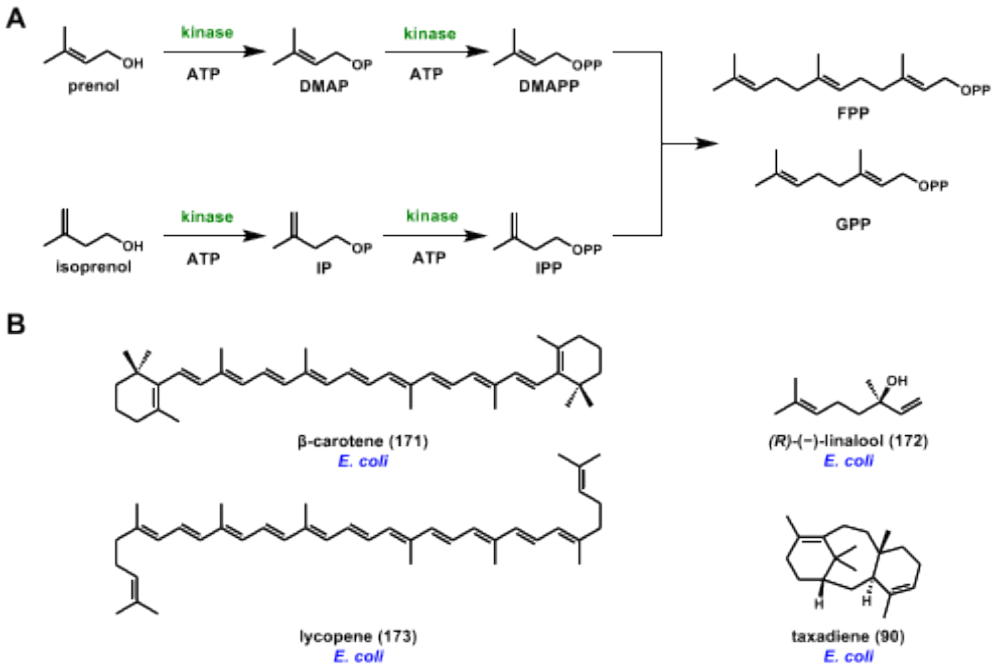


Figure 37. Phosphorylation-based two-step IPP/DMAPP biosynthetic pathway for isoprenoid production. (A) This approach starts with feeding of prenol and isoprenol alcohols; (B) Representative compounds that have been produced in bacteria using this approach.

Terpenoid BGCs frequently encode P450 enzymes that are difficult to express in *E. coli* due to redox partner availability and requirements for various cofactors. Strategies for expression of eukaryotic P450s have been developed, as discussed in a recent review by Hausjell et al.[384] One commonly used technique is the implementation of synthetic protein scaffolds, as demonstrated by Deuber, Prather, Keasling, and coworkers.[385] This method has led to improved P450 functions in bacteria by bringing the P450 monooxygenase and associated P450 reductase partner in close proximity (Figure 38). This can be accomplished by the use of scaffolding proteins to tether the oxidase and reductase together, or by direct genetic fusion of the two proteins.[386] Improved production of terpenoid products such as lutein (**174**) and (+)-nootkatone (**175**), as well as nonterpenoid compounds such as L-DOPA and apigenin (**176**) have been achieved in *E. coli* using this strategy. Overexpression of four genes involved in heme biosynthesis also increased titers in some cases.[387] A similar scaffolding approach to tether the P450 CYP82D26, a P450 reductase, and an $NADP^+$-dependent aldehyde reductase led to a 9-fold increase in daidzein (**177**) titer in *E. coli*.[388]

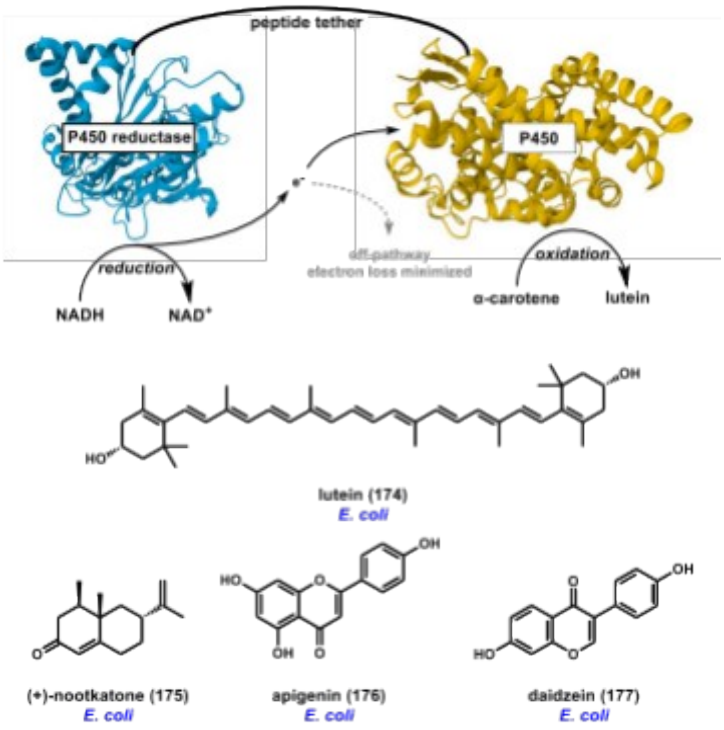


Figure 38. Spatial engineering to improve P450 functions in *E. coli.* (PDB: 7SUX (blue) and 3EJB (yellow)).

3.2.3.5 BGCs without Core Enzymes

As noted in the introduction (Section 1.2), many natural products are synthesized from BGCs that do not have a readily identifiable core enzyme. There are accumulating examples of such BGCs from bacteria and fungi. Heterologous expression has played an important role in identifying the biosynthetic pathways of these compounds, using iterative expression of different combinations of biosynthetic enzymes. One notable example is the identification of a PLP-dependent enzyme SbzP as the core-generating enzyme in the biosynthesis of altemicidin (**19**) (Figure 39A). The enzymatic basis and the building blocks of the characteristic 6-azatetrahydroindane moiety found in **19** and related natural products were a mystery and could not be deciphered using genetic knockout studies in the native host. Expression of each unassigned gene in the *sbz* BGC, in *S. lividans* TK21, was deployed as a functional assay.[34] Expression of SbzP alone led to formation of a product with the 6-azatetrahydroindane core. Biochemical characterization of SbzP showed it catalyzes an unusual net (3+2) annulation between two primary cofactors in the cell, β-NAD^+ and *S*-adenosylmethionine (SAM) to forge the bicyclic structure. Heterologous expression also led to functional characterization of eight SbzP homologs encoded in various BGCs.

Another example is the biosynthetic characterization of fluopsin C (**20**), a copper-containing antibiotic produced by the five-enzyme *flc* BGC from the bacterium *Pseudomonas aeruginosa* (Figure 39B).[35] Heterologous expression in *P. fluorescens* SBW25 afforded production of **20**. *P. fluorescens* was chosen as a heterologous host due to the lack of homologous genes in its genome, minimizing the risk of interference from endogenous proteins. The pathway involves oxidative modification of L-cysteine to *N*-hydroxymethanethioamide using two nonheme diiron oxygenases. *P. fluorescens* was also the heterologous host of choice in the elucidation of the 4-formylaminooxyvinylglycine (**178**) BGC by the same group (Figure 40).[389]

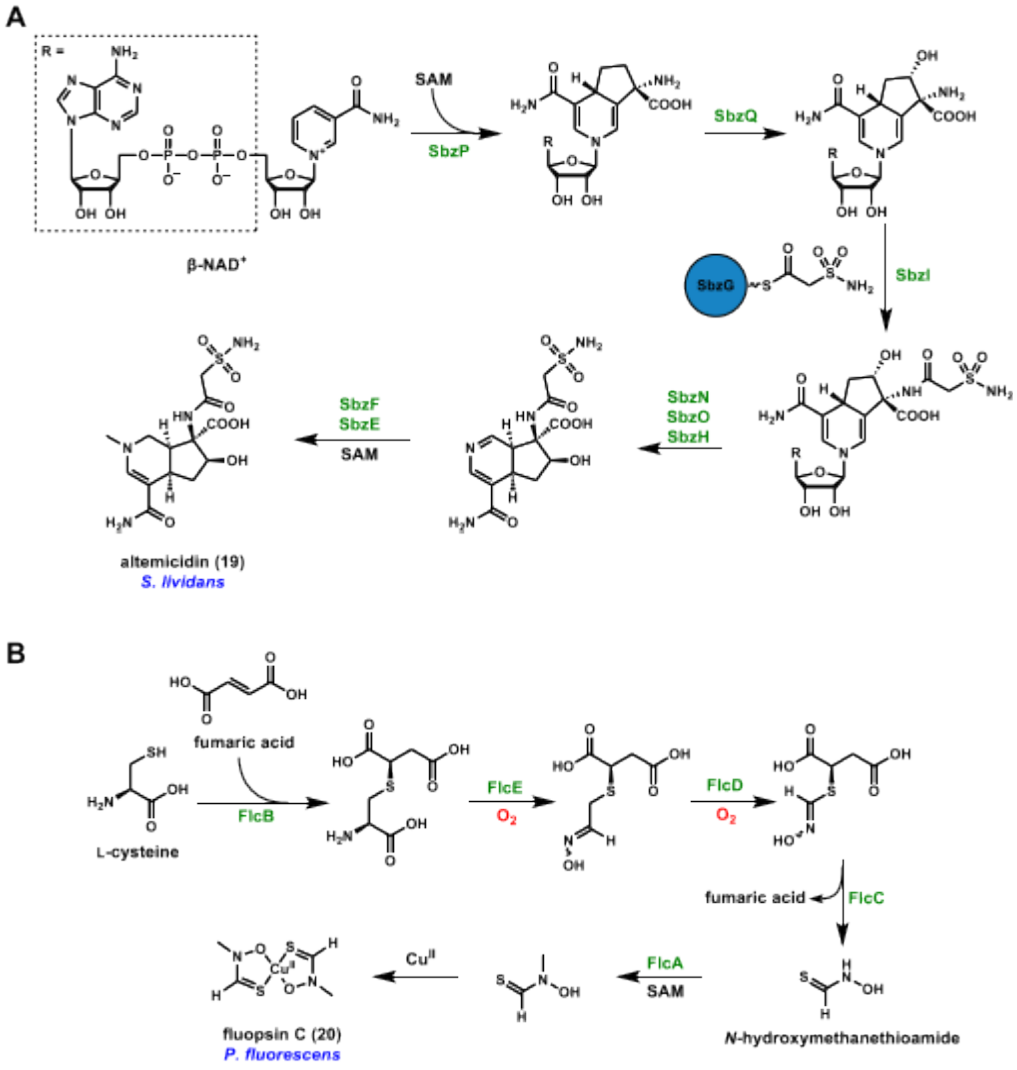


Figure 39. Examples of natural products biosynthesized without known core enzymes. Heterologous expression played an important role in pathway elucidation. (A) Biosynthetic pathway of **19**; (B) biosynthetic pathway of **20**.

Additional examples of natural products biosynthesized from BGCs without core enzymes are shown in Figure 40 and briefly discussed below.

i) *E. coli* was used for functional reconstitution of a two-gene operon that is conserved in over 50 proteobacteria. The two genes include an unannotated ThiF-nitroreductase didomain enzyme and an *N*-acyltransferase. Analysis of *E. coli* expressing the two enzymes led to the discovery of a new class of oxazolone-containing compounds including **179**.[390]

ii) Heterologous expression in *E. coli* followed by mass spectrometry analysis led to the discovery of angustmycin A (**180a**) and C (**180b**).[391] Angustmycin C is the penultimate intermediate of the pathway prior to the final dehydration step catalyzed by AgmF. Removal of *agmF* led to targeted accumulation of **180b**. The production of these two metabolites was also achieved in *S. albus* G153 in the same year.[392]

iii) A direct cloning strategy was used to heterologously express the *cyb* cluster encoding the biosynthetic pathway for cyanobacterin (**181**), an amino acid-derived phytotoxin containing a γ-butyrolactone moiety.[393] The *cyb* BGC was edited to remove intergenic regions and putative transcriptional terminators, followed by expression in *E. coli*. While **181** itself was not produced by the heterologous host, the detection of four new metabolites provided insight into the required C-C bond formation during the γ-butyrolactone core assembly.

iv) Azoxy-containing compounds such as valanimycin (**43**) and calvatic acid have interesting bioactivities; for example, **43** is prized as an antibiotic.[105] Heterologous reconstitution of the *vlm* cluster in *S. albus* J1074 was key to understanding the previously-unknown enzymatic logic behind the biosynthesis of the azoxy moiety. The enzyme VlmO was shown to be a hydrazine synthase, catalyzing an intramolecular rearrangement of the N-O bond in *O*-seryl-isobutylhydroxylamine to yield an N-N bond-containing molecule. This rearrangement resembles those of cupin-family enzymes implicated in bacterial N-N bond formation.[394] In the following step, the hydrazine moiety from the VlmO product was converted into an azoxy group in a reaction catalyzed by the enzyme VlmB.

v) Biosynthesis of the heme-like natural product class of tolyporphins including tolyporphin A (**182**) was investigated in *E. coli.* An intermediate in heme biosynthesis, coproporphyrinogen III, was shown to

undergo oxidative decarboxylation and multiple C-C bond cleavages to yield unsubstituted β sites in the pyrrole moieties.[395]

vi) The heterologous production of (–)-cispentacin (**183**) was demonstrated in *S. albus* G153, which forms part of the compound amipurimycin (**184**), a nucleoside antibiotic.[396] Amipurimycin biosynthesis was also characterized in the same strain.[397]

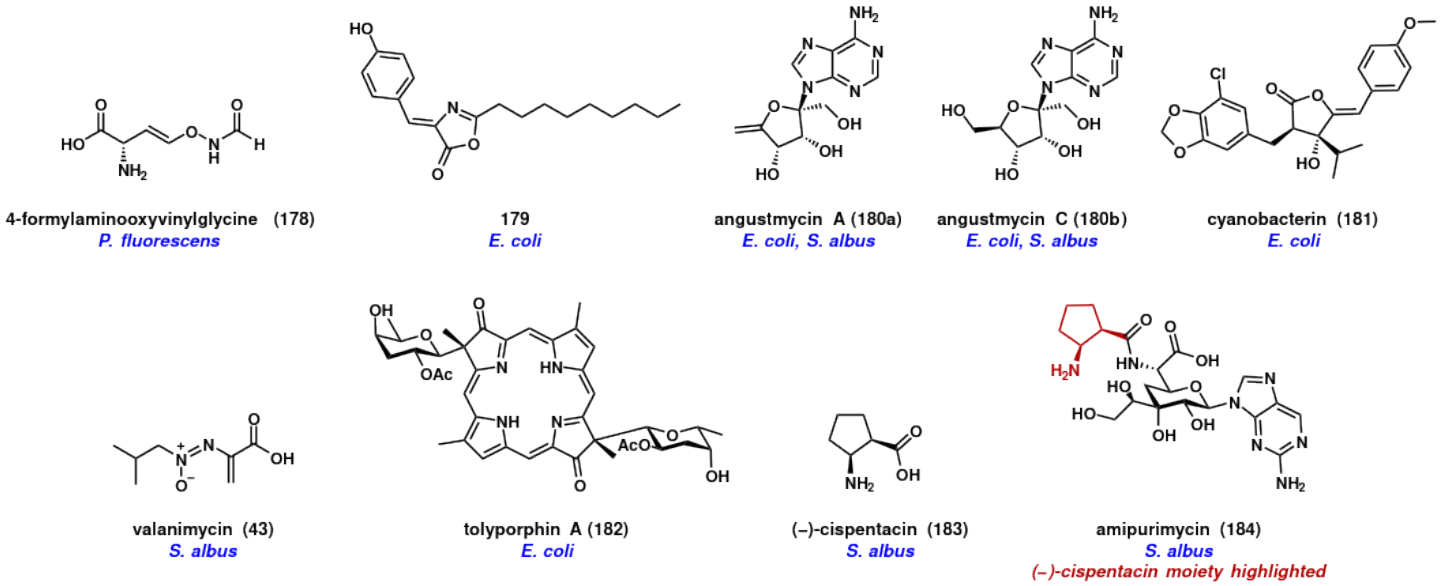


Figure 40. Examples of natural products from BGCs without well-known core enzymes.

3.3 Heterologous Expression in Yeasts

Heterologous reconstitution of natural products of eukaryotic origin, whether from fungi, plants, or animals, has seen the most success in single cell yeast organisms, represented prominently by *Saccharomyces cerevisiae*, also known as Baker's Yeast. Other yeasts used for such endeavor include the oleaginous yeast *Yarrowia lipolytica* and the methylotrophic yeast *Pichia pastoris*. Although all these organisms are classified as fungi, our discussion of their roles in natural product biosynthesis will be separated from the ascomycete fungi section (Section 3.4).

3.3.1 Baker's Yeast as a Heterologous Host

S. cerevisiae (referred to as yeast in this section) is the eukaryotic organism of choice for synthetic biology pursuits as it was the first eukaryote to be fully sequenced.[398] *S. cerevisiae* is the gold standard for heterologous expression of plant

natural products, including terpenes, alkaloids, phenylpropanoids, flavonoids, and anthocyanins.[399] This single-celled organism is characterized by fast growth and well-established genetic tools and has GRAS status.[400] Baker's yeast has cellular compartments similar to other higher order organisms, thereby supporting the expression of plant enzymes, whie its endomembrane systems are conducive to expression of complex plant proteins as well. For example, the plant cytochrome P450 oxygenases are anchored to the ER in the native organism, and these enzymes cannot be reliably expressed in prokaryotic organisms.[401] Yeast is also highly adaptable to automation advances in molecule biology (BioFoundry), which makes it an attractive host for the DBTL paradigm. Discussed below are recent examples of heterologous biosynthesis grouped into compound families.

3.3.1.1 Sesquiterpene and Diterpene-derived Natural Products

Yeast has been an attractive heterologous host for producing terpenes and terpenoid compounds since the early 2000s. Expression of plant and fungal terpenoids has been particularly successful. This is owing to the robust isoprenyl metabolism in yeast from the MEV pathway, as well as the ability to reconstitute P450 functions as described above. The groundbreaking example is Keasling lab's biosynthesis of precursors to the anti-malarial artemisinin (**8**) at titers up to 100 mg/L (Figure 41A).[232] **8** is traditionally extracted from *Artemisia annua L*, but this sourcing is not sustainable to meet global demands.[402] Hence, achieving high titers from yeast and semisynthesis is highly impactful. Subsequent metabolic engineering achieved titers of 40 g/L of the precursor amorphadiene (**185**) and 25 g/L of artemisinic acid (**94**), which can be readily converted to **8** via semisynthesis with an industrial photochemical conversion process (Figure 41B).[26,403]

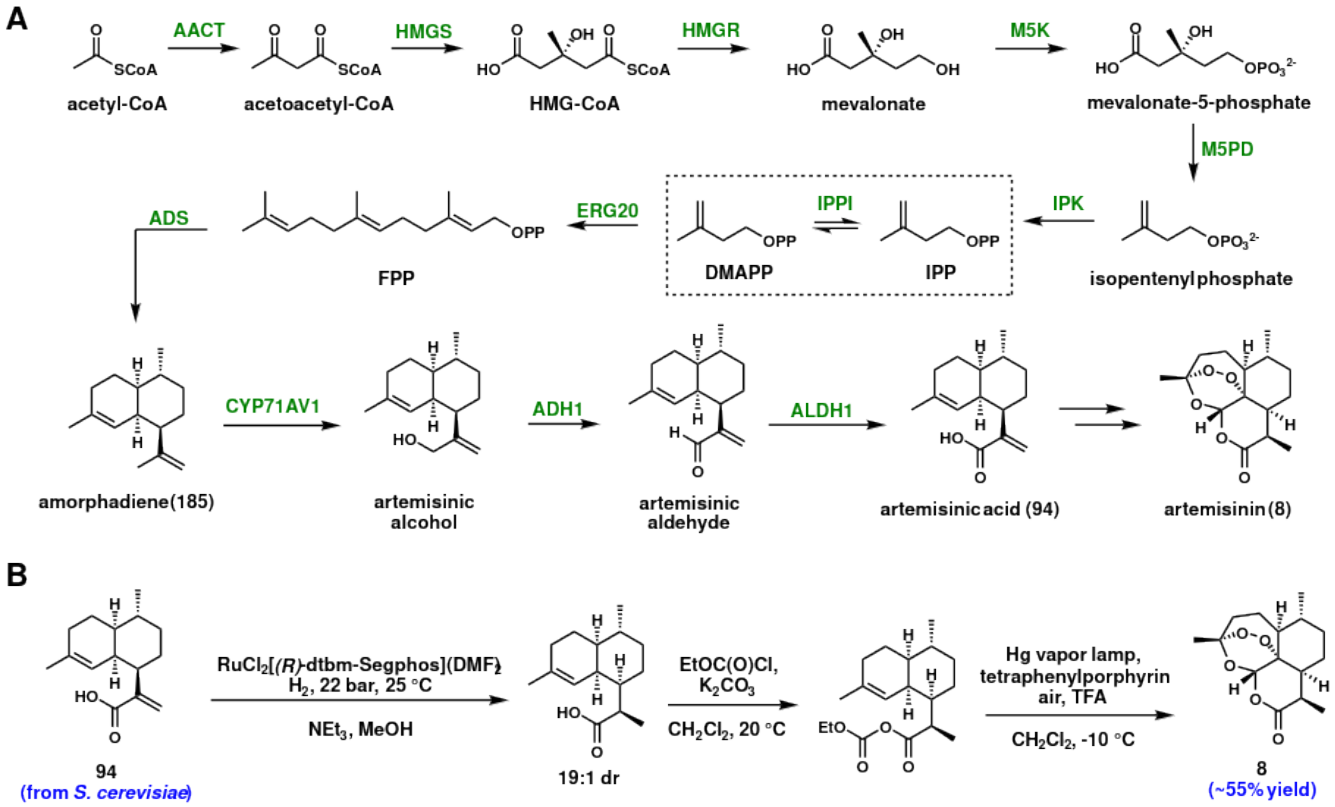


Figure 41. Production of **94** in yeast followed by semisynthesis to **8.** (A) biosynthetic pathway to **94** is reconstituted in yeast at high titer; (B) Sanofi's semisynthetic preparation of **8** from **94**.

Terpenes and terpenoids are common ingredients in perfumes and fragrances. Santalol is the major component of sandalwood oil that has a unique aroma. Zi et al. demonstrated this sesquiterpenoid can be heterologously produced in yeast, and the titers can be improved to over 1 g/L in fed-batch fermentation by downregulating expression of the squalene synthase gene *ERG9*.[404] One requirement for industrial santalol extraction from plants, however, is to maintain a ratio of α-santalol (**186**) to β-santalol (**187**) between 2:1 to 3:1 (Figure 42). Using a product-specific santalene synthase as the starting point, multiscale simulations led to a point mutation Phe441Val that can produce both isomers at the desired ratio at a combined titer of ~ 700 mg/L.[405] This work demonstrates an additional degree of control in heterologous biosynthesis through targeting a desired product mixture ratio.

A yeast-based expression system for terpene cyclases was used to characterize such enzymes from the mugwort plant *Artemisia argyi*, leading to the identification of 54 terpene products, including a novel compound named cyclosantalol.

[406] Activity screening revealed that a sesquiterpene compound, (+)–intermedeol (**188**), exhibited excellent insect-repellant activity. A biosynthetic titer of 2.34 g/L of **188** was achieved in yeast.

Glycosylated diterpenes such as rubusoside (**189**) and rebaudiosides, including rebaudioside A (**190**) (Figure 42), are artificial sweeteners that are extracted from plant species. These steviol glycosides are hundreds of times sweeter than sucrose with a low caloric count. A multimodular approach was used to establish a yeast strain to make these compounds. [407] The diterpene titer was first boosted through overexpression of rate-limiting enzymes in the MEV pathway and minimizing loss of isoprenyl precursors to primary metabolism. Expression of downstream P450s and glycosyltransferases led to the accumulation of the target compounds at an initial titer of 4.5 mg/L. Bottleneck steps were identified and addressed using a combination of spatial engineering of P450s enzymes and expanding the capacity of ER for overexpressed P450s. Additional yeast cell engineering steps include identification and overexpression of transporters, stress-response factors, and knockout of genes predicted by genome-scale metabolic models to benefit steviol glycosides titers. Collectively, these efforts resulted in final titers of 1.4 g/L of **189** and 133 mg/L of **190** in fed-batch fermentation.

Multiplexed expression level tuning was applied to *S. cerevisiae* to optimize terpene titers. Four bacterial transcription regulators that respond to four distinct small-molecule inducers: vanillic acid, xylose, anhydrotetracycline, and isopropyl β-D-1-thiogalactopyranoside (IPTG), were used. Each of the four promoters within this array can be independently controlled, offering a wide dynamic range spanning from low background to 40- to 5000-fold increase in gene expression. The authors demonstrated the utility of this approach by optimizing a four-gene heterologous pathway involved in biosynthesis of the terpene linalool (**172**). Bottlenecks in the pathway were successfully identified and characterized.[408]

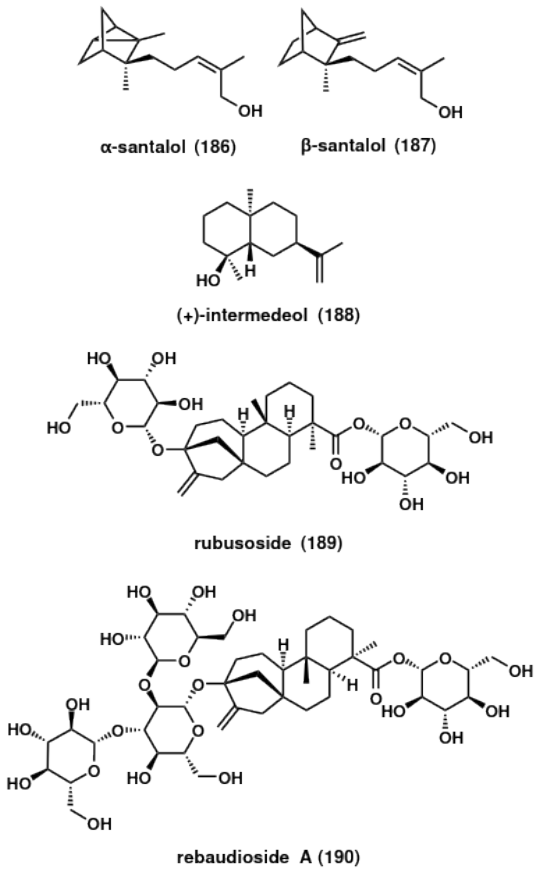


Figure 42. Examples of sesquiterpenes and diterpenes that have been produced in yeast.

3.3.1.2 Triterpene and Sterol-derived Natural Products

Triterpenoids are natural products that are most commonly derived from the C30 precursor squalene.[409] Squalene is a precursor to sterols in mammals, ergosterols in fungi and β-amyrin in plants. The Keasling lab reported the complete yeast biosynthesis of the potent vaccine adjuvant QS-21 (**191a-b**) (Figure 43), of which the supply is limited by low yields from conventional plant extraction.[410] 38 enzymes were introduced to produce the highly decorated triterpenoid at 0.0012% w/w QS-21 per dry cell weight. Although the titer was low in this remarkable initial demonstration, there is no doubt that with further strain optimization, the supply issue associated with QS-21 can be resolved with the yeast platform. The production of **191a-b** and related compounds *in planta* is described in detail in Section 3.5.

Ginsenoside Rh2 (**192**), an anticancer drug from ginseng, was produced in yeast (Figure 43).[411] The authors identified the bottleneck enzyme as the glucosyltransferase that decorates the triterpene skeleton. Through semi-rational design, catalytic efficiency of the glucosyltransferase was improved by approximately 1800-fold. This, combined with other improvements, including improving UDP-glucose precursor supply, led to a 900-fold increase in titer of **192**. Ganoderic

acids (GAs) from medicinal mushrooms are triterpenoids with benficial biological activities. The Xiao group sequenced the producing strain *Ganoderma lucidum* and used *S. cerevisiae* to iteratively screen 158 candidate P450 enzymes that may be involved in GA biosynthesis.[412] Two P450s were identified that can transform lanosterol into two desired GA molecules including ganoderic acid A (**193**). Integration of multiple copies of the P450 enzymes and a reductase partner led to titers of GAs in excess of 50 mg/L. An improved yeast strain to produce campesterol (**194**) was recently reported.[413] The authors optimized sterol acyltransferase activity and improved the supply of FPP, the immediate precursor to squalene. Genome sequencing of strains recovered from adaptive laboratory evolution was performed to identify genes linked to improved sterol production. These genes were then overexpressed in the final strain.

Triterpenes that do not originate from the squalene precursor, but rather are cyclized from hexaprenyl pyrophosphate derived from the head-to-tail extension of GGPP, were identified through heterologous expression of three bifunctional terpene cyclases in engineered yeast.[414] This study led to the production of talaropentaene (**195**), macrophomene (**196**), and colleterpenol (**197**) (Figure 43).

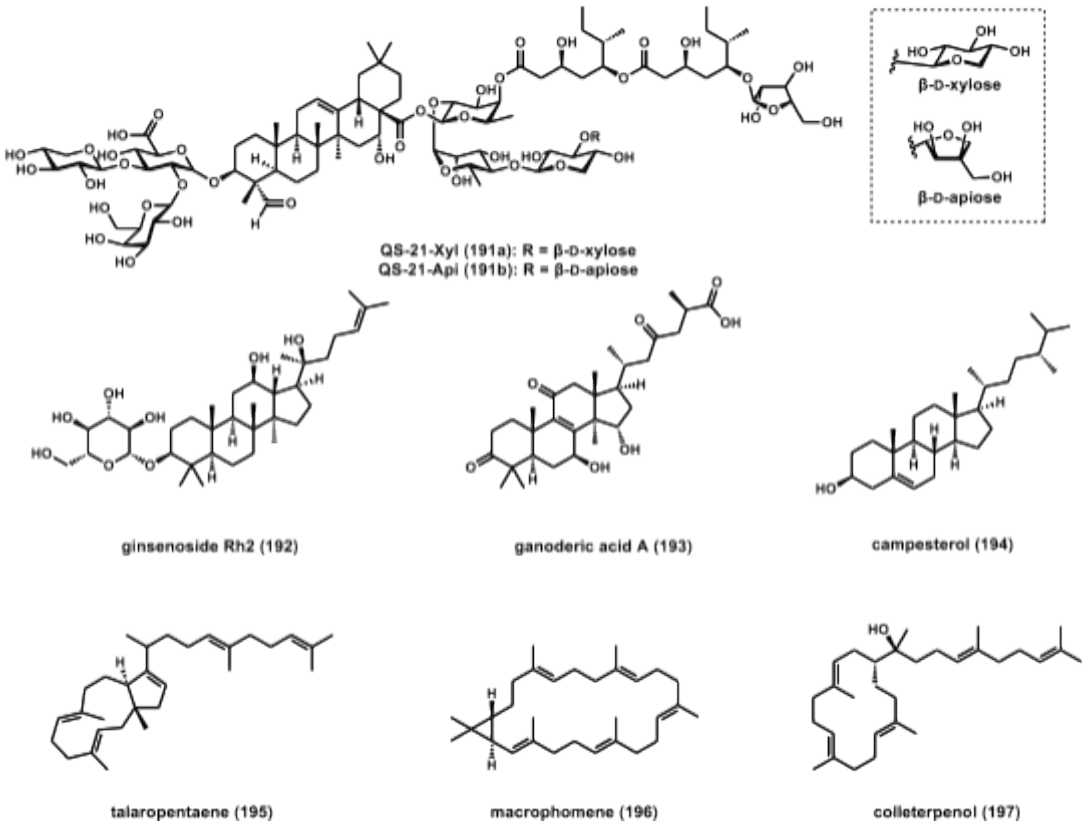


Figure 43. Examples of triterpene and sterol natural products that have been produced in yeast.

3.3.1.3 Alkaloid Natural Products

Alkaloids of many subfamilies have been reconstituted in yeast, including tetrahydroisoquinolines (THIQs), tropanes, monoterpene indole alkaloids (MIAs), ergots etc.[415–417] Over 3,000 THIQ alkaloids have been identified from plant organisms, among which are the opioids morphine (**9**) and codeine (**198**). These metabolites are derived from two molecules of L-tyrosine via an extended pathway that involves a number of P450 enzymes, with (*S*)-reticuline (**199**) as a key biosynthetic intermediate. An impressive tour-de-force accomplishment is Smolke's reconstitution of thebaine (**200**) and hydrocodone (**92**) in yeast, in which over twenty enzymes originating from all kingdoms of life were assembled (Figure 44).[236] This synthetic biology feat sets the stage for engineered biosynthesis of non-addictive opioids from a microbial host. Based on this approach, the same lab also reconstituted the biosynthesis of noscapine (**95**) (Figure 44), a pthalideisoquinoline alkaloid isolated for centuries from the native *Papaver somniferum* plant, at mg/L titers.[231] A yeast-based platform was also developed for the production of **9** and neomorphine from **200**.[418] The Martin group developed a yeast strain that can produce the key THIQ intermediate **199** at 4.6 g/L.[236] A key strategy was inactivating competing oxidoreductases that shunts the precursor 4-hydroxyphenylacetaldehyde (4-HPAA) to the corresponding acid or alcohol. The authors observed that amino acids other than L-tyrosine can be incorporated by the pathway to form substituted THIQs. The Dueber lab noted that the enzyme norcoclaurine synthase (NCS), which catalyzes formation of (*S*)-norcoclaurine (**201**) from dopamine and 4-HPAA, is toxic to yeast when overexpressed.[419] To address this issue, NCS was expressed as a tagged protein that is compartmentalized into the peroxisome, of which the membranes are permeable to both the substrates and **201**. Furthermore, to accommodate the overexpressed NCS, the authors expressed engineered transcriptional factors that induce expansion of available peroxisomes.

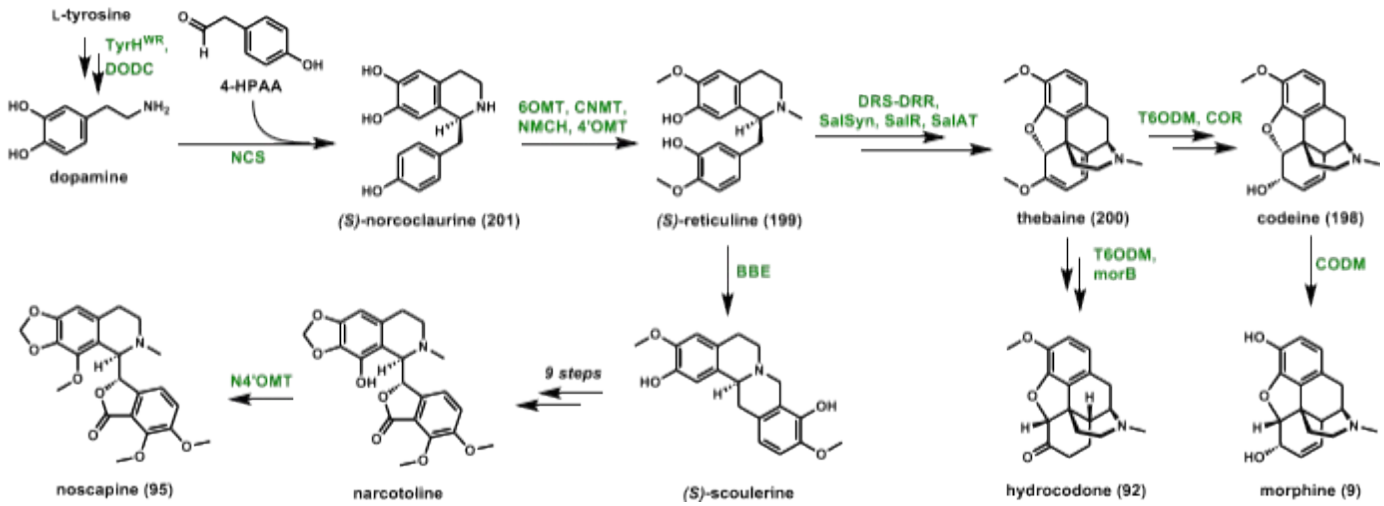


Figure 44. Heterologous biosynthesis of THIQ alkaloids in yeast.

The Smolke lab also demonstrated the biosynthesis of various tropane alkaloids, namely hyoscyamine (**202**) and scopolamine (**203**), as well as *de novo* production of an analog cinnamoyltropine (**204**) (Figure 45).[416,420] This was accomplished in yeast using a modular approach to overcome several bottlenecks. The tropane pathway originates from phenylalanine and putrescine (**205**), the latter derived from L-arginine. The authors significantly enhanced putrescine production by more than 70-fold using rational strategies including the expression of genes involved in arginine metabolism, disrupting regulatory mechanisms that control polyamine levels, reconstructing an alternative biosynthetic pathway for putrescine production, etc. Additionally, production of the key precursor tropine (**206**) reached 6 mg/L after overexpressing genes such as the type III PKS *abPYKS*, which condenses two malonyl-CoA units with *N*-methylpyrrolinium (**207**) to form 4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoic acid (**208**). Further P450-catalyzed ring closure, followed by reduction of the resultant ketone to an alcohol, yields **206**. A late-stage compartmentalization strategy was used to mimic the spatial organization found in the native plant producer. Identification and overexpression of plant transporters that improve intermediate transport in yeast led to further increases in titers of **202** and **203** titers to 480 µg/L and 172 µg/L, respectively.[421]

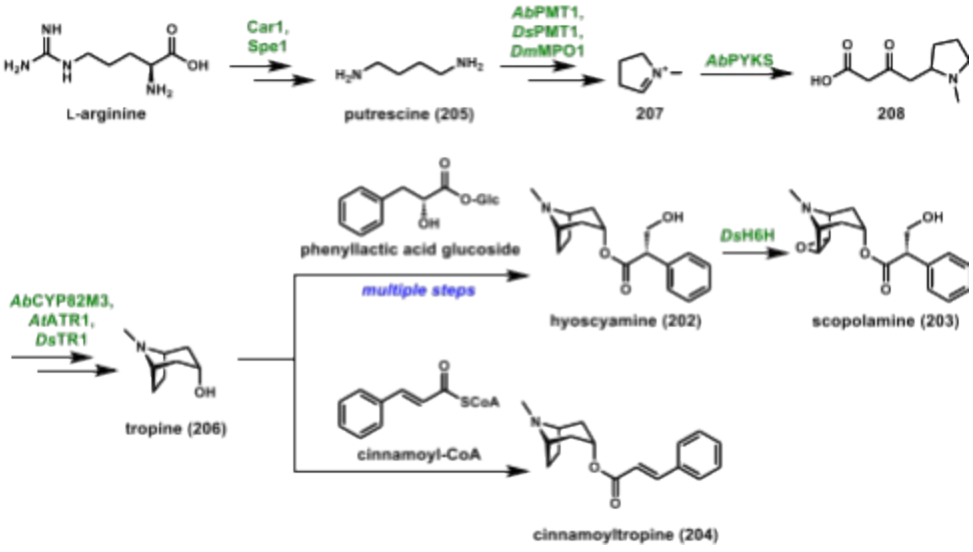


Figure 45. Heterologous biosynthesis of tropane alkaloids in yeast.

MIAs are plant alkaloids that include thousands of bioactive compounds, many of which have therapeutic value that treat diseases such as cancer, Alzheimer's disease, and malaria. Most notable compounds in this family include

vinblastine (**209**), vincristine (**210**), ajmaline (**211**), serpentine (**212**), alstonine (**213**), mitragynine (**214**), etc. (Figure 46). The universal biosynthetic intermediate is the molecule strictosidine (**50**), which derives from a ten-gene pathway that involves terpene (GPP), amino acid (L-tryptophan) and sugar (UDP-glucose). Biosynthesis of this molecule has been intensively studied by many groups, originating with the first yeast reconstitution by the O'Connor lab in 2015.[422] Subsequent identification of missing genes in the pathway enabled more robust titers (30-50 mg/L) from yeast, including efforts by the Keasling, Lian, and Tang labs in recent years.[423–426] Yee and coworkers compartmentalized the biosynthesis of geraniol precursor to the mitochondria to take advantage of the high flux of acetyl-CoA, as well as to avoid consumption of GPP by the downstream enzyme FPPS in the cytosol.[427] This resulted in an increase of ~30 fold in titer of the monoterpene compared to cytosolic expression. The high titers of **50** enabled reconstitution of downstream pathways in yeast, highlighted by full reconstitution of vinca alkaloid **209** by the Keasling and Jensen labs, in which 56 gene edits were performed, including expression of 34 heterologous genes.[423] Parallel efforts by others have resulted in the reconstitution of **211**-**214**.[428–430] Most recently, Keasling, Jensen, and coworkers reported the *de novo* production of **214** in yeast, achieving a titer of 290 µg/L and producing fluorinated and hydroxylated derivatives of **214**.[431]

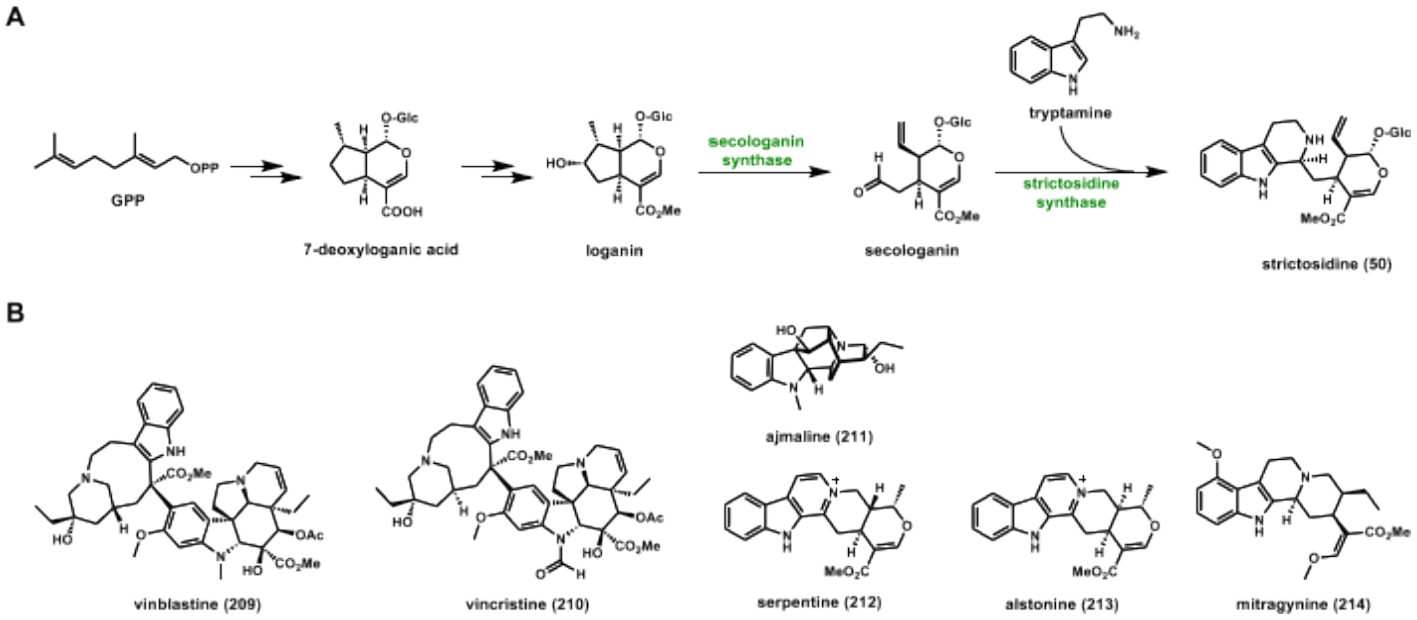


Figure 46. Production of MIAs in yeast. (A) Biosynthetic pathway to **50**, precursor to MIAs; (B) Examples of MIAs produced in yeast.

Ergot alkaloids are fungal alkaloids derived from grass pathogens such as the fungus *Claviceps purpurae*.[432] These compounds are used in the treatment of neurological disorders because they bind to the receptors for the neurotransmitters serotonin, dopamine, and others. The biosynthesis of ergot alkaloids starts with the formation of D-lysergic acid (**10**), which involves prenylation of L-tryptophan to form 4-dimethylallyl tryptophan (4-DMAT, **215**) followed by oxidative cyclization to yield chanoclavine I (**216**) (Figure 47). The seven-enzyme pathway was recently reconstituted in yeast in a final titer of 1.7 mg/L.[399] A modular compartmentalization approach was used to achieve overproduction of agroclavine (**217**), which is a late-pathway intermediate to **10** and is a potent anti-depressant.[433] The pathway was segmented into two modules, with bottlenecks in each module separately identified and addressed. By anchoring the downstream module (three enzymes) to the ER, a two-fold increase in titer of **217** was observed. Further increasing NADPH availability via overexpression of the NADH kinase POS5 led to additional increases in titer. Collectively, these modifications resulted in 150 mg/L of **217** in fed-batch fermentation.

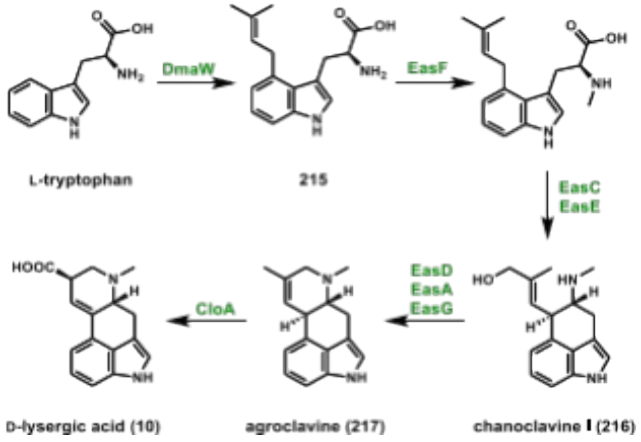


Figure 47. Heterologous biosynthesis of ergot alkaloids in yeast.

3.3.1.4 Polyketides, NRPs, and Meroterpenoids

Polyketide biosynthesis in yeast has been extensively explored, especially iterative PKSs that are smaller compared to multimodular assembly line PKSs. The yeast strain BJ5464-NpgA is especially engineered for PKS reconstitution and contains genetic knockouts of vacuolar proteases such as PEP4 and PBP1, as well as integration of the fungal phosphopantetheinyl transferase NpgA.[434] The lovastatin biosynthetic pathway, which includes two PKSs, LovB and LovF, was reconstituted in yeast.[235,435] Yeast has also been engineered to biosynthesize the semisynthetic derivative simvastatin using a precursor directed approach (See Section 4.2.9.3).[234] Harvey and coworkers developed the strains DHY213 and

JHY686, the latter which not only incorporates NpgA and a fungal P450 reductase partner, but also shows improved growth and mitochondrial stability. Using the DHY strain as a host, the authors developed the HEx (heterologous expression) platform to rapidly assemble synthetic DNAs into pathways for expression in yeast.[436] The auto-inducible ADH2 promoter from yeast, and homologous promoters from other yeast species, were used to drive the expression of biosynthetic enzymes upon depletion of glucose. Using this strain and automated plasmid assembly, 41 fungal BGCs from diverse fungi were investigated, including a number of PKS-anchored BGCs, 22 of which produced detectable compounds.

The yeast strain has also been used to express PKSs from animals, of which the metabolites play unique roles in the native hosts (Figure 48). For example, the EcPKS1 from sacoglossan mollusk *Elysia chlorotica* was expressed from BJ5464-NpgA to reveal polypropionate-derived polyene-pyrones **218** and **219**, which are proposed to serve as UV protectants for the host.[437] In addition, the MuPKS from the parrot *Melopsittacus undulatus* was expressed in yeast and resulted in the formation of yellow pigments known as psittacofulvins (**220**) that are responsible for coloration in the feathers.[438] A single mutation in MuPKS in the parrot resulted in blue pigmentation.

The production of NRPs in yeast has been explored and is exemplified by the heterologous production of penicillin G (**96**).[439] Coexpression of enzymes including ACV synthase, NpgA, and PclA and PenDE that convert ACV tripeptide to **96** in *S. cerevisiae* BY4741 were performed. Following promoter screening, a titer of 5 μg/L of **96** was achieved. The authors speculated that implementing additional metabolic engineering strategies, including further boosting peroxisomal expression of PclA and upregulating peroxisome formation, could lead to increased yields.

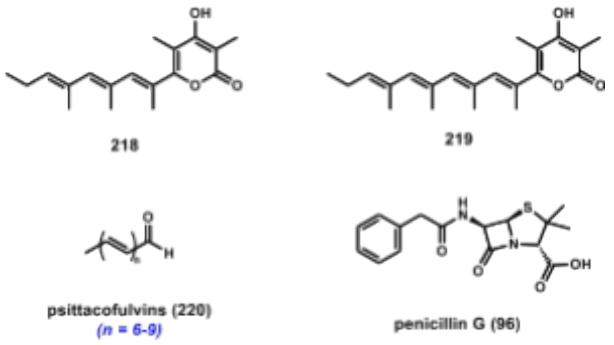


Figure 48. Representative polyketides and NRPs produced in engineered yeast.

Yeast has been used as a host for the *de novo* production of cannabinoids (Figure 49).[233] Cannabinoids such as tetrahydrocannabinol (THC, **221**) and cannabidiol (CBD, **222**) are meroterpenoid natural products, consisting of terpene

(GPP) and polyketide (olivetolic acid, **223**) building blocks. The biosynthesis of cannabinoids was partially elucidated, including the plant PKS that synthesizes **223**, and the cyclases that form either THCA (**224**) or CBDA (**225**). The native prenyltransferase that prenylates **223** to cannabigerolic acid (**93**) was not known prior to the yeast reconstitution study, although engineered versions from microbial sources are available.[440] Keasling and coworkers accomplished full reconstitution of a number of cannabinoids in yeast at mg/L titers. This was enabled by a number of breakthroughs, including i) achieving high titer of the PKS starter unit hexanoyl-CoA (**226**); ii) elevating the cellular GPP concentration; and most importantly iii) identifying the plant prenyltransferase CsPT4 that performs the geranylation step. The full reconstitution of the cannabinoid pathway enabled engineered biosynthesis of rare cannabinoids that are derived from alternative starter units for medicinal applications. Indeed, the authors were able to provide the yeast cell with different short-to-medium chain acids that can be incorporated into the final cannabinoids. In a follow up study, the same group achieved a 0.5 g/L titer of the key intermediate **93**.[441] A suite of metabolic engineering strategies was employed, including tuning the level of fatty acid synthase (FAS) to minimize off-pathway consumption of **226**, and improving the stability and activity of CsPT4 in yeast through fusion to an auxiliary protein.

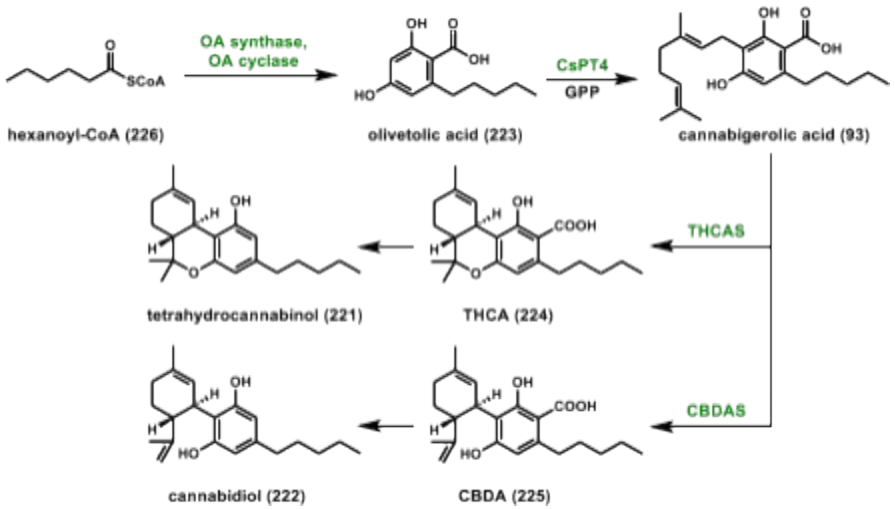


Figure 49. Heterologous biosynthesis of cannabinoids in yeast.

3.3.1.5 Flavonoids

Yeast is a popular choice for the heterologous production of flavonoids, which are plant derived metabolites with diverse biological activities. The Koffas group among others have provided excellent review of application of synthetic biology to microbial flavonoid production.[442–444] Flavonoids are synthesized by a type III PKS with a starter unit derived from

aromatic amino acid catabolism, such as *p*-coumaric acid (**227**)**,** caffeic acid (**228**), or ferulic acid (**229**), followed by several rounds of polyketide chain extension using malonyl-CoA as an extender unit (Figure 50). The nascent PKS products, such as naringenin chalcone, are further modified by tailoring enzymes into the diverse members of the family. One flavonoid of medical interest is xanthohumol (**230**) synthesized by hops. **230** is a prenylated and methylated derivative of naringenin chalcone, and has displayed anticancer, antidiabetic, and other activities. The Zhou group engineered *S. cerevisiae* to synthesize **230** through a series of rational approaches including balancing three parallel biosynthetic pathways (aromatic amino acid, malonyl-CoA and isoprenoid) and engineering bottleneck enzymes to achieve an 83-fold improvement in titers.[445] Glycosylation is an important and common modification of flavonoids that enhances water solubility and efficacy. The Shan group engineered a pathway in *S. cerevisiae* for synthesizing flavonoid-7-*O*-disaccharides. Through metabolic and promoter adjustments, the UDP-rhamnose and UDP-glucose pathways were optimized and resulted in desired glycosylated products at hundreds of mg/L titer.[446]

To increase the level of the precursor *p*-coumaric acid (**227**) in yeast, a recent study optimized the carbon allocation between glycolysis and aromatic amino acid metabolism, resulting in g/L titers.[447] The Yuan group optimized biosynthetic titers of **228** in yeast using a combination of rational and library-based methods. To enhance the synthetic capabilities of yeast for related compounds, screening of yeast libraries using the **S**ynthetic **C**hromosome **R**e**a**rrangement and **M**odification **b**y **L**oxPsym-mediated **E**volution (SCRaMbLE) strategy was performed to provide new insights into how yeast regulates the synthesis of aromatic amino acids.[448] As biosynthesis of the phenolic acids require multiple cofactors such as NADPH, $FADH_2$, and SAM (for ferulic acid), rewiring the supply and recycling of these cofactors in yeast led to significant increases in the titers of **227** (5.5 g/L) and **228** (3.8 g/L).[449] The Borodina group integrated CRISPRi (see section 2.2.3) and high throughput screening to identify new genes targets for enhancing **227** production, using the colorimetric betaxanthins as a readout.[450] Large guide-RNA (gRNA) libraries deregulating 1000 metabolic genes *in S. cerevisiae* were screened and six targets that increased the titer were identified. Further screening of a gRNA multiplex library identified the unanticipated combination of knockdown of PYC1 (pyruvate carboxylase) and NTH2 (trehalose) which resulted in a 3-fold increase in titer of **227**.

The coumarin-derived osthole (**231**), an anti-asthmatic derived from *Cnidium monnieri,* was produced from engineered yeast. The authors initially developed an umbelliferone (UMB)-producing yeast strain, followed by integration

of prenyltransferase and methyltransferase genes from different organisms. Typical challenges such as protein expression, cofactor supply, precursor availability and rate-limiting steps were addressed to result in hundreds of mg/L of **231**.[451] The related courmarin product marmesin (**232**) was produced from *S. cerevisiae*. The authors boosted aromatic amino acid biosynthesis, in particular that of L-tyrosine, to increase precursor supply. Rational engineering of the umbelliferone 6-dimethylallyltransferase and marmesin synthase increased **230** production further.[452]

A *S. cerevisiae* strain was developed to synthesize raspberry ketone (RK, 4-(*p*-hydroxyphenyl)-2-butanone, **233**), which is a high value compound used in flavoring and fragrances. **233** can be accessed from L-tyrosine in four steps (and from **228** in three steps) and the genes have been identified. A multitude of metabolic engineering strategies were used to arrive at a final titer of 310 mg/L of 4-hydroxybenzalacetone, which can be readily hydrogenated to arrive at **233**. These strategies involved promoter tuning, modular pathway engineering and coculturing.[453]

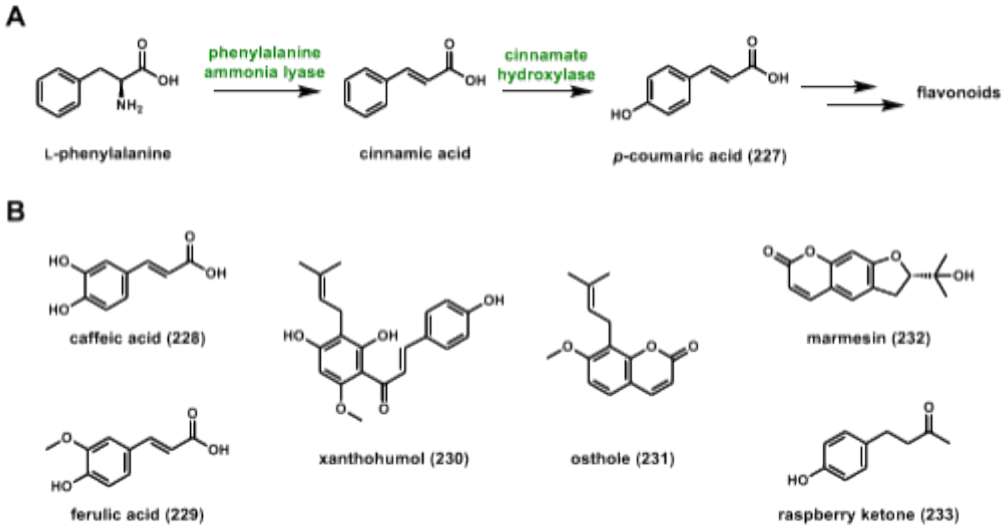


Figure 50. Examples of flavonoids produced heterologously in yeast**.** (A) Early stages of flavonoid biosynthesis from L-phenylalanine; (B) Selected examples.

3.3.2 *Yarrowia lipolytica* and *Pichia pastoris* as Yeast Hosts

Yarrowia lipolytica is an oleaginous yeast, notable for its ability to accumulate upwards of 20% lipids in dry cell weight. This type of yeast is mainly used to synthesize lipid-based compounds.[454] *Y. lipolytica* was recently engineered to synthesize resveratrol (**234**) along with other metabolites derived from the shikimate pathway, *cis,cis*-muconic acid (**235**), and salicylic acid (**236**) (Figure 51). Through further genetic modifications, including the incorporation of feedback-

insensitive alleles of key genes and the addition of extra copies of the heterologous biosynthetic genes, titer of **234** reached g/L levels in a bioreactor.[455] The Stephanopoulos group was successful in synthesizing β-carotene (**171**) at ~40 g/L from this host by overcoming product inhibition of lycopene cyclase.[456] A single mutation was sufficient to attenuate inhibition without compromising enzymatic activity. Cytosine base editors (CBEs) were used in *Y. lipolytica* to increase titers of the flavonoid naringenin (**237**).[457] The biosynthetic pathway responsible for cordycepin (**238**), derived from the adaptogenic mushroom *Cordyceps militaris,* was reconstituted into *Y. lipolytica*.[458] A combination of promoter engineering, precursor flux optimization and protein engineering resulted in the titer of **238** reaching hundreds of mg/L.

Pichia Pastoris is a methylotrophic yeast strain commonly used for recombinant protein production.[459] This strain is noted for its potential for recombinant protein overexpression.[241] *P. pastoris* was recently engineered for enhanced lycopene (**173**) production.[459] The authors developed a CRISPR/Cpf1-based gene repression system and optimized gene editing techniques to achieve this goal. Additionally, they utilized the sterol regulatory element-binding protein SREBP (Sre) to regulate lipid metabolic pathways, thereby promoting overproduction of **173**.[460] As a result of these efforts, the engineered strain achieved a **173** titer of 7.24 g/L and 75.48 mg/g dry-cell-weight in fed-batch fermentation, marking the highest reported yield in *P. pastoris* to date. Furthermore, a CRISPR-based synthetic biology toolkit was developed for integrating and assembling multigene biosynthetic pathways into the chromosome of *Pichia pastoris*. [461] A range of constitutive and methanol-inducible promoters, varying in strength, were characterized. This toolkit was used to successfully produce lycopene derivatives including β-carotene (**171**), zeaxanthin (**239**), and astaxanthin (**240**).

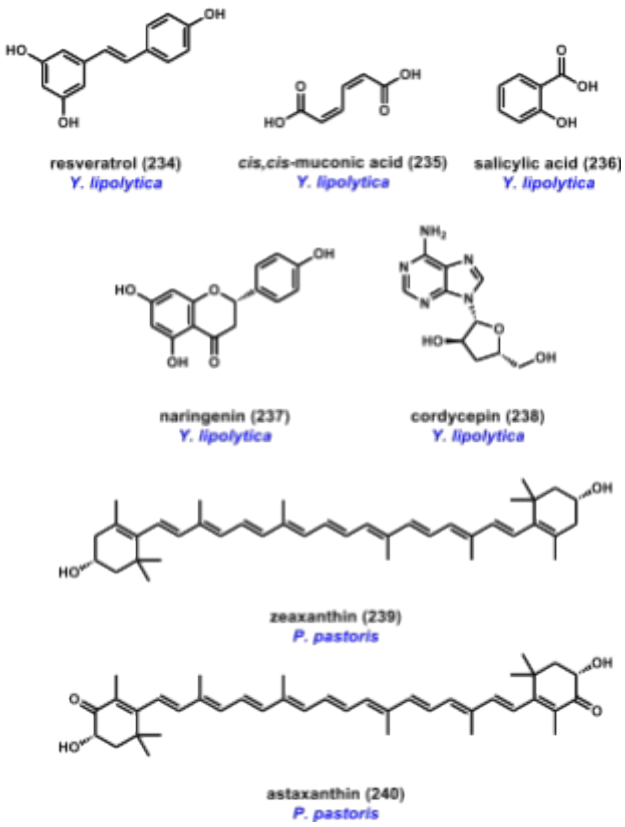


Figure 51. Examples of natural products produced in *Y. lipolytica* or *P. pastoris*.

3.4 Heterologous Expression in Filamentous Fungi

3.4.1 Introduction

While genetic protocols have been developed for a number of well-studied fungi, it remains challenging to perform genetic knockouts or overexpression for a majority of fungal species that encode natural product BGCs of interest. As a result, heterologous expression has been widely used to characterize fungal BGCs. Yeast is a robust host for terpene biosynthesis (see Section 3.3) and is continuing to be used for genome mining and pathway reconstruction of fungal terpene BGCs.[81] The HEx approach described by Harvey et al. takes advantage of yeast's powerful homologous recombination to construct BGCs from different fungi using synthetic DNAs.[436] Notwithstanding its role in natural product biosynthesis, yeast has several key drawbacks for studying complex fungal BGCs. First, yeast has minimal intron splicing machinery that cannot splice introns from other fungi, thereby requiring the use of intron-free genes that can either be constructed from synthetic DNA or from cDNA libraries.[462] Synthetic DNA can be prohibitively expensive, especially when dealing with those encoding megasynthases (often exceeding 5 kb each), while native strains may be difficult to source,

especially those that are pathogenic or phytopathogenic. In addition, expression of membrane bound enzymes, including P450s, remains inconsistent in yeast and often results in failure. As a result, a majority of fungal BGCs are now heterologously expressed in more closely related filamentous fungal hosts. Yeast remains a critical component of the workflow as a tool to generate large vectors through homologous recombination of DNA fragments.[463]

3.4.2. *Aspergillus nidulans* as a heterologous host

3.4.2.1 Molecular biology workflow of using *A. nidulans* for heterologous pathway expression

Several fungal strains have been developed as heterologous hosts. The model organism *A. nidulans* is advantageous due to well-developed genetics and relatively fast growth characteristics. Several auxotrophic markers in *A. nidulans* have been identified and have aided the genome integration or episomal vector propagation. The frequently used strains are derived from *A. nidulans* A1145, which is auxotrophic for three metabolites: uracil (Δ*pyrG*), pyridoxine (Δ*pyroA*), and riboflavin (Δ*riboB*). Several variants of *A. nidulans* A1145 have been further developed as a heterologous host, such as *A. nidulans* A1145ΔSTΔEM, which contains knockouts of the BGCs responsible for biosynthesis of sterigmatocystin (**241**) and emericellamide A (**242**), two major metabolites that interfere with detection and isolation of new metabolites .[464,465] The *A. nidulans* LO8030 strain was engineered to inactivate eight major endogenous secondary metabolite BGCs, including those associated with **241**, **242**, monodictyphenone (**243**), F-9775A (**244**) and F-9775B (**245**), austinol (**246**), terrequinone A (**247**), asperthecin (**248**), and asperfuranone (**249**) (Figure 52).[466]

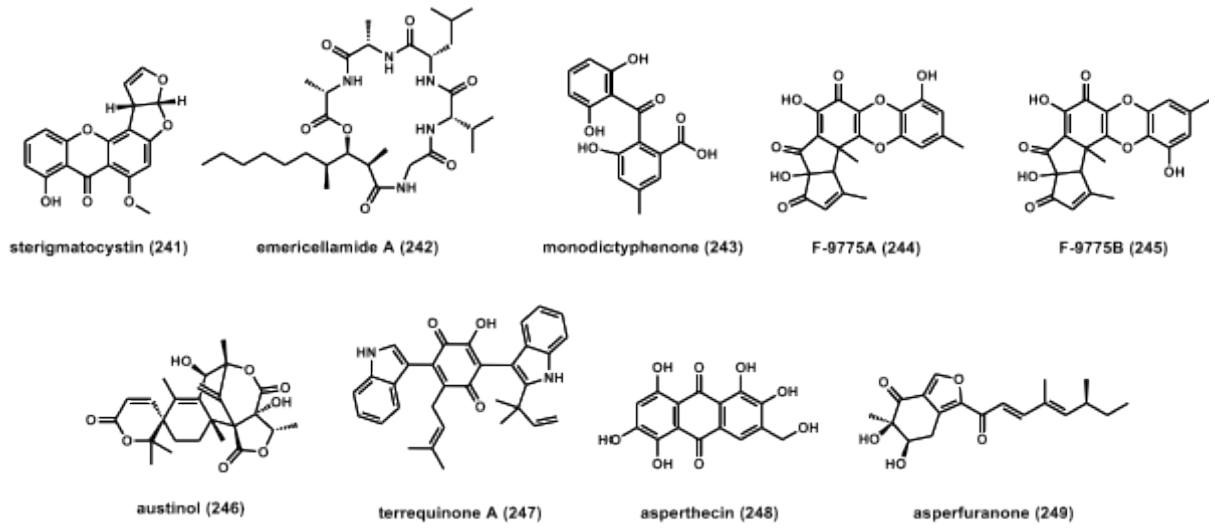


Figure 52. Endogenous secondary metabolites in *A. nidulans*.

For vector-based heterologous expression in *A. nidulans*, a BGC can be refactored into three vectors with the AMA1 origin of replication that can be maintained in *A. nidulans* (Figure 53).[465] Each vector can be selected on synthetic media with uracil, pyridoxine, or riboflavin. During refactoring, each heterologous gene is outfitted with new promoters and terminators. The promoter can be either constitutive, such as P*gpdA*, or inducible, such as P*amyB*, depending on the toxicity of the gene or the metabolite. With each vector carrying 4 to 5 genes, a BGC with 15 or less genes can be refactored. The three vectors can be propagated in *A. nidulans* (for metabolite production), yeast (for plasmid construction via homologous recombination) and *E. coli* (for storage and amplification). As an alternative to vector-based heterologous expression, integration of biosynthetic genes into *A. nidulans* is also frequently used to achieve more stable strain for metabolite production. The choice of vector or integration depends on the researcher and offers flexibility if *A. nidulans* is used over other fungal hosts.

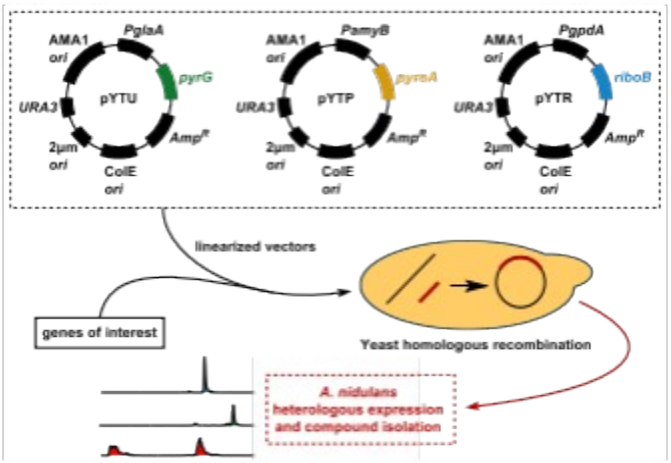


Figure 53. Three-plasmid system used to perform heterologous expression experiments in *A. nidulans* A1145.

Based on the *A. nidulans* expression host, several molecular biology tools have been developed to improve workflows. Using the 2A peptide from picornaviruses allows for the expression of multiple genes simultaneously through the use of a polycistron, an mRNA coding for multiple genes controlled by a single promoter and terminator.[467] A vector system was designed using this principle for heterologous expression of psilocybin (**98**) in *A. nidulans*.[239] This was the first reported example of heterologous expression of **98**, and was followed by studies to produce **98** in yeast and *E. coli* with the aim of increasing titer.[468,469] Polycistronic expression also facilitated the manipulation of the austinol (**246**) biosynthetic pathway in *A. nidulans,* allowing access to **246** congeners.[470] Given the widespread usage of yeast-based homologous

recombination to combine DNA fragments into expression vectors for heterologous expression, one study developed a similar DNA assembly system in *A. nidulans* with deficiencies in non-homologous end joining (NHEJ). This tool was shown to be successful for the construction of AMA1-based vectors.[471]

Fungal artificial chromosomes (FACs) have been used for heterologous expression in fungal hosts. Analogous to BAC technology discussed previously, FACs allow library creation and screening of cryptic fungal BGCs from the genome of a native producer. FACs are derived from BAC vectors to allow shuttling between *E. coli* and a fungal host such as *A. nidulans* RJW256 through incorporation of the AMA1 origin. FACs have been used in the heterologous expression of compounds such as terezine D (**250**)[472] and various terpenes.[473] FACs were used to identify the BGC for pestalamide A (**251**) during construction of a genomics-metabolomics dataset for ascomycete strains as part of an omics-based approach for natural product discovery.[474] A related construct is the hybrid yeast-fungal artificial chromosome (YFAC), which allows for yeast transformation-assisted recombination and subsequent cloning in *A. nidulans*.[475] These YFACs have been used for heterologous expression of nanangelenin A (**252**), alternapyrone (**253**), elsinochrome A (**254**), and phomacin D among other compounds (Figure 54).[475–478]

Recently, the FAC-MS platform was developed to improve scalability of the FAC-based workflow. FAC-MS incorporates a metabolomic scoring system to screen BGCs based on hits within the host's genome or in other FACs. Higher scores are assigned to MS signals that are only detected in one FAC strain, allowing prioritization of BGCs or compounds. In the initial study, 56 BGCs from various fungal species were expressed in *A. nidulans*, and 15 new metabolites were identified, including the macrolactone valactamide A (**255**), which is produced by an NRPS-PKS-anchored BGC.[479] The FAC-MS technology has been used to probe the biosynthesis of (-)-benzomalvin A (**256**) in *Penicillium* due to the high-throughput screening potential of this technology.[480] The biosynthesis of acu-dioxomorpholine (**257**) from *Aspergillus aculeatus* was also investigated using FAC-MS, and the biosynthetic pathway was heterologously expressed in *A. nidulans* (Figure 54).[481]

Other methods utilizing techniques such as USER cloning have been developed.[482] This strategy allows for targeted DNA integration into chromosome 1 of *A. nidulans*. USER cloning allows rapid construction of the targeted expression cassette, ready for integration after propagation in *E. coli.* Proof-of-concept study was carried out using the *mpaC* gene coding for a PKS involved in the biosynthesis of mycophenolic acid (**258**) (Figure 54).

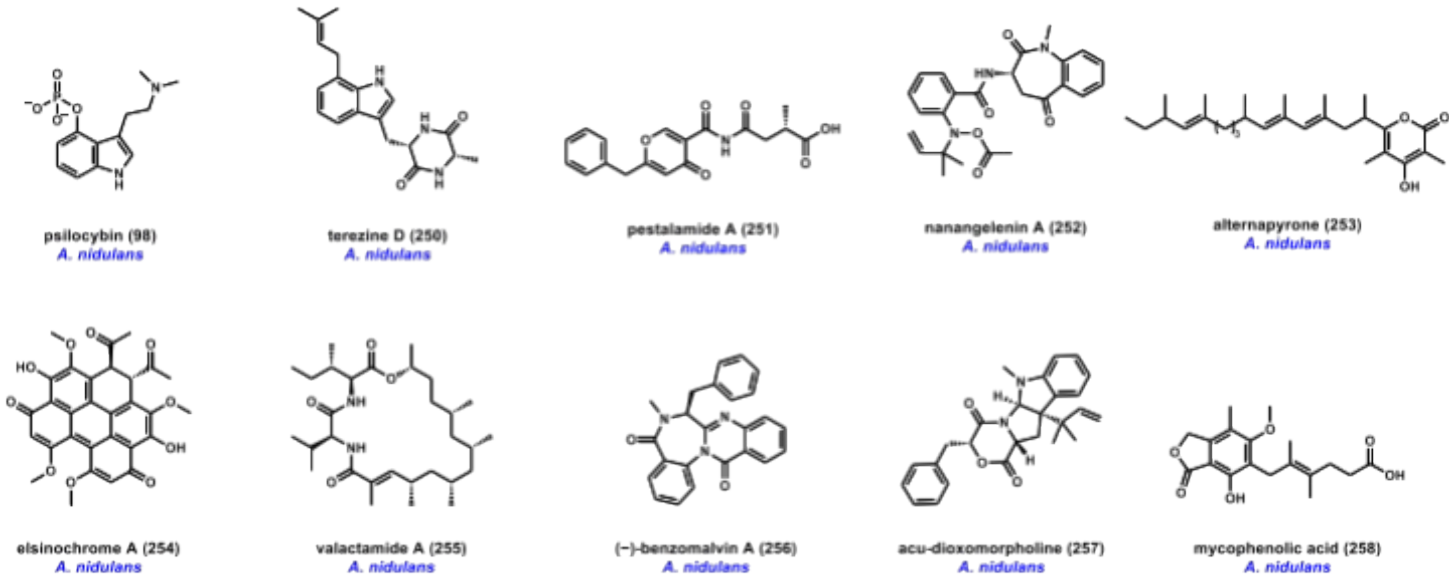


Figure 54. Examples of natural products heterologously produced in *A. nidulans.*

CRISPR gene editing technology has found many applications in filamentous fungi, including *Aspergilli*. In order to increase the number of accessible genomic sites for Cas9-mediated double strand break formation, additional PAM sites have been investigated, such as the 5'-TTN-3' PAM site recognized by Cpf1 from *Lachnospiraceae bacterium.* Codon-optimized Cpf1 was optimized for use in *A. nidulans* and shown to be effective in causing gene disruption events. Cpf1 was also shown to target different loci in *A. nidulans* and *A. niger*, resulting a more diverse set of possibilities for strain editing depending on the choice of heterologous host.[483]

3.4.2.2 Application of *A. nidulans* to Genome Mining and BGC Reconstitution

The flexibility of heterologous expression in *A. nidulans* has enabled many examples of genome mining of natural products and reconstitution of BGCs. Focusing on BGCs containing two or more core enzymes has been fruitful in mining compounds with complex structures. Such biosynthetic logic is exemplified by the cyclosporine (**259**) pathway in which a PKS is involved in the biosynthesis of an amino acid for incorporation by the NRPS into the final cyclic peptide structure.[484] Using this criterion to mine BGCs in *A. nidulans* yielded products such as aculene C (**260**), flavunoidine (**261**), paeciloxazine (**262**) and CJ-12662 (**263**) from combinations of NRPS and terpene cyclase; and thermolide A (**264**) from PKS and NRPS; and octacyclin A (**265**) from three single-module NRPS or NRPS-like enzymes (Figure 55).[485–490]

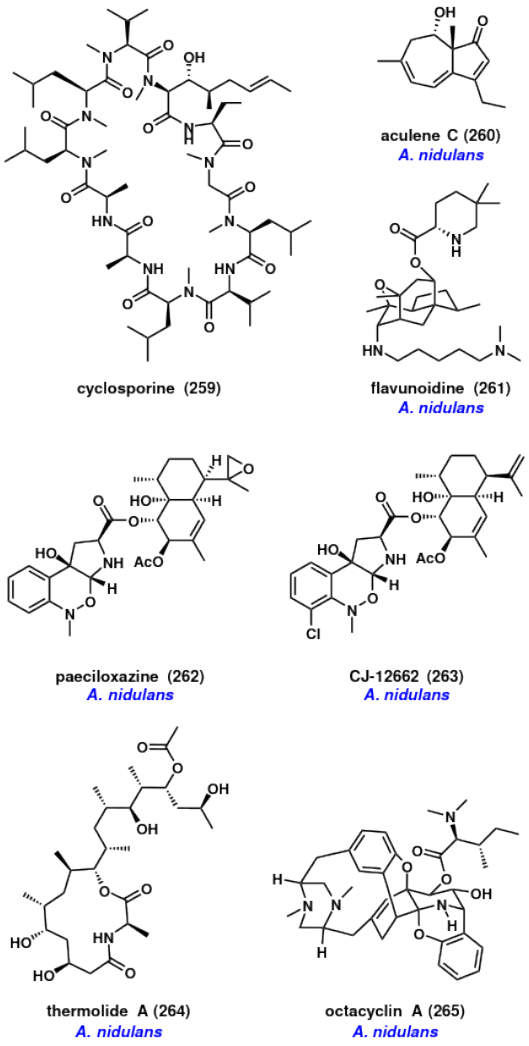


Figure 55. Examples of natural products biosynthesized by multiple core enzymes.

Heterologous expression results have revised biosynthetic pathways proposed from knockout studies in the native host. This is exemplified by reassigning the role of flavin-dependent halogenase AoiQ.[491] Based on inactivation studies and shunt metabolite analysis, this enzyme was initially predicted to catalyze chlorination of an unactivated sp^3 carbon in orthosporin (**266**) to yield demethyldichlorodiaporthin (**267**) (Figure 56).[492] Through heterologous expression in *A. nidulans*, a 1,3-diketone compound **268** synthesized by a different non-reducing PKS DiaA was shown to be the substrate of AoiQ. AoiQ catalyzes gem-dichlorination of the activated α-carbon through a canonical flavin-dependent halogenase mechanism, followed by C–C bond scission and reduction by DiaC to give **267**.

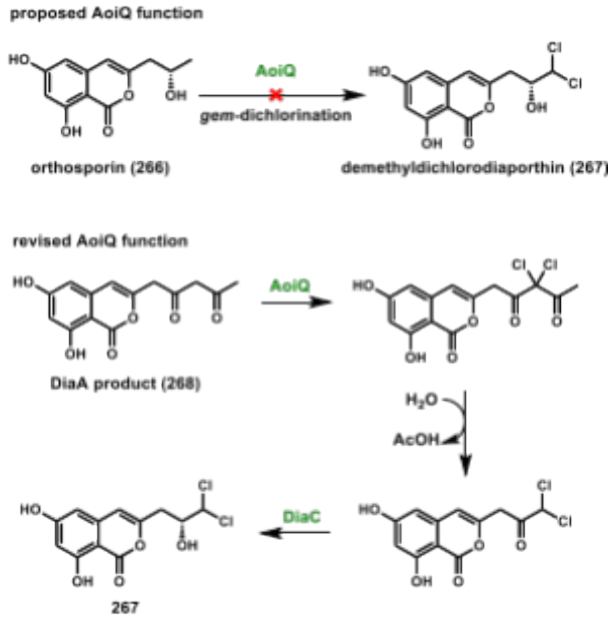


Figure 56. Revision of AoiQ function using *A. nidulans* for pathway reconstruction.

Using *A. nidulans*, BGCs with no identifiable core enzymes (we call them the unknown-unknowns) from fungi were mined and reconstituted. Yee et al. identified one large family of such BGCs that use an arginine-containing cyclodipeptide synthases (RCDPSs) to synthesize diketopiperazines with an L-arginine residue. [36] One such BGC further elaborates the diketopiperazine into a pseudopeptide natural product NK13650 A (**269**), which is a histone acetylase 300 inhibitor related to **52** (Figure 57).[493]

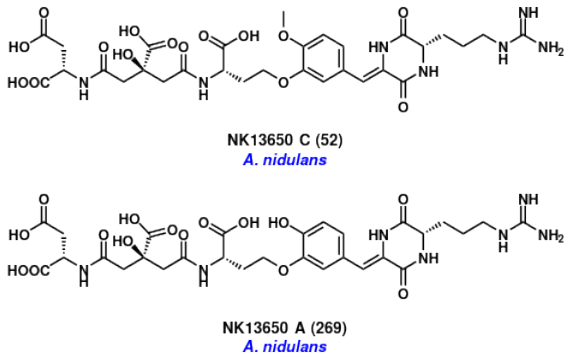


Figure 57. Representative natural products not derived from well-known core enzymes in *A. nidulans*.

A. nidulans has been used to reconstitute biosynthesis of compounds of which the targets are indicated by a self-resistance enzyme (SRE) (Figure 58). One example is the immunosuppressant (-)-FR901483 (**28**), which is derived from a dityrosyl piperazine intermediate.[494] The target of **28** was proposed to be PPAT, which was encoded in the BGC. This was

verified in a follow-up study using biochemical and structural approaches.[76] Similarly, the BGC of the potent CYP51 inhibitor restricticin (**30**) was identified by search for CYP51-homologs in BGCs. Heterologous reconstitution yielded the acetylated analog of **30**, presumably due to response of the *A. nidulans* host to counter the bioactivities.[79] The enzymatic target of the well-known natural product harzianic acid (**27**) was elucidated following full reconstitution in *A. nidulans*, during which an encoded acetolactate synthase in the BGC was reassigned to be the SRE. Acetolactate synthase catalyzes the first step in branched chain amino acid biosynthesis and is a highly pursued target for herbicide and antifungal discovery due to its absence in human and plants.[495] **27** was shown through X-ray crystallography to inhibit fungal acetolactate synthase in a unique mode of binding that is not affected by known resistance mutations. In a recent example from the biotechnology industry, an SRE-guided approach followed by heterologous expression in *A. nidulans* led to the rediscovery of roseopurpurin C (**33**), an inhibitor of human cyclin-dependent kinase 2. This compound was shown to induce cell cycle arrest at the G1 phase in human cells, making it a potential starting point for anticancer therapeutic development.[82] The biosynthesis of compound cerulenin (**270**), the first reported inhibitor of fatty acid synthase, was heterologously reconstituted in *A. nidulans* LO8030 after a SRE-guided genome mining approach was used to identify the corresponding *cer* BGC.[496] This led to the identification of *cerF*, a colocalized copy of fatty acid synthase subunit α, with the HRPKS *cerA*. Heterologous expression validated the *cer* BGC as responsible for the production of **270**. As the last example in this section, the biosynthesis of substituted pyridones harzianopyridone (**271**) and atpenin B (**272**) was probed. **271**, **272**, and related compounds such as atpenin A5 are nanomolar inhibitors of mitochondrial complex II. The biosynthetic pathways were reconstituted in *A. nidulans*, aided by the identification of the SRE, succinate dehydrogenase, in the atpenin BGC.[497]

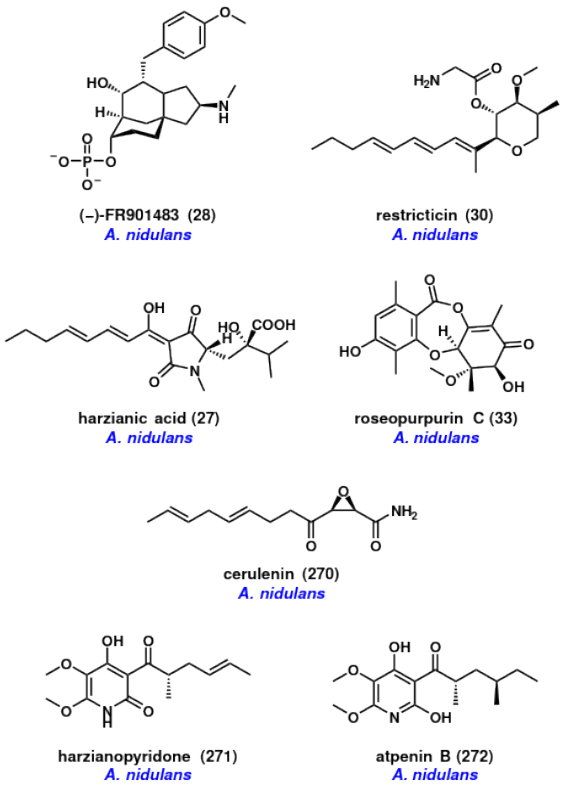


Figure 58. *A. nidulans* as a host for reconstituting BGCs contain an SRE.

Genome mining in *A. nidulans* can be useful in identification of secondary metabolites involved in fungal pathogenesis of animals and plants. For example, transcriptomics analyses have shown that BGCs anchored by polyketides are highly expressed during the biotropic and necrotropic phases of plant infection by devastating fungal pathogens.[498–500] Direct refactoring of these BGCs can therefore lead to elucidation of associated metabolites. For example, the *sth* BGC in the wheat pathogen *Parastagonospora nodorum* was found to be highly upregulated during plant infection, and showed homology to the *bet* BGC in *Phoma betae Fr.*, a fungus that causes disease in sugar beet plants.[501] Heterologous reconstitution of the *sth* cluster in *A. nidulans* led to identification of stemphyloxin II (**273**) (Figure 59).[502] In another example, a PKS-encoding BGC overexpressed by the wheat pathogen *Zymoseptoria tritici* is well conserved in other plant-interacting fungi. The heterologous reconstitution of a homologous cluster from *Trichoderma* sp. yielded a polyketide derived treconorin (**274**), that was esterified to the 4'-OH of glucose. The polyketide portion is a 5,7-bicyclic structure that resembles a terpene product. Stepwise reconstitution and biochemical studies in *A. nidulans* revealed a likely cyclization pathway for the polyketide, as well as the original of the glucose ester. The polyketide chain is offloaded from the PKS by the disaccharide trehalose, followed by α-glucosidase cleavage to give the glucose ester. In parallel, esterification of polyketide

to the disaccharide glucinol was discovered through heterologous production of luteodienoside A (**275**) in *A. nidulans*.[503] The *ltb* BGC was originally found in an endemic Australian fungus *Aspergillus luteorubrus* MST-FP2246.

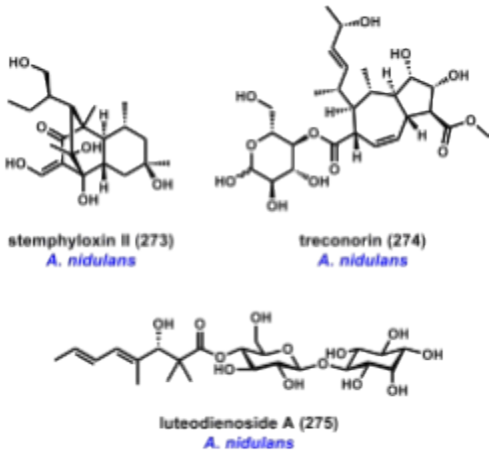


Figure 59. BGCs from phytopathogenic fungi expression in engineered *A. nidulans* resulted in new natural products.

BGCs often encode transporter proteins such as Major facilitator superfamily (MFS) transporters. A recent review described these transporter proteins and their roles in improved production of β-lactam antibiotics penicillin and cephalosporin C.[504] In *Penicillium chrysogenum*, a native penicillin producer, the final two enzymes in the penicillin biosynthetic pathway (phenylacetyl-CoA ligase, PCL; and acyl-CoA:6-amino penicillanic acid acyltransferase, AT) are localized within the peroxisome.[505,506] Two MFS transporters, PenM and PaaT, have been shown to be involved in transport of isopenicillin N and phenylacetic acid into peroxisomes. In addition to MFS transporters, targeted expression of enzymes can improve heterologous natural product titers. For example, heterologous production of penicillin in *A. nidulans* was improved by targeting the expression of two upstream biosynthetic enzymes, ACV synthetase (AcvA) and isopenicillin N synthetase (IpnA), to peroxisomes (Figure 60). Increasing the number of peroxisomes further enhanced penicillin G (**96**) production.[507] Such efforts have validated the idea of “spatial engineering” and show the benefits of both transporter proteins and targeting enzymes to specific organelles.[504]

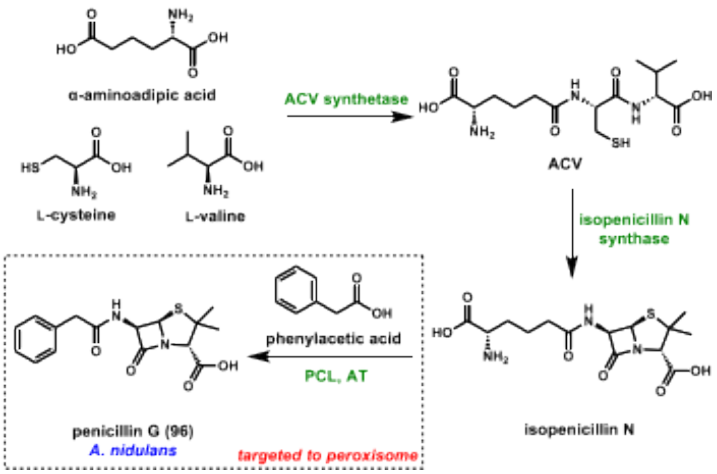


Figure 60. Biosynthesis of **96** in engineered *A. nidulans* with an emphasis on compartmentalization.

3.4.2.3 Enzyme Discovery from *A. nidulans* reconstitution

Heterologous biosynthesis, in combination with biochemical assays and structural studies, have led to discoveries of new enzymes from fungal biosynthetic pathways. A group of natural products that has been studied extensively using *A. nidulans* are those derived from pericyclic reactions. The reconstituted biosynthesis of natural products such as Sch210972 (**276**), myceliothermophin A (**277**), leporin B (**278**), pyridoxatin (**279**), ilicicolin H (**280**), pyrrocidine A (**281**), fischerin (**282**), (+)-premalbrancheamide (**283**), etc. have yielded new enzymes collectively known as pericyclases that catalyze various pericyclic reactions (Figure 61).[508–515] Using pericyclases as beacons, genome mining efforts have led to further discoveries of new natural products, including varicidin that is formed via an inverse-electron demand Diels-Alderase.[516]

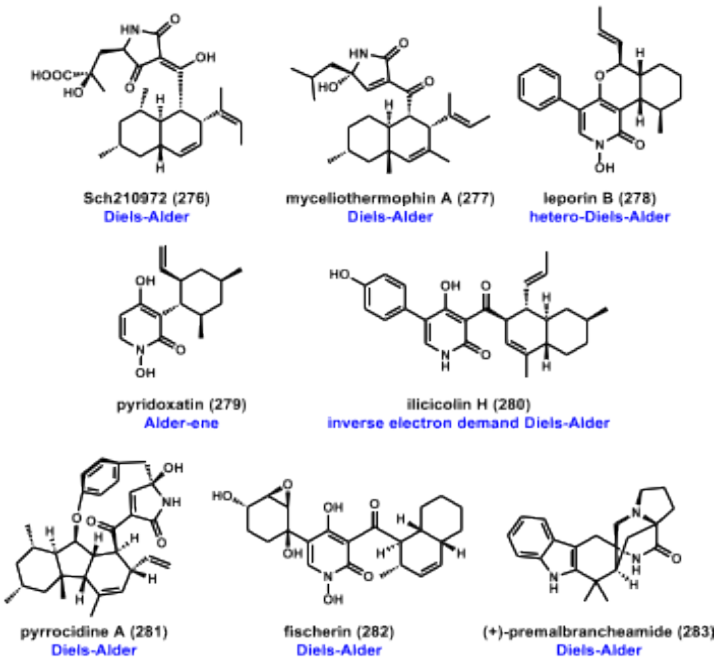


Figure 61. Fungal natural products of which biosynthesis involves a pericyclase-catalyzed reaction.

Three examples of new pericyclases emerged from *A. nidulans* expression are the norbornene-forming SdnG, multifunctional PfB/BruB and those involved in intermolecular [4+2] cycloaddition (Figure 62). SdnG is an annotated hypothetical protein of unknown function found in the sordarin (**284**) BGC. Analysis of products from the heterologous expression in *A. nidulans* without SdnG suggested the Diels-Alder reaction that forms the norbornene core requires enzyme acceleration to avoid shunt product formation. Coexpression of SdnG eliminated the shunt products and led to more efficient **284** biosynthesis. The role of the enzyme was confirmed by *in vitro* biochemical analysis and X-ray crystallography. Interestingly, SdnG is proposed to utilize iminium catalysis, by forming an iminium adduct between an active site lysine and the aldehyde-containing substrate, to accelerate the Diels-Alder reaction.[517,518]

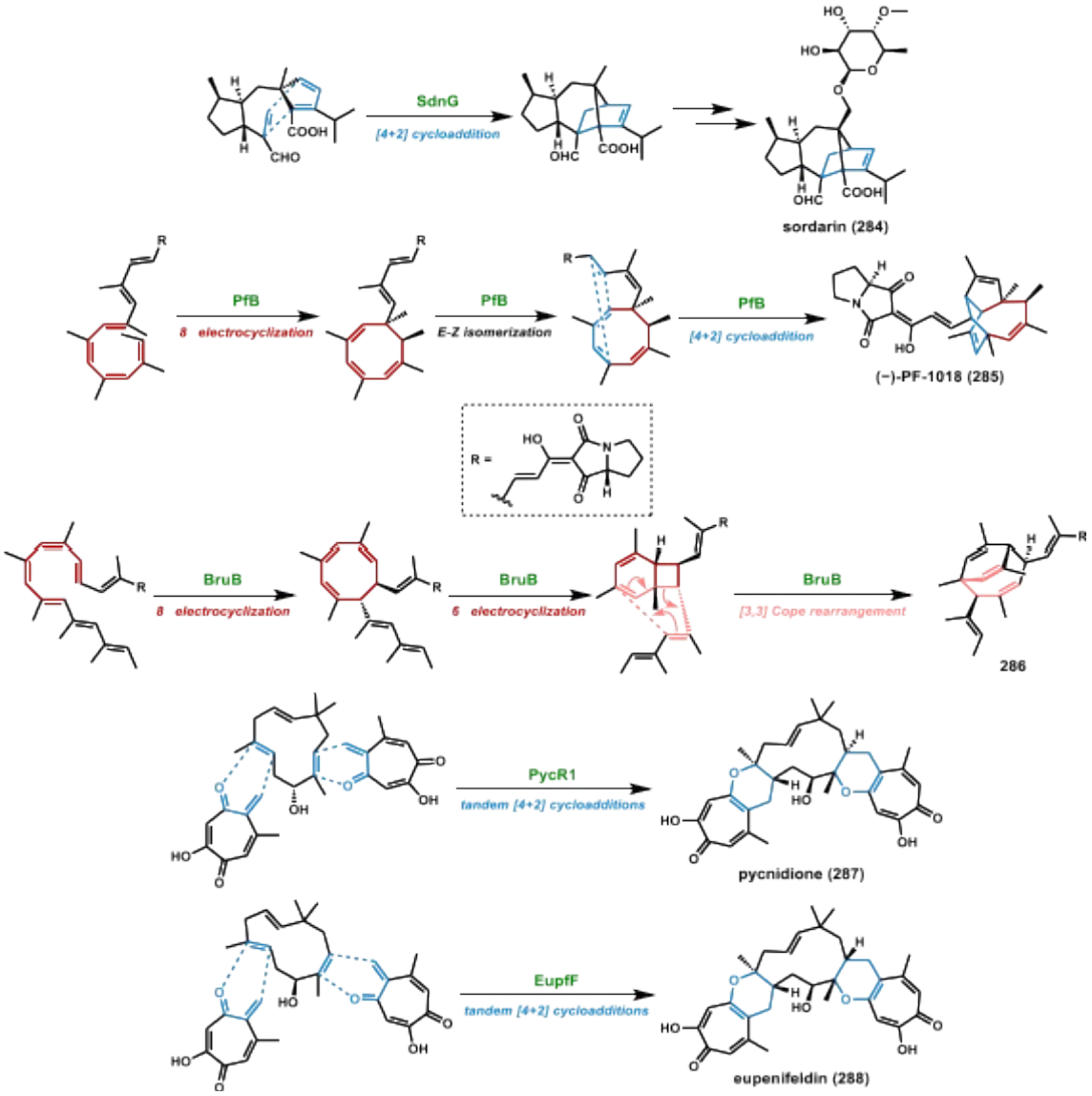


Figure 62. Examples of pericyclic reactions in natural product biosynthesis that have been identified and characterized using heterologous expression in *A. nidulans*.

Heterologous expression in *A. nidulans* was key to the characterization of two multifunctional enzymes, PfB and BruB, from two homologous, compact BGCs that encode a PKS-NRPS megasynthetase. PfB and BruB were shown to catalyze regioselective and stereoselective electrocyclization of polyene products from the respective PKS-NRPS, followed by either a [4+2] cycloaddition to give (-)-PF-1018 (**285**), or [3,3]-Cope rearrangement to give **286**, respectively (Figure 62).[519] A set of Diels-Alderases catalyzing tandem intermolecular [4+2]-cycloadditions were identified and characterized from the biosynthetic pathways of pycnidione (**287**) and eupenifeldin (**288**). These enzymes, named EupfF and PycR1 respectively,

were characterized through crystallization, mutagenesis, and computational studies. Heterologous expression of each BGC in *A. nidulans* afforded production of the respective final products.[520]

The PLP-dependent family of enzymes has emerged from heterologous reconstitution studies in *A. nidulans* to play a major role in scaffold building during fungal natural product biosynthesis. BGC analysis has revealed a group of PLP-dependent enzymes with sequence homology to cystathionine-γ-synthase, which perform γ-replacement using *O*-succinyl-homoserine (OAH) with cysteine thiol to form cystathionine.[521] These homologs have been shown to catalyze diverse γ-replacement reactions, including C–C, C–O, and C–N bond formations (Figure 63).[36,486,522–524] The prototypical C–C bond forming enzyme is CndF from the citrinadin BGC. CndF was reconstituted in *A. nidulans* to catalyze the reaction between acetoacetate and OAH to form a new C-C bond and the cyclic imine **289.** Stereoselective reduction by an imine reductase yields the amino acid (2S, 6S)-6-methyl-pipecolic acid (**290**) that is incorporated by the citrinadin NRPS into the natural product as the methylpiperidine moiety. CndF was shown to have broad substrate scope towards β-ketone nucleophiles, and *A. nidulans* was used as a host for biosynthesis of a series of substituted pipecolate building blocks. A similar C-C bond forming PLP-dependent enzyme was found in the biosynthetic pathway of fusaric acid (**291**), a well-studied mycotoxin.[525] Fub7, together with an imine-reductase, forms 5-substituted pipecolates using a polyketide-derived aldehyde (**292a**) as the carbon nucleophile. In parallel, FlvA from the BGC for flavunoidine (**261**) was reconstituted in *A. nidulans* to synthesize 5,5-dimethyl-pipecolic acid (**293**), which is esterified to the terpene core by a monomodule NRPS.[486] Recently, the biosynthetic pathway of fusarilin A (**294**) was reconstituted to reveal a PLP-dependent enzyme (FusB) that can perform an additional aldol condensation following the γ-replacement using an enamine (**292b**) as a nucleophile, in essence a net (3+2) annulation, to afford the cycloleucine (**295**) that serves as the core in **294**.

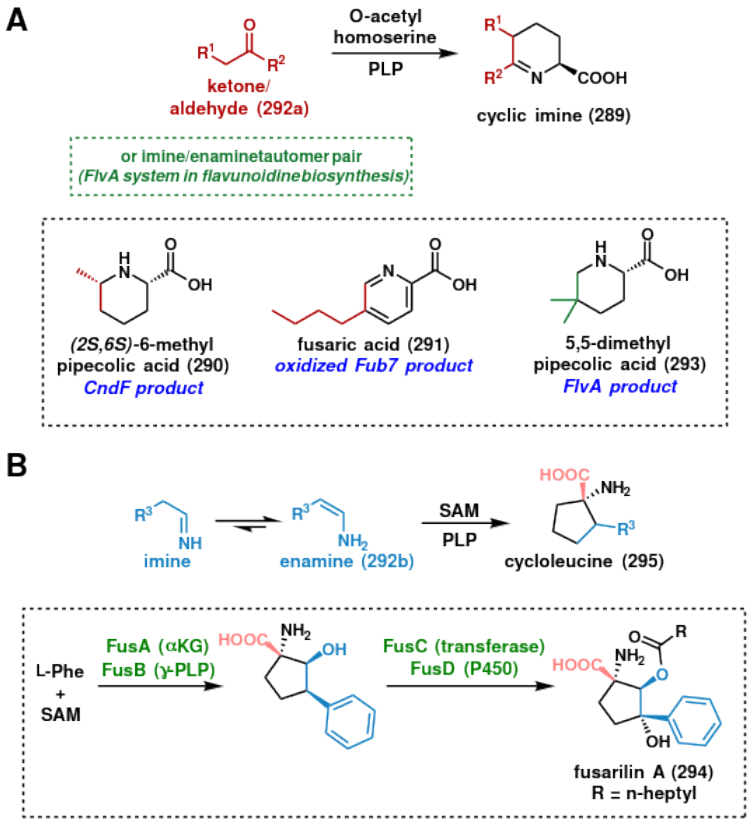


Figure 63. PLP-dependent enzymes as scaffold building enzymes in fungal natural product biosynthesis. (A) PLP-dependent pipecolic and picolinic acid formation. (B) Cycloleucine biosynthesis investigated via heterologous expression in *A. nidulans.*

Other recent examples of enzyme discovery through *A. nidulans* are listed below (Figure 64). These include: i) Reconstitution of the BGC of cytochalasins, such as aspergillin PZ (**296**) in *A. nidulans,* identified a berberine bridge enzyme (BBE)-like oxidase AspoA that catalyzes an unusual olefin isomerization.[526] This is the first example of successful heterologous reconstitution of the cytochalasin structure; ii) A P450 enzyme, CtdY, was shown to catalyze amide bond cleavage, which led to the formation of an iminium cation during maturation of the 2,5-diazabicyclo[2.2.2]octane system found in 21R-citrinadin A (**297**);[527] iii) In combination with resistance-guided genome mining and heterologous expression in *A. nidulans* LO8030, a P450 enzyme HknE was shown to be responsible for the formation of a cyclopentenone moiety during conversion from neosartoricin B (**298**) to hancockinone A (**299**), a complex metabolite consisting of a prenylated 6/6/6/5 tetracyclic skeleton;[528] iv) Expression of the *ftchy* cluster from *Fusarium tricinctum* in *A. nidulans* provided insight into the formation of the 4(3*H*)-quinazoline scaffold of chrysogine (**300**).[529] This 4(3*H*)-quinazoline scaffold is present in over 200 alkaloid natural products. A two-module NRPS, ftChyA, was shown to be involved in chrysogine biosynthesis through the

formation of a linear γ-L-glutamyl-L-alanyl-anthranilate tripeptide; v) heterologous expression of the flavipucine BGC in *A. nidulans* resulted in production of both flavipucine (**301**) and dihydroisoflavipucine. Six additional flavipucine products could be isolated and identified, including one containing the unusual 1,3-dihydro-2-azepinone structure;[530] vi) heterologous expression of the phomactin diterpenoids was shown to require only two enzymes: a cyptic terpene cyclase to convert GGPP into phomactatriene (see Figure 10), and a P450 monooxygenase to catalyze oxidation at multiple sites along the phomactatriene scaffold to yield phomactin B4 (**40**);[103] and vii) heterologous expression of the *fun* BGC from *F. tricinctum* CGMCC 3.4731 identified an ABBA-type prenyltransferase, FunA, that was shown to be the first prenyltransferase to accept free L-histidine as a substrate in the biosynthesis of fungerin (**302**);[531] and viii) the biosynthesis of (–)-vinigrol (**303**).[532]

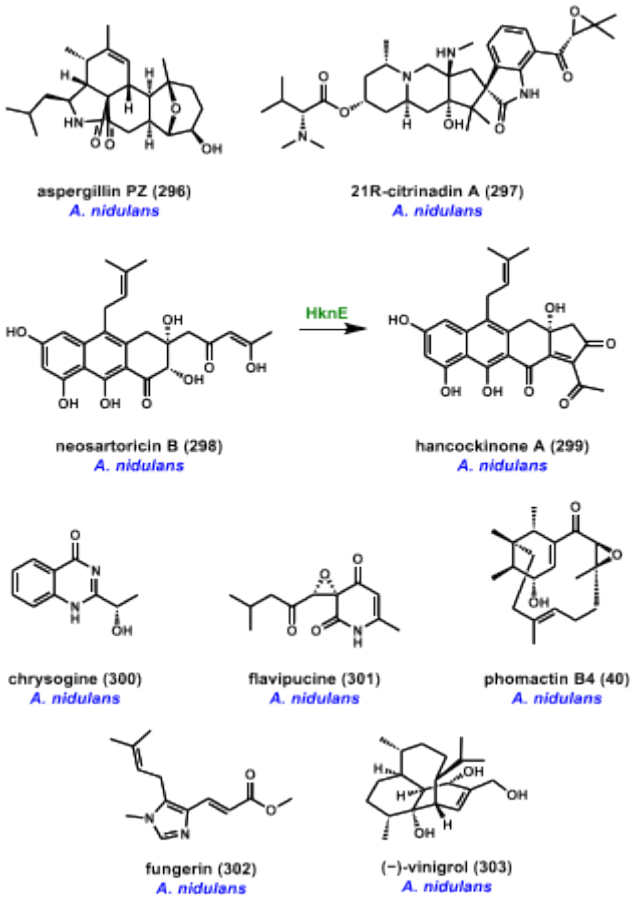


Figure 64. Representative natural products heterologously biosynthesized in *A. nidulans*.

3.4.3 *Aspergillus oryzae* as a heterologous host

A. oryzae is an important strain used in the fermentative production of soy sauce, sake and other East Asian food. The safety profile of *A. oryzae* and rapid growth characteristics make this strain an attractive host for heterologous biosynthesis. Two widely used laboratory strains for use as heterologous hosts are *A. oryzae* M-2-3, which is auxotrophic for arginine (knockout of *argB*); and *A. oryzae* NSAR1, a quadruple auxotroph (*argB-, niaD-, sC-,* and *adeA*-).[533,534] Additionally, deletion of endogenous BGCs such as that responsible for biosynthesis of kojic acid has resulted in cleaner background metabolic profiles.[535]

A. oryzae has been extensively used for the heterologous reconsitution of fungal biosynthetic pathways, as documented in the review by Oikawa.[536] A pionneering example of heterologous expression of a complete fungal BGC was the expression of the citrinin (**304**) BGC in *A. oryzae* by Sakai and coworkers (Figure 65).[537] This is followed by reconstitution of the tenellin (**305**) BGC in *A. oryzae* M-2-3, in which the titer was five times higher than that from the native producer. Around the same time, biosynthesis of the meroterpenoid pyripyropene A (**306**) from *Aspergillus fumigatus* was elucidated through heterologous expression of the BGC in *A. oryzae* M-2-3.[538] Reconstitution of this pathway established the biosynthetic logic of meroterpenoid biosynthesis in fungi, including starter unit for the PKS, prenylation of the PKS product, cyclization of the terpene fragment and oxidative modification of the pathway intermediates. *A. oryzae* has since been used extensively for the reconstitution of nearly all known meroterpenoid BGCs in fungi, including a recent study into the biosynthesis of austalide F (**307**).[539] Readers can find a comprehensive review by Abe and coworkers documenting these successes.[540] In parallel, indole diterpene natural products, which are formed through the prenylation of the indole moiety, followed by oxidative cyclization and rearrangement, have also been widely produced in *A. oryzae.* This is exemplified by the elucidation of the paxilline (**16**) biosynthetic pathway by Oikawa and coworkers, as well as the recent elucidation of the biosynthesis of lolitrem B (**308**) by the same group.[541,542] A review by Oikawa et al. describes additional examples of pathway elucidation for fungal indole diterpene natural products.[543] Other fungal alkaloids that have been produced in *A. oryzae* include ardeemin FQ (**309**) and (±)-penindolene A (**310**).[544,545]

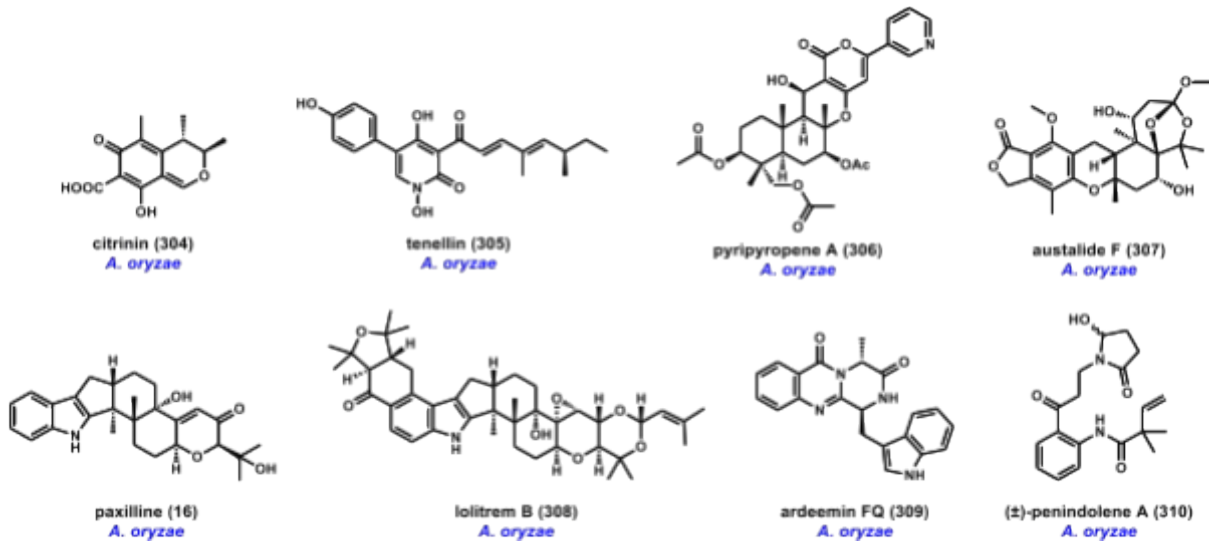


Figure 65. Representative natural products produced in engineered *A. oryzae*.

Just as in *A. nidulans,* selected examples of *A. oryzae* as a heterologous host have been chosen for this review (Figure 66). These include: i) heterologous reconstitution of the *cal* BGC associated with the biosynthesis of calbistrin A (**311**) showed that the PKS CalA is responsible for the production of both the polyene and the decalin-containing portions of the molecule. CalA was therefore shown to be an interesting dual-functional PKS.[546] Similarly, the biosynthesis of gregatin A (**312**), which contains an alkylated furanone core, was shown to involve the fusion of two polyketide fragments generated by a single PKS and *trans*-enoylreductase through heterologous expression;[547] ii) the biosynthesis of byssochlamic acid (**313**) and agnestadride A, two fungal polyketide-derived maleidrides, was shown to involve two ketosteroid isomerase-like and two phosphatidylethanolamine-binding protein-like enzymes;[548] iii) the biosynthetic pathways for chrysophanol (**314**), skyrin (**315**), and rugulosin A (**316**), three bioactive fungal anthraquinones/bisanthroquinones, were elucidated from *A. oryzae*;[549,550] iv) the biosynthesis of stachysalicyloid C (**317**) was shown to involve a ketoreductase-domain-mediated polyketide chain release;[322] v) an unusual P450 enzyme, FsoE, catalyzing C-C bond cleavage was characterized during fuscoatroside (**318**) biosynthesis through expression in *A. oryzae;* [551] vi) the BGC of the diterpene pleuromutilin (**97**) was expressed in *A. oryzae*, representing the first basidiomycete BGC to be expressed in an ascomycete host.[237] **97** is the semisynthetic precursor to approved antibiotics and the heterologous production may provide a high-titer source; vii) heterologous expression uncovered the biosynthetic pathway of the furanosteroid demethoxyviridin (**319**), which was shown to utilize a Baeyer-Villiger monooxygenase, an esterase, and a dehydrogenase in order to cleave a pregnane side chain, rather than a proposed P450-mediated pathway;[552] viii) heterologous expression in *A. oryzae* NSAR1 was used to

identify a set of "domainless" enzymes that were characterized to be terpene cyclases, leading to the discovery of fungal onoceroid triterpenoids such as moserinol (**320**);[553] ix) the Oikawa and Minami groups have studied DUF3328 enzymes (belonging to the UstYa family from the BGC for ustiloxin B) in the context of fungal RiPP biosynthesis.[554] Specifically, DUF3328 enzymes such as CctP2, O, and R in hydroxycyclochloritine (**321**) biosynthesis, and AprY in asperipin-2a (**322**) biosynthesis have been functionally expressed in *A. oryzae*;[555–557] and x) to increase the throughput of screening cryptic BGCs, automation of the heterologous expression workflow was established. To demonstrate the potential of this approach, different combinations of predicted terpene cyclases and modification enzymes in the genomes of several fungal strains were cloned into a plasmid library and heterologously expressed in *A. oryzae*. 185 distinct terpenoids were discovered from 26 of these clusters.[558]

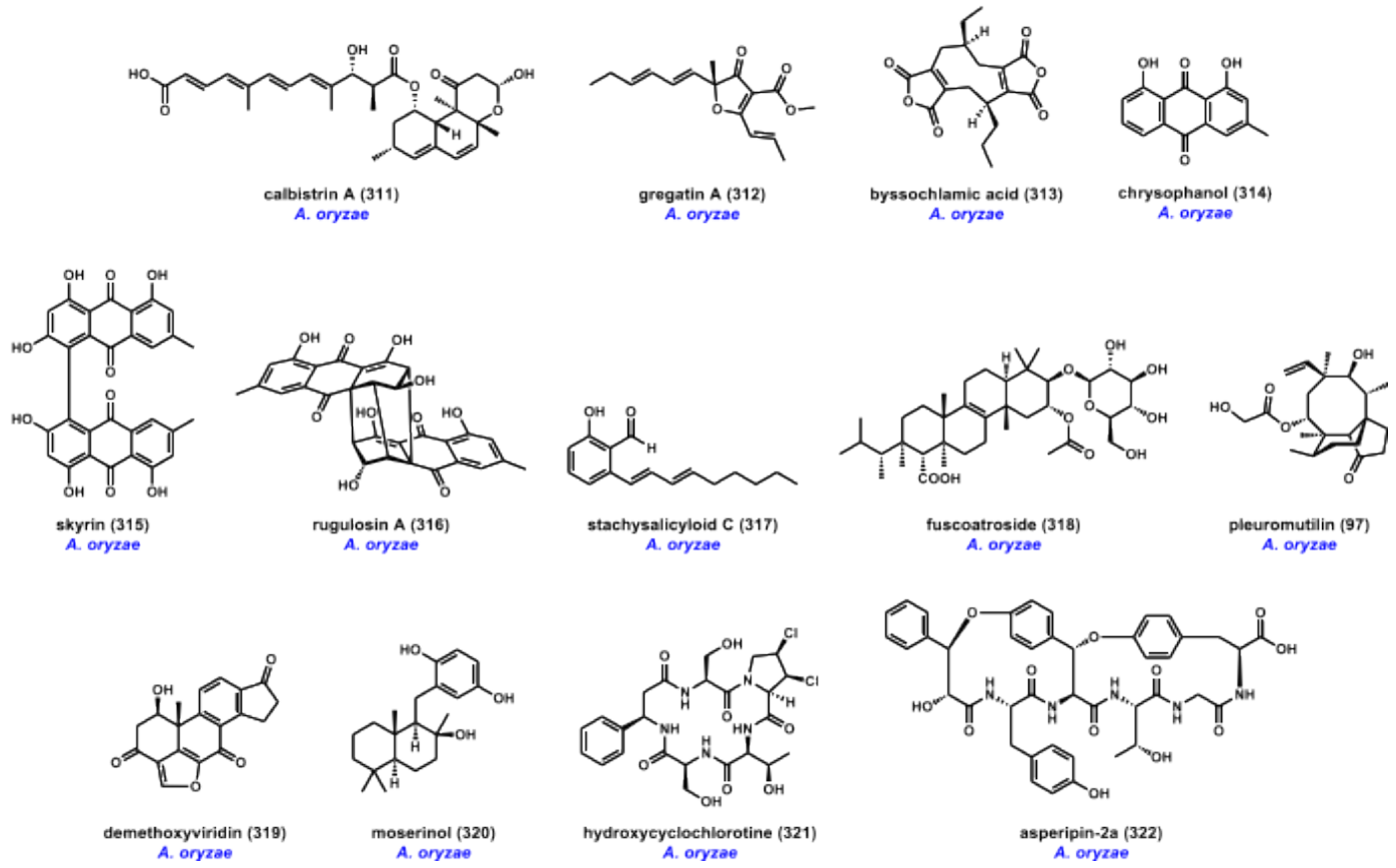


Figure 66. Representative natural products produced in engineered *A. oryzae*.

3.4.4 *Aspergillus niger* as a heterologous host

Aspergillus niger is an important fungal host in microbial biotechnology, and is used in the industrial production of citric acid and gluconic acid.[559,560] Known as the "black mold", it has been classified as a GRAS organism for fermentation.[561]

The *A. niger* genome shows considerable biosynthetic potential based on BGC predictions.[562,563] One of the first examples of heterologous expression of a complete fungal biosynthetic pathway is the reconstitution of the penicillin BGC from *Penicillium chrysogenum* in *A. niger*. This work was instrumental in showing that the genes responsible for the production of penicillin constituted a fungal BGC.[238] Since then, *A. niger* has had a less prominent role as a host for natural product biosynthesis compared to *A. nidulans* and *A. oryzae*.[562,563] *A. niger* has been used to heterologously produce compounds such as fungal cyclodepsipeptides (Figure 67). The 351-kDa NRPS ESYN was expressed in *A. niger,* and gram-scale production of enniatin A (**323**) was achieved after engineering the host to utilize intracellular α-ketovaleric acid.[564] Other cyclodepsipeptides such as beauvericin (**86**) and bassianolide (**324**) were also heterologously produced in *A. niger*.[565] A recent effort to develop *A. niger* as a heterologous host led to deletion of the *akuB* gene from *A. niger* to yield a NHEJ-deficient strain.[566] Targeted gene editing techniques using CRISPR-Cas9 have been developed for multiplexed gene editing experiments, expanding the strain engineering potential for use of *A. niger* as a heterologous host.[567] In cases where screening of several organisms to identify an optimal heterologous host is desired, tools such as DIVERSIFY can be used to create a single gene expression cassette that can then be simultaneously expressed in several species. A proof-of-concept study was conducted by expressing a three-gene cassette in *A. nidulans, A. oryzae, A. niger,* and *A. aculeatus*.[568]

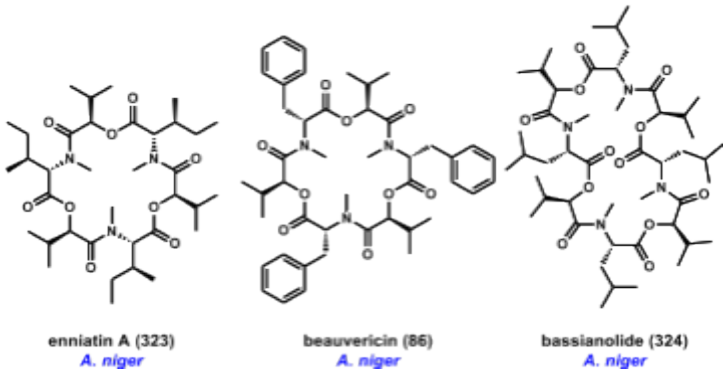


Figure 67. Representative natural products biosynthesized from *A. niger*.

3.4.5 Other Fungi as Heterologous Hosts

Other fungal species such as *Penicillium rubens* and *Trichoderma reesei* have been used as heterologous hosts for secondary metabolite production. An *Agrobacterium tumefaciens*-mediated transformation (ATMT) system was developed for *P. rubens.*[569] In addition, *P. rubens* was engineered to inactivate the BGCs of several metabolites, including chrysogine (**300**) and penicillin G (**96**), to achieve a cleaner metabolic background.[570] This strain was then used to heterologously

produce decumbenones and related polyketide calbistrin A (**311**). Another *Penicillium* strain, *P. paxilli,* has been used in heterologous reconstitution experiments as well.[571–573] *T. reesei* is used to produce heterologous proteins such as cellulases at high levels.[574,575] Recent explorations have shown that *T. reesei* holds potential as a host for heterologous production of secondary metabolites from waste products.[576] Additionally, members of the genus *Fusarium*, such as *F. heterosporum*, have been utilized for the heterologous expression of compounds such as the tetramic acid-containing Sch210972 (**276**).[577]

3.5 Heterologous Expression in Plants

Reconstituting a plant biosynthetic pathway into a microbial host and achieving high titers is often challenging due to differences in cellular physiology between the native and heterologous organisms. A more direct approach is to use model plant hosts for heterologous expression. The importance of plants as heterologous hosts for natural products, as well as recent biotechnological developments to improve their genetic tractability and feasibility for industrial purposes, can be found in two insightful reviews.[578,579]

The most commonly used plant for heterologous expression is *Nicotiana benthamiana,* due to its ease of cultivation and genetic tractability.[580] A suite of tools has been developed to facilitate both stable and transient transformation of *N. benthamiana* plants (Figure 68). While transgenic expression offers a stable plant cell line that is useful for long term and permanent expression of the pathway, constructing the transgenic species remains challenging and time consuming. A powerful approach that has emerged is the *Agrobacterium*-mediated transient gene expression in *N. benthamiana*. This technique was previously developed for the rapid and high-level production of proteins in plants.[581] The potential of this host for plant metabolite biosynthesis was shortly realized thereafter. Pathway genes of interest are cloned into T-DNA of plasmids and transformed into *A. tumefaciens.* Plant leaves are infiltrated with a suspension of *A. tumefaciens* transformants. Following entry into the intercellular space under pressure, the bacteria can transfer the T-DNA into the plant cells, which results in expression of the heterologous genes. The plants are then grown under greenhouse conditions for several days, followed by leaf harvest and extraction of organic metabolites for analysis. Pathway precursors and intermediates can also be infiltrated into the plant leaves, making the plant a host for biotransformation. This method has been widely adopted in the plant natural product community, and has led to significant breakthroughs in the elucidation, reconstitution and engineering of a number of high-profile plant natural products. Indeed, combined with modern plant transcriptomics and coexpression analysis, multiple gene candidates can be rapidly introduced into *N. benthamiana* to test

proposed roles in biosynthesis. This approach can also be used for combinatorial biosynthesis or preparative synthesis.[582] The latter was demonstrated by Osbourne and coworkers in the biosynthesis of different amyrin-derived triterpenes from *N. benthamiana* at as high as ~4 mg/g dry leaf, with each gram of dry leaf corresponding to approximately 2-3 plants.[583]

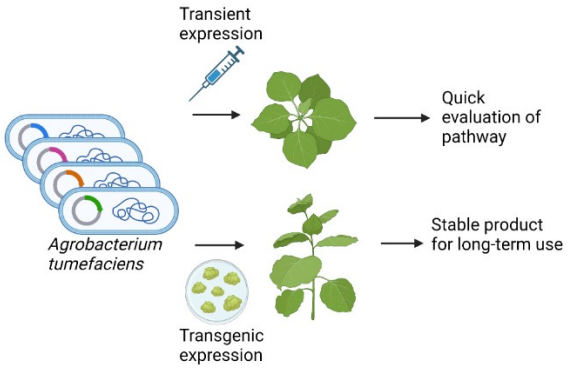


Figure 68. Overview of *Agrobacterium*-mediated transient expression of heterologous genes in *N. benthamiana*.

Created in BioRender. Dror, M. (2024) https://BioRender.com/c51d443

A milestone demonstration of the combination of chemical hypothesis, transcriptomics and metabolomics using *N. benthamiana* as a host is the biosynthetic reconstitution of (-)-podophyllotoxin (**11**), a phenylpropanoid natural product found in the mayapple plant.[230] **11** is also the semisynthetic precursor to etoposide, which is an inhibitor of DNA polymerase II and an important anticancer drug. Sattely and coworkers selected 29 candidate genes and transiently expressed them combinatorially in *N. benthamiana* followed by mass spectrometry analysis (Figure 69). The lab demonstrated 10 genes were sufficient for the biosynthesis of 4'-desmethyl-epipodophyllotoxin (**325**), which is a more immediate precursor to etoposide semisynthesis. Among the genes identified is a nonheme, α-ketoglutarate-dependent oxygenase that catalyzes oxidative C-C bond formation that converts the intermediate (-)-yatein (**326**) to the fused 5-6-6-5 tetracyclic framework in (-)-deoxypodophyllotoxin (**327**) that is the characteristic core of podophyllotoxins and etoposide. This enzyme has been used by the Renata group in the chemoenzymatic synthesis of **11**.[584] The Sattely lab subsequently optimized the host by increasing flux to coniferyl alcohol (**328**), the precursor to the pathway.[585] By expressing a total of 16 genes in tobacco, up to 4.3 mg/g dry plant weight **325** was accumulated.

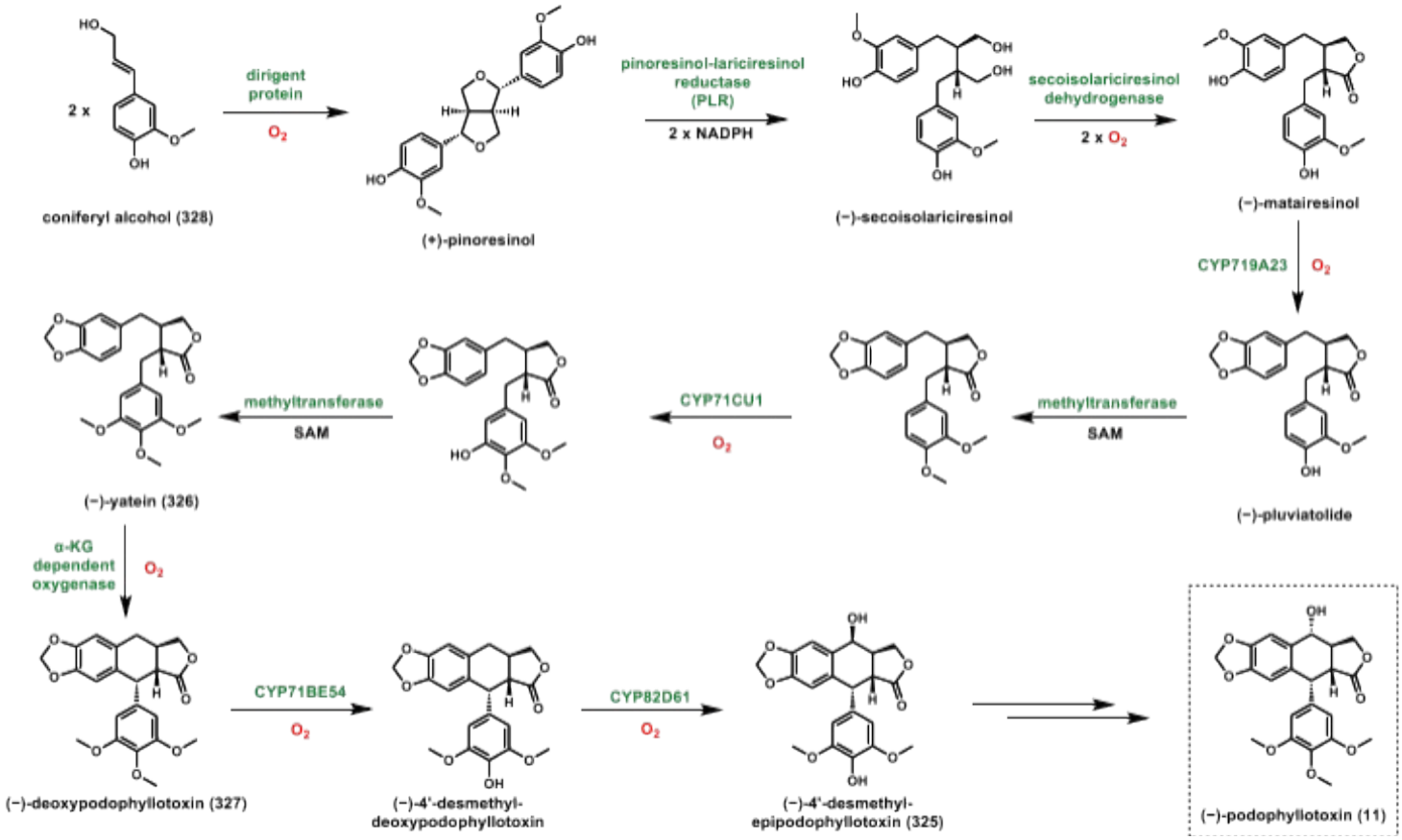


Figure 69. Biosynthetic pathway of **11**, identified using heterologous expression in *N. benthamiana.*

An additional important phenylpropanoid-derived plant natural product reconstituted using *N. benthamiana* is colchicine (**91**) from autumn crocus (Figure 70).[229] **91** is an inhibitor of microtubule assembly and has been used in traditional medicines for treating inflammatory disease. The most interesting structural feature is the 6-7-7 ring system, in which the central cycloheptene ring is generated through an oxidative coupling of phenyl rings derived from L-tyrosine and L-phenylalanine, via the known (*S*)-autumnaline intermediate (**329**). Nett and Sattely used the combination approaches outlined above to show 16 genes are sufficient to arrive at *N*-formyldemecolcine (**330**) in the first report, including the P450 enzyme CYP75A110 that catalyzes the aforementioned aryl coupling reaction to yield isoandrocymbine (**331**). Six methyltransferases required in the pathway were identified from 11 candidates that were expressed at high levels during periods when colchicine production was turned on in the native producer. In a follow up study, the same authors identified the remaining three genes responsible for converting **330** to **91**. When combined with the previously determined pathway genes, total biosynthesis of **91** in *N. benthamiana* was accomplished.

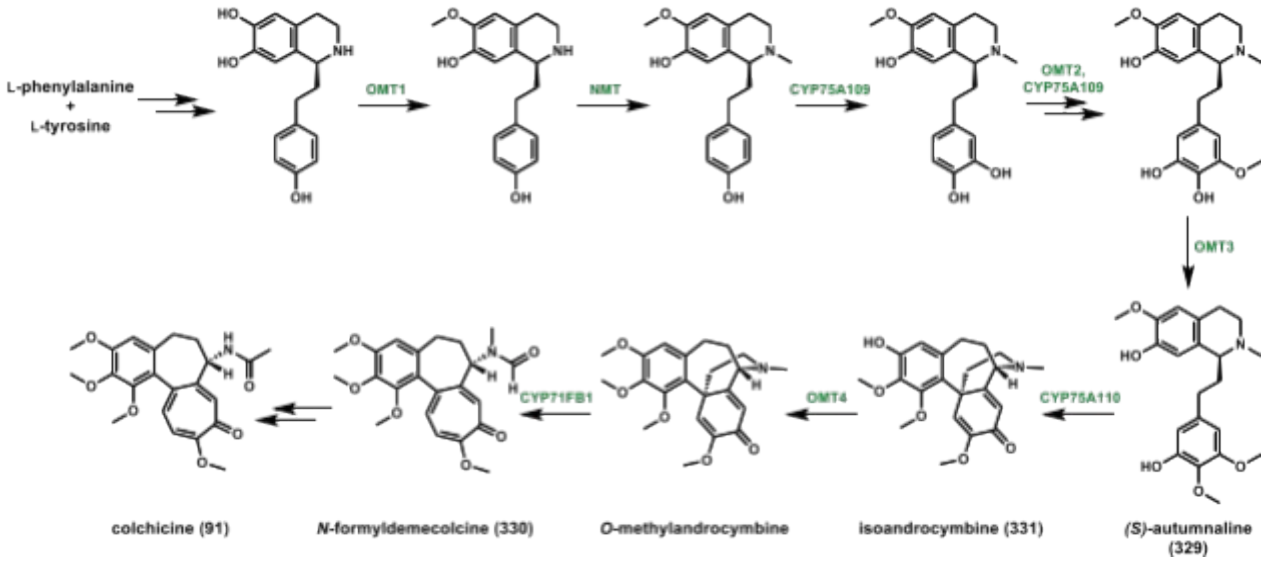


Figure 70. Biosynthetic pathway of colchicine (**91**) reconstituted in *N. benthamiana*.

In addition to phenylpropanoids, terpene-derived natural products are abundant in plants, and many have essential therapeutic values to humans. The Osbourne and Sattely groups targeted the limonoid family of triterpenes, a very large (>1500) family of compounds containing the insecticide azadirachtin (**332**) from the neem plant (Figure 71).[586] The *N. benthamiana* platform is particularly well-suited for studying triterpene biosynthetic pathways due to the availability of the 2,3-oxidosqualene (**333**) precursor used by plants for β-amyrin (**334**) biosynthesis. Osbourne and Sattely first identified the three enzymes conserved in many limonoid-producing plants that are involved in the transformation of **334** to a protoliminoid precursor, melianol (**335**).[587] Based on this information, the same group of authors significantly extended the pathway in *N. benthamiana* to two complex limonoids, azadirone (**336**) and kihadalactone A (**337**), both containing the limonoid class-defining furan structure. As a testament to the scalability of this host, many of the key intermediates in the pathway were isolated and confirmed by NMR spectroscopy. The pathway was recently updated with the identification of a carboxylesterase that markedly improved the yield of **336** in the heterologous *N. benthamiana* host through removal of the C21-acetoxy group present in **338** as a deprotection step, leading to more efficient formation of the furan ring in **336**.[588]

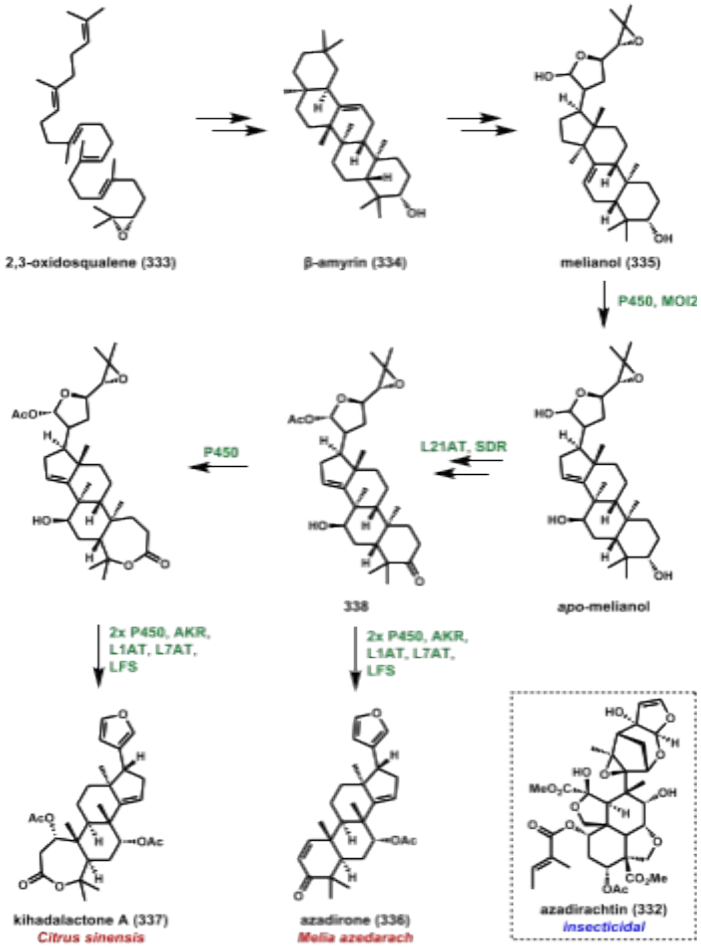


Figure 71. Biosynthetic pathway of limonoids **336** and **337** reconstituted in *N. benthamiana*.

While the yeast reconstitution of QS-21[410] (**191a-b**) was discussed above (Section 3.3.1.2), biosynthetic reconstitution in *N. benthamiana* played an indispensable role in elucidating all the genes in the pathway (Figure 72). Using methyl jasmonate as an elicitor for triterpene saponin biosynthesis in the plant *Saponaria vaccaria*, Scheller and coworkers identified numerous gene candidates from comparative transcriptomics.[589] Using both *N. benthamiana* and *S. cerevisiae* as heterologous hosts, the author confirmed the involvement of genes involved in the oxidative modification of **334** to the aglycon quillaic acid (QA, **339**), formation of the rare nucleotide sugar UDP-D-fucose, and the glycosyltransfer steps. The Osbourne group worked with the soapbark tree *Quillaja saponaria*, and used *N. benthamiana* to reconstitute the biosynthesis of QA-TriX-FRXX (**340**), a bridgehead intermediate that can be used for adjuvant pathway engineering. A total of 16 genes were necessary to complete the pathway. Furthermore, the authors identified three enzymes that can elaborate the pathway to produce QS-7 (**341**) that has desired adjuvant properties.[590] To extend the complexity of the pathway, the same lab reconstituted the complete, 20-step biosynthetic pathway to QS-21, which requires the installation

of an arabinosylated acyl chain to **340**. This information, together with further engineering to provide a *de novo* source to the acyl chain, led to the complete biosynthesis in yeast.[410]

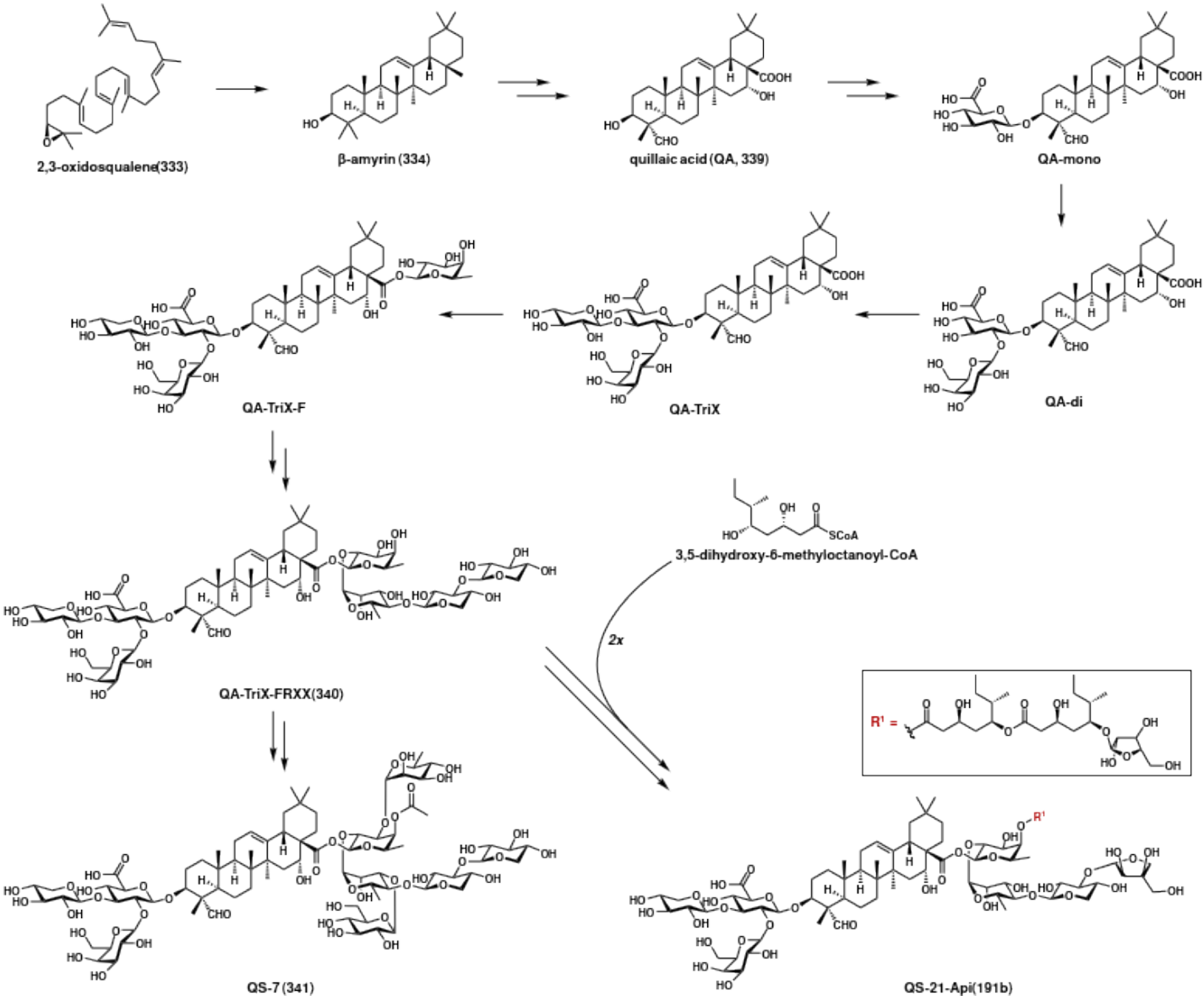


Figure 72. Biosynthetic pathways of **341** and **191b** characterized using heterologous expression in *N. benthamiana*.

The biosynthetic pathway of the anticancer diterpenoid paclitaxel (**7**) from the Taxus (Yew) trees has been intensely pursued for over thirty years, starting from the mature trees and cell cultures of *Taxus*, to microbial hosts such as *E. coli* and yeast, to culminate in recent completion in *N. benthamiana.*[591] Genes involved in the generation of the carbon core taxadiene (**90**) (taxadiene synthase), hydroxylation of the skeleton at C_5, C_{13}, C_{10}, C_2 and C_7, as well as downstream acylation steps have been characterized and reconstituted in various hosts (Figures 73 and 74).[592–597] However, key steps involved in

the formation of the highly functionalized baccatin III (**343**), including the enzymatic origin of the key oxetane feature in **7**, are missing. Functional expression of yew-originating enzymes in *N. benthamiana*, including taxadiene synthase, taxadiene C5-α-hydroxyalse (T5αH) and a reductase partner was demonstrated by Wang and coworkers.[598] Studies using GFP fusions showed T5αH is localized in the chloroplasts and ER in tobacco, respectively. Subsequent redirection of T5αH into the plastid, as well as overexpressing bottleneck enzymes in the MEP precursor pathway led to significant increases in taxane titer. Sattely and coworkers characterized the activity of T5αH in *N. benthamiana* in detail and revealed the accumulation of numerous over-oxidized products that decreased the titer of the desired product mono-hydroxylated taxadien-5α-ol (**100**). [599] The expression level of T5αH was then tuned through promoter swapping to increase the level of **100** threefold. This enabled the simultaneous expression of the first six genes in the pathway in *N. benthamiana,* which resulted in the formation of the advanced intermediate taxadiene-5α,10β-diacetoxy-13α-ol (**342**) (Figure 73). This tour-de-force analysis demonstrated the complexity of the pathway and the necessity to fine-tune the expression of early pathway genes to avoid shunt product formation, which can be detrimental to the efforts to reconstitute downstream steps.

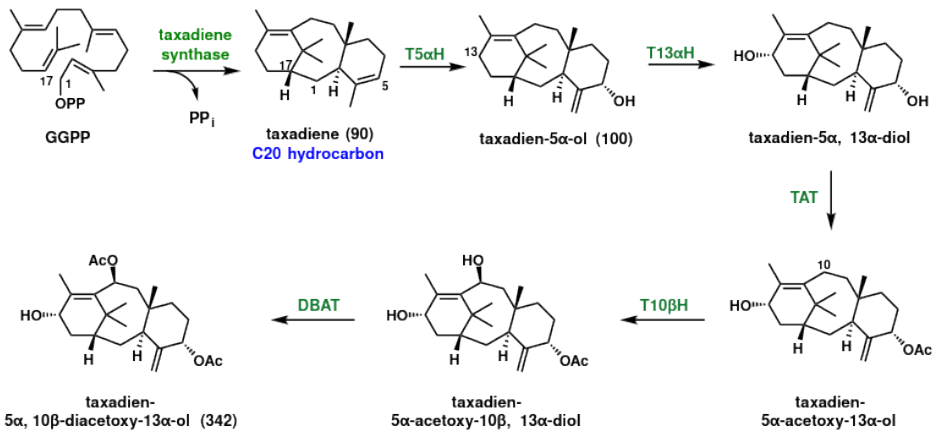


Figure 73. Early stages of the paclitaxel biosynthetic pathway as characterized by heterologous expression in *N. benthamiana.*

A key breakthrough in paclitaxel biosynthesis was the identification of the enzyme required in forming the signature oxetane ring. Three groups simultaneously reported the identification of the missing enzyme, albeit with different conclusions regarding the gene identity, timing, and mechanism of the reaction. Zhang and Fernie used *N. benthamiana* to screen enzymatic candidates responsible for C_9 hydroxylation, C_9 oxidation, C_1 hydroxylation, and oxetane formation.[600] They identified the gene candidates, including an α-ketoglurarate, nonheme iron-dependent epoxidase that

is proposed to catalyze epoxidation of the exocyclic olefin. Expression of 13 genes led to formation of ~150 ng/g plant weight of baccatin III (**343**). The same group subsequently reconstituted the biosynthesis of the phenylalanine-derived acyl chain that is transferred to C_{13}-OH. Expression of five genes together with feeding **343** to *N. benthamiana* resulted in the detection of low amounts of **7** (64 ng/g of plant material). Kampranis and coworkers used yeast and *N. benthamiana* to demonstrate that CYP725A4, the assigned T5αH, has the additional function of generating the oxetane ring. Initial expression of CYP725A4 with CPR resulted in observation of dioxygenated taxanes, in agreement with that reported by Sattely.[601] However, preparative scale isolation using the transient expression in *N. benthamiana* led to one compound that has the oxetane installed at this early stage. Biotransformation showed that endotaxadiene and exotaxoadiene can be converted to the oxetane, which led to the proposal that CYP725A4 catalyzes two consecutive epoxidation reactions *en route* to oxetane formation.

The Yan and Lei groups reached yet another conclusion by screening members of the CYP725A subfamily from *Taxus* using taxa-4(20),11-diene-2α,5α,7β,9α,10β,13α-hexaol-hexaacetate as the substrate.[602] The authors identified a new enzyme, TOT1, to be the responsible enzyme, using a combination of biotransformation, *in vitro* biochemical characterization, and computational modeling (Figure 74). The authors proposed a Fe^{III}-bound oxonium intermediate for direct oxetane formation rather than the previously hypothesized mechanism involving an epoxide intermediate. Additional identification of the C_9 hydroxylase (T9αH1) allowed the authors to achieve complete reconstitution of **343** in *N. benthamiana*, while showing the redundant functions of two previously identified enzymes, T10βH and DBAT. While it is certainly intriguing that the three groups reached different conclusions, especially the heterologous reconstitution of **343** using different oxetane-forming candidates, these studies completed the long-sought after biosynthetic pathway of **7** and open up new synthetic biology opportunities. Subsequently, the Sattely lab identified an NTF2-like protein, FoTO1, that favored the production of **100** from **90** and suppressed the formation of over-oxidized shunt products when taxadiene synthase and T5αH are coexpressed.[603] Additionally, seven new genes were identified that enabled the heterologous reconstitution of the *de novo* biosynthetic pathway of **343** at levels comparable to the natural abundance in the needles of *Taxus.* A key to this success was to transcriptionally profile various cell states instead of conventional transcriptomics and co-expression analysis. This study is therefore valuable not only for its contributions towards unlocking the entire biosynthetic pathway of **7,** but also for the novel approach to identifying biosynthetic genes in a complex plant system.

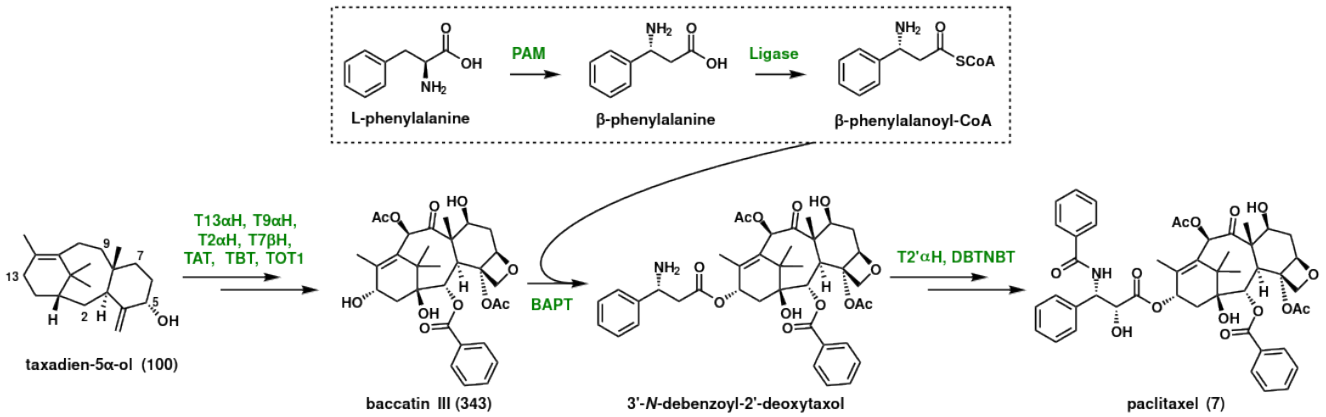


Figure 74. Late-stage biosynthetic pathway of paclitaxel characterized by heterologous expression in *N. benthamiana*.

In addition to these examples discussed above, many important pathways have been reconstituted in *N. benthamiana* using the combination of transcriptomics and metabolomics approaches described above. These include pathways producing strychnine (**344**), cocaine (**345**), ginkgolide B (**346**), lycopodium alkaloids such as huperzine A (**347**), *Solanum* alkaloids such as α-tomatine (**348**), Amaryllidaceae alkaloids such as the FDA-approved Alzheimer's drug galantamine (**349**), isoflavone phytoalexins such as capsidiol and glucoraphanin, gramine, and more (Figure 75).[604–614] It is not an understatement to say that the tobacco model plant has completely revolutionized plant natural product research and engineering.

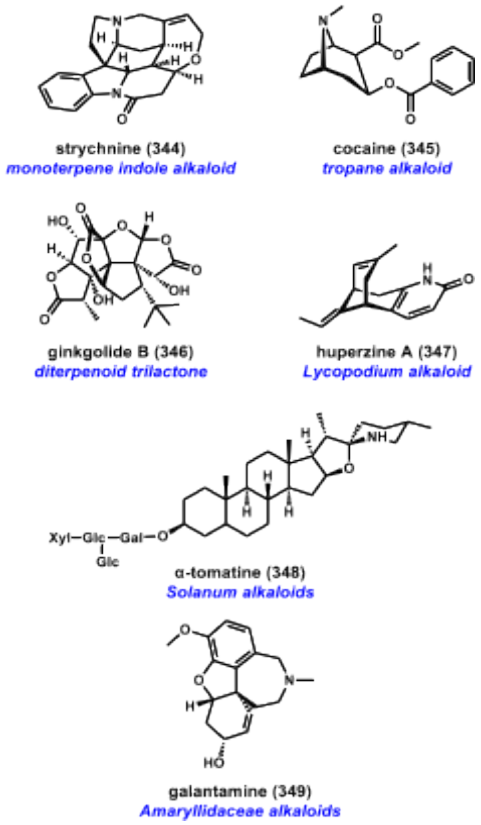


Figure 75. Notable natural products produced heterologously in *N. benthamiana*.

In addition to *N. benthamiana,* the ripening tomato fruit was shown to be a possible heterologous host due to its highly active plastidial MEP pathway, providing a suitable platform for FPP and terpenoid overproduction.[615] Metabolic engineering of the cultivated tomato plant, *Solanum lycopersicum* cv. Ailsa Craig, was shown to be suitable for the production of sesquiterpenes such as patchoulol. Tomato glandular trichomes, epidermal structures capable of secreting and storing secondary metabolites, have also been explored in the context of natural product heterologous expression and could lead to new innovations and tools in the future.[616–618]

4 Structural Diversification of Natural Products

In the previous sections, we summarized the impact of synthetic biology on the genome mining of natural products and enzymes discovery, both in the native producing host and in model heterologous hosts. Once the link between BGC and natural product is established, and enzyme functions are interrogated, the next goal is to engineer the biosynthetic pathways towards the production of unnatural analogs or "new-to-nature" compounds. The goal of these efforts is to alter the structures of natural products to improve their pharmacological properties, such as their potency and stability, as well as lower toxicity. Many classic examples of pathway engineering have been documented, starting with the earlier approaches that do not rely on biosynthetic information, to more recent examples that pinpoint exact enzymatic step to be engineered. In this section, a review of several key approaches will be presented. First, we will highlight recent strategies to engineer assembly-line megasynthetases that are based on new definitions of domain and module boundaries (Section 4.1). Second, pathway engineering based on building block diversification will be discussed, including precursor directed biosynthesis (PDB), mutasynthesis, and, to a lesser extent, biotransformation (Section 4.2). These strategies have been applied towards the diversification of cores for numerous classes of natural products, including polyketides, NRPs, RiPPs, terpenes and terpenoids, and alkaloids, as well as manipulating the cosubstrates utilized by tailoring enzymes for core decorations. Lastly, the concept of combinatorial biosynthesis in which enzymes from related or unrelated pathways are mixed to generate analogs will be discussed (Section 4.3). All these approaches have benefited from advances in synthetic biology, most notably protein engineering methods such as directed evolution.

4.1 New Approaches to Engineer Assembly-line Enzymes

The colinear, assembly-line logic of bacterial type I polyketide synthases (PKS) and nonribosomal peptide synthetases (NRPS) has been established for over 30 years.[39,619,620] The prospect of engineering these megasynthases for new natural product is emboldened by the modularity of the assembly lines. The implication is that it should be possible to rationally manipulate the modules through deletion, addition, or swapping to combinatorially access new-to-nature compounds. Yet, while many successes have been reported to support this promise, overall rational reprogramming attempts are mostly hampered by significantly reduced or abolished activity of the resulting enzymes.[20] In recent years, multiple X-ray and electron microscopy structures of these megasynthases have been elucidated,[621–626] and together with

modern bioinformatics and machine learning approaches, have enabled transformative approaches to reenergize pursuits of PKS and NRPS engineering. Here, we will review several of the most recent successes.

4.1.1 PKS Module Engineering

Over the past three decades, efforts to engineer PKSs through module-swapping have relied upon the "traditional" definition of a module, that consists of the ketosynthase (KS) domain in the position furthest upstream and the thiolation (acyl-carrier protein, ACP) domain in the position most downstream (Figure 76A). Deletion, addition and insertion of modules have resulted in formation of desired analogs in some examples. However, in most cases this strategy has not been successful in producing substantial quantities of targeted compounds. It is widely accepted that such swapping attempts destabilize the intimate protein-protein interactions required for assembly-line biosynthesis.[627] A new definition of PKS modules was proposed by Keatinge-Clay based on bioinformatics analysis of *trans*-AT type I PKSs.[628] The authors noticed that sets of domains that genetically migrate together during PKS evolution are not consistent with the traditional module definition. Instead, the tailoring domains and the ACP domain genetically migrate with the downstream KS domain, rather than the upstream one. This led to a new module definition, in which an ACP is kept with the downstream KS in a module as shown in Figure 76A. To test this hypothesis, Keatinge-Clay and coworkers used the venemycin PKS that produces venemycin (**350**) as a model system.[629] The three module PKS consists of two polypeptides: VemG contains the loading and the first module, while VemH contains the second module and the releasing thioesterase (TE) domain, as traditionally defined. In the new definition (evolutionary definition), VemG module 1 consists of the loading module (AT-ACP) and the first KS domain, to give a module 1 of AT-ACP-KS. Module 2 is split between the two polypeptides to consist of AT-ACP from the traditionally defined module 1 in VemG and the first KS domain in VemH. The new module 3 consists of the remaining AT-ACP-TE domains in VemH.

Figure 76. Evolutionary module boundaries are more robust in module swapping than traditional boundaries. (A) Comparison of traditional boundaries and new, evolutionary boundaries used to define modules for venemycin PKS. (B) Hybrid Pik and Vem PKSs constructed using both traditional and new boundaries result in synthesis of **351**. However, the hybrid PKS from new boundaries is far more robust when the turnover rates are measured.

Using both the tradition and evolutionary definitions, module swapping with the pikromycin PKS module 1 was performed, which led to the production of the propionate primed α-pyrone **351** in both cases (Figure 76B). Kinetic analysis showed the hybrid PKS built with the new domain boundaries displayed >25-fold higher turnover number than that built with traditional boundaries. Further domain swapping experiments between the venemycin and pikromycin PKSs led to the consistent observation that swapping using newly proposed boundaries led to significantly more robust hybrid PKSs. Combinatorial swapping was then performed to generate a library of compounds that can be obtained at preparative scale. Building on this success, the Keatinge-Clay lab built a three-peptide hybrid PKS Pik1567 (Figure 77). By optimizing expression levels in *E. coli* with T7 promoter tuning, the hybrid four-module PKS produced the tetraketide product **352** at

~100 mg/L titer.[630] With this new paradigm of how a module is defined, excised, and swapped, it is anticipated that more efficient engineering of even larger type I PKSs will be successful.

Figure 77. A four-module PKS built using evolutionary boundaries led to efficient synthesis of tetraketide **352.**

Recently, Piel and coworkers used multiple-sequence alignments to gain insight into the natural recombination sites that have evolved between KS and downstream domains. The goal is to use these sites as fusion points for the rational engineering of functional PKSs.[631] In-depth analysis of sequence alignments and the covariance of extracted motifs revealed protein sequences near a subdomain C-terminal to KS domains, referred to as the flanking subdomain (FSD), are highly conserved. The FSD was previously demonstrated to be involved in polyketide organizational dynamics by facilitating lateral PKS interactions required for higher-ordered PKS assemblies.[632,633] Specifically, computational analysis pinpointed a motif with the sequence of LPTYPFX_5W, which was previously shown to also be present between KS and AT domains in *cis*-AT PKSs.[634,635] Proposing this sequence to be universal to nearly all type I PKS modules as a site for engineering, the authors performed a large scale campaign to modify PKSs. The LPTYPFX_5W motif was used as a site to engineer diverse analogs of the complex natural products lacunalide A (**353a**) and B (**353b**) (Figure 78). This resulted in a high level of success, as demonstrated in the biosynthesis of truncated lacunalide analogs **353c-j** through module deletions and chimeric assembly-line construction. Therefore, this simple yet powerful approach enabled the authors to produce synthetically inaccessible polyketides in a highly rational and efficient manner.[631]

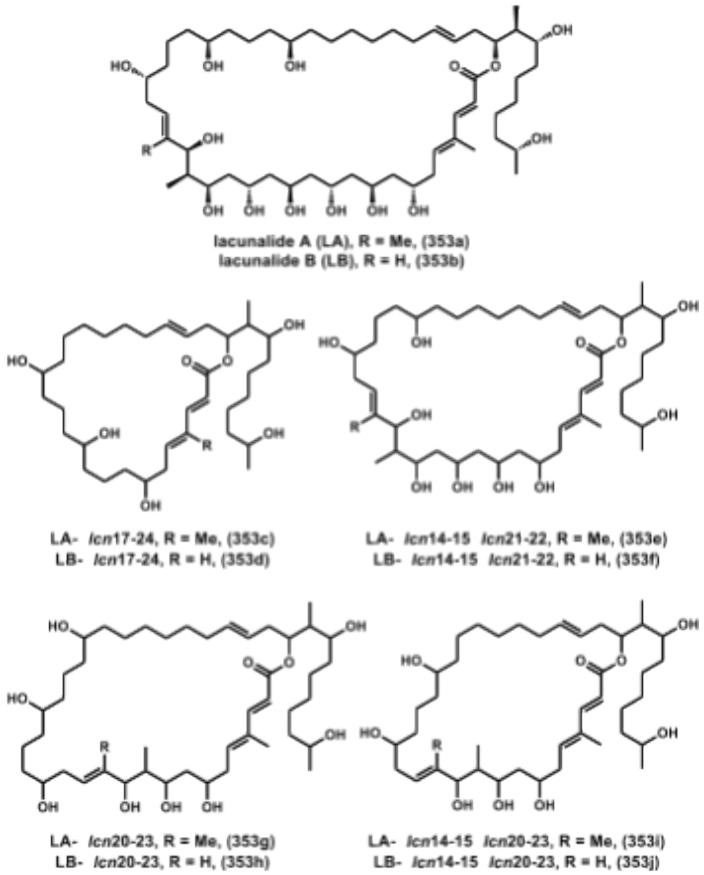


Figure 78. The engineering strategy employed by Piel and coworkers enabled the generation of a library of **353a-b** derivatives **353c-j**.

4.1.2 NRPS Module Engineering

Historically, a minimal module in an NRPS assembly line has been defined as C-A-T (condensation-adenylation-thiolation), and C-A didomains must be kept intact during engineering efforts to avoid perturbations to assembly line functionality.[636] However, just as is the case with type I PKSs, recent insights gained from bioinformatics analyses suggest that these traditionally accepted boundaries should be revised, as indicated by findings that natural recombination hotspots tend to be less invasive when used as module and domain boundaries compared to those traditional definitions.[42,637–641] Numerous studies have challenged so-called canonical domain boundaries and have instead focused on identifying evolutionary sites of recombination that are more tolerant for engineering efforts.[42,326,641–643] For example, in-depth *in silico* analysis of the evolution of NRPS diversity by Baunach et al. indicated recombination (e.g. inter- and intragenomic recombination, speciation, horizontal gene transfer) to be the key driver of NRPS diversification and directly translates into

the structural diversity of their products.[642] A key finding was that recombination events take place almost exclusively in A domains, specifically resulting in partial substitutions within the A_{core} subdomain.[642]

Recently, Bode and coworkers developed a series of new rules for engineering NRPSs using functional units called e**X**change **U**nits (XUs), rather than prototypical modules (Figure 79A).[640] In a thorough structural analysis of NRPS domain-domain interactions in *Xenorhabdus* and *Photorhabdus* proteobacteria, the authors defined an XU is comprised of an A-T-C domain architecture, rather than the traditional C-A-T.[640,644,645] The XUs could serve as swappable units based on a conserved motif in the interdomain C-A linker. The study demonstrated that one or two XUs could be successfully swapped without the significant drop in product titers frequently observed when using the traditional module boundaries.

While the XU strategy has been validated in different model systems and enabled the successful production of new peptides, a key limitation stems from the inability of downstream C domains to elongate the unnatural upstream peptidyl substrate as a result of swapping.[646] To address this limitation, Bode and coworkers introduced the e**X**change **U**nit between **C**ondensation domains (XUC) concept.[639] Through structural analysis of C domains, a new fusion point to target for domain exchanges was identified.[639,640] The identified region is in a flexible loop within the C domain and divides the C domain into its *N*- and *C*-terminal subdomains. Such regions serve as boundaries of the swappable XUC units bearing the module architecture C_{sub}-A-T-C_{sub}. An advantage of XUC-defined boundary is that the original interface between the C domain and the adjacent A domain is preserved.[639] Only 80 XUC building blocks are required to build any peptide using the 20 proteinogenic amino acids, whereas 800 XU building blocks would be required to achieve the same feat.[639] However, one disadvantage is an XUC containing a C domain cannot be combined with an XUC containing an epimerase (E) domain. Another caveat is that XUC building blocks lack compatibility between different genera, hence limiting the use of XUCs to NRPSs sourced from the same genera.

Figure 79. NRPS engineering using XU, XUC and XUT exchangeable units developed by the Bode lab. (A) Types of modular units used; (B) Application of the exchange units towards *de novo* synthesis of the proteasome inhibitor **34b** inspired from **34.**

Most recently, Bode and coworkers used computational modeling and principal component analyses to further investigate intergenomic recombination sites.[42] The authors' systematic analysis led to their discovery of a previously unreported recombination site within the T domain with a conserved FFxxGGxS motif. To validate the e**X**change **U**nit between **T** domains (XUT) approach, the authors sought to design an artificial biosynthetic assembly line to produce bioactive peptides that inhibit the eukaryotic proteasome, with the fungal natural product fellutamide B (**34**) as a molecular inspiration. Using both XU and the new XUT fusion sites, Bode and coworkers were able to draw from numerous NRPS fragments and effectively recombine them to construct a designer NRPS assembly line, which produced the targeted compound **34b** at 21.6 mg/L, and displayed desired biological property (Figure 79B). A crystal structure of **34b** in complex

with the yeast 20S proteasome was also obtained to validate the molecular design. This impressive demonstration is an important step towards realizing the true potential of the NRPS assembly lines.

4.2 Redirecting Biosynthetic Pathways with Precursors

The simplest and most classic method to diversify the structures of natural products from a producing organism is through the alteration of precursors and/or biosynthetic intermediates. Whether through feeding synthetic substrates to native or engineered hosts, or directly generating substrate analogs *in vivo*, precursor-oriented diversification has been used to produce new derivatives of natural products. This method requires gatekeeping enzymes incorporating unnatural substrates to be promiscuous, and that downstream tailoring enzymes are tolerant of structurally altered biosynthetic intermediates.

4.2.1 Precursor-Directed Biosynthesis

Precursor-directed biosynthesis (PDB) is a method that in its simplest form, requires little engineering of microbial hosts.[20] PDB entails the hijacking of native or engineered biosynthetic machinery to incorporate unnatural precursors into the product.[647–649] In 1948, Behrens and coworkers fed various short chain acids to microbial cultures of *Penicillium chrysogenum* (Figure 80), probing their potential to generate new penicillin derivatives for bioactivity assays without detailed insight into biosynthesis of penicillins.[650] In modern approaches of PDB, knowledge of the biosynthetic pathway and biochemical property of enzymes are leveraged in the preparation of unnatural precursors. In principle, precursors can be fed to producing organisms or crude extracts targeting both early and late steps in the biosynthetic pathway, as exemplified in the generation of unnatural erythromycin (**1**)[651] and halogenated monoterpene indole alkaloid (MIA) analogs, as examples.[429,652]

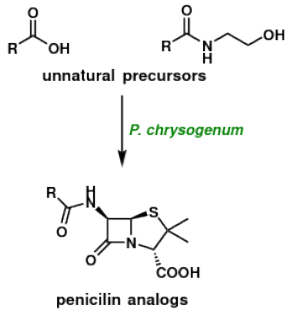


Figure 80. Earliest example of PDB involved feeding simple precursors to liquid fermentation cultures of *P*. *chrysogenum*. Diverse R-groups were incorporated in the penicillin products.

Although used interchangeably at times, PDB is typically associated with providing unnatural precursors to scaffold-building enzymes such as PKS, NRPS, etc.; while late-stage substrate feeding to engineered strains is known as mutational biosynthesis (mutasynthesis, see 4.2.2). During PDB, the precursor must be able to penetrate the cell membrane and be stable and nontoxic to the producing organism.[653] The precursor must also be kinetically competitive against the native substrate present at intracellular concentrations. When the unnatural building blocks cannot outcompete natural substrates which is often the case, PDB will result in accumulation of both natural and unnatural products, with the latter being minor or trace co-products even when precursors are provided at saturating concentrations.[654–658] These difficulties may be overcome through engineering gatekeeping enzyme substrate promiscuity or removing the supply of competing natural precursors as in mutasynthesis.

One powerful, and now classic, example of PDB is the generation of antimycin A3b (**354a**) analogs by Yan *et. al*.[659] The authors simultaneously manipulated the precursor supply of three parts of the biosynthetic pathway: NRPS starter unit, PKS extender unit, and post assembly-line tailoring (Figure 81). Specifically, 6-fluorotrytophan was fed to *Streptomyces* sp. NRRL 2288 to generate the unnatural 4-fluoroanthranilate that can be incorporated as a starter unit by the NRPS-PKS assembly line. A wide variety of C_7-alkyl modifications were generated by feeding nine enoyl-CoA variants, which produced unnatural alkyl malonyl-CoAs by action of the promiscuous reductive-carboxylase AntE. C_8-acyl modifications were generated by feeding 33 variants of acyl-CoAs for the final *O*-acylation step catalyzed by AntB. These enzymes displayed unusual substrate promiscuity and 380 analogs were generated and tested for biological activity. An impressive 65% of the analogs displayed enhanced cytotoxicity against the murine leukemia P388 cell line, while 21% of them demonstrated enhanced antifungal activity against *Candida albicans* compared to **354a**. In particular, **354b** displayed a 66-fold increase in cytotoxicity, while **354c** exhibited 2.0-fold antifungal activity compared to **354a.**

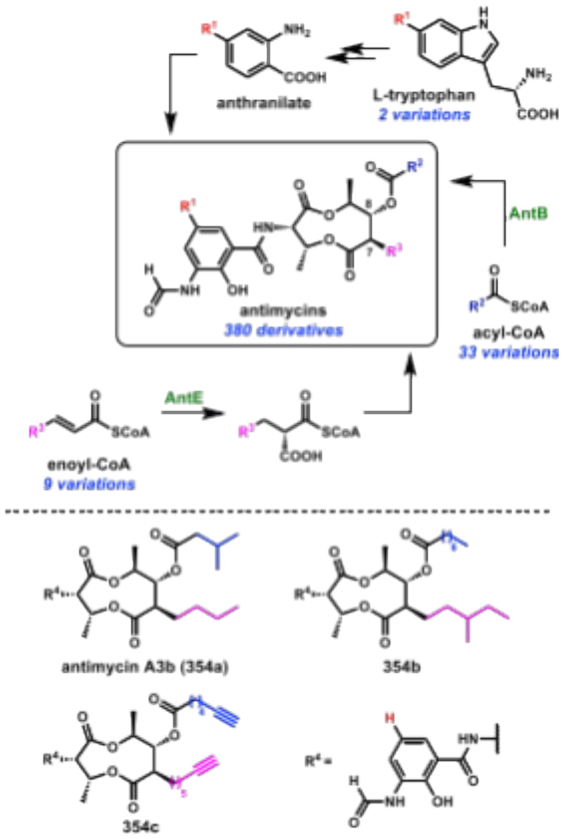


Figure 81. PDB of 380 antimycin derivatives by exploring the substrate promiscuity of three pathway enzymes.

Recent applications of PDB span numerous classes of natural products (Figure 82). For examples, i) Chen and coworkers exploited 4-halo, 4-methoxy, and 4-methylphenylalanine derivatives to generate analogs of the uridyl-peptide antibiotic sansanmycin H (**355**) from *Streptomyces* sp. SS., with fermentation titers greatly exceeding that of **355**.[660] Both 4-bromophenylalanine and 4-chlorophenylalanine derivatives of **355** exhibited four-fold enhanced antibacterial potency against *Staphylococcus aureus*, which was attributed to increased lipophilicity; ii) Codd and coworkers applied PDB towards the production of a series of desferrioxamine B (**356**) analogs with variations in the diamine portions of the molecule, whereas semisynthetic efforts had previously only allowed for modifications to ancillary fragments chemically coupled with the natural product.[661,662] Aliphatic diamines fed to *Streptomyces pilosus* outcompeted cadaverine to generate a variety of derivatives with altered metal-chelating abilities; iii) Nett and coworkers supplied various substituted aryl carboxylic acids to the pseudochelin pathway in *M. xanthus* and generated a series of pseudochelin A (**357**) and myxochelin B derivatives with activity towards human 5-lipooxygenase;[663] and iv) similarly, by feeding aryl carboxylic acid substrates to *Chrysosporium* sp., Capon and coworkers generated analogs of chrysosporazine A (**358**) and neochrysosporazines from

strains CMB-F214 and CMB-F294, respectively.[664] Analogs produced were useful in structure-activity relationships (SAR) studies as inhibitors of P-glycoprotein. The benzamide, acetamide, and methyenedioxy moieties were found to be essential towards its inhibitory activities.

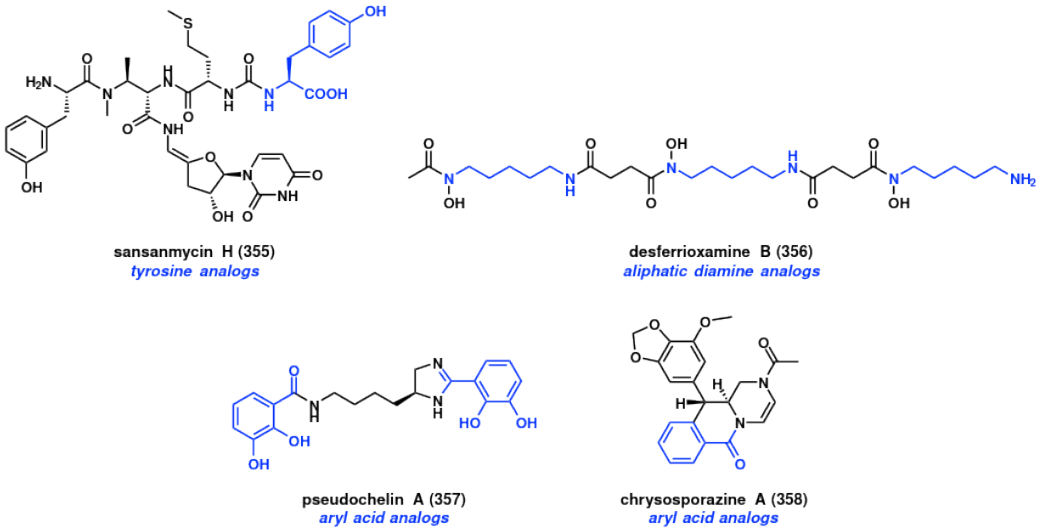


Figure 82. Application of PDB to enhance the bioactivities of natural products.

4.2.2 Mutational Biosynthesis

While mutational biosynthesis, or mutasynthesis, and PDB are often conflated, they represent distinct approaches toward pathway engineering.[665,666] Mutasynthesis directly addresses issues associated with PDB regarding competition of unnatural precursors with native substrates. Enzymes associated with natural precursor supply are genetically inactivated or downregulated, making unnatural analogs the only substrates for incorporation. The unnatural or synthetic analogs are referred to as mutasynthons. Mutasynthesis was initially applied over 50 years ago on mutants of the neomycin producer *Streptomyces fradiae,*[666,667] but has since been applied to many other compound classes including alkaloids,[668,669] peptides,[670,671] polyketides,[651] and their hybrids.[672–674] For example, i) the prodiginine biosynthetic pathway in the producing strain *P. putida* KT2440 was inactivated via knockout of the *pigD* gene (see section 2.1.2.5), which is the condensation enzyme involved in the biosynthesis of the precursor 2-methyl-3-amylpyrrole (MAP, **54**). In the native pathway, MAP is then coupled with 4-methoxy-2,2'-bipyrrole-5-carboxaldehyde (MBC, **53**) as a cosubstrate by PigC.[675] Feeding monopyrroles with varying alkyl substitutions to the deletion mutant enabled access to an array of prodigiosin (**55**) derivatives (Figure 83A); ii)

as an example of mutasynthesis using a heterologous host, Nett and coworkers recently generated a series of physostigmine (**359a**) derivatives by feeding synthetic analogs of the natural intermediate 5-*O*-methylcarbamoyl-*N*-acetylserotonin (MAS) to *M. xanthus* engineered to express the last three pathway enzymes (PsmBCD) (Figure 83B).[668] Other enzymes omitted from the biosynthetic pathway are responsible for the biosynthesis of MAS from L-tryptophan. The authors reported production of the drug candidate phenserine (**359b**), a compound which was only previously accessible via total synthesis; and iii) platensimycin and platencin (**105**) are both diterpenoid carboxylic acids amidated with 3-amino-2,4-dihydroxybenzoic acid (ADHBA).[676] By deleting *ptmB1* required for biosynthesis of ADHBA in the producer *Streptomyces platensis* SB12029, Shen and coworkers incorporated 18 differentially substituted aminobenzoic acid substrates to generate analogs.

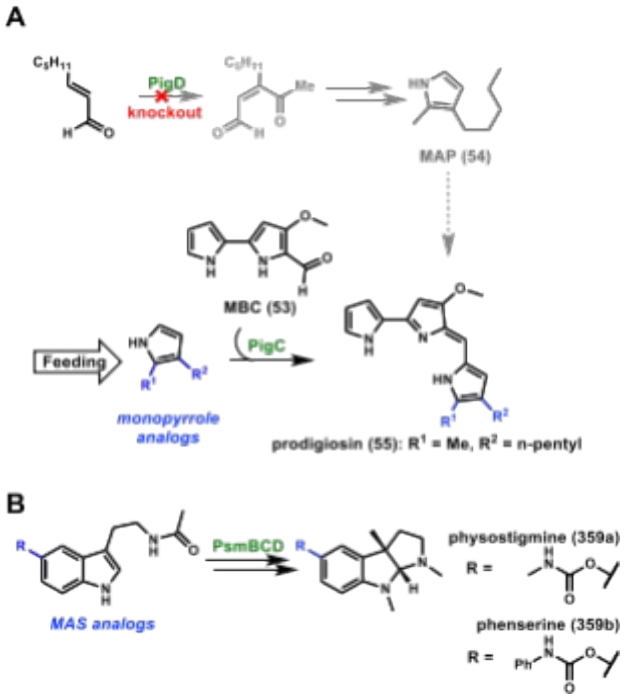


Figure 83. Examples of mutasynthesis. (A) Generation of prodigiosin (**55**) analogs; (B) generation of physostigmine (**359a**) analogs.

4.2.3 Biotransformation

Biotransformation is a classic form of structure diversification. The endogenous enzyme activities of a host are harnessed to modify the structures of fed natural products, effectively "extending" a biosynthetic pathway towards analog

generation.[677] Biotransformation can generate diverse natural product analogs, with stereoselective and regioselective modifications that are otherwise difficult to accomplish through semisynthetic means. Whole cell biotransformation also has the benefit of protecting and stabilizing enzymes within the intracellular environment and recyclable cofactors.[678] Microbial biotransformation has been utilized on industrial scale in conjunction with chemical synthesis for the manufacturing of chiral intermediates and end products, with applications in the production of pharmaceuticals including indinavir and atazanavir.[679]

The steroid industry has applied whole cell transformation, mainly taking advantage of the redox enzymes,[680] to generate pharmaceutically relevant compounds using precursors like phytosterols, sapogenins, and cholesterol as starting materials.[681] A key discovery by Peterson and Murray showed that progesterone could be readily hydroxylated in yields up to 95% at C_{11} by incubation with *Rhizopus nigricans*.[682,683] This is now an industrial biotransformation during which *R. nigricans* is used to modify 16-epoxyprogesterone (generated semisynthetically from the yam-derived diosgenin) into the 11α-hydroxy-16-epoxyprogesterone intermediate, which is diversified to various commercial corticosteroids, such a dexamethasone, prednisone, and prednisolone.[684] Recent work demonstrated additional biotransformation hosts that can modify natural products with the steroidal skeleton.[685,686] As an example, the plant-origin sapogenin neoruscogenin (**360**) isolated from *Ruscus,* when fed to the fungus *Alternaria eureka,* resulted in the accumulation of new derivatives with the cholestane and furostanol-type steroidal frameworks (Figure 84A).[687] Similarly, cyclocephagenol (**361**) fed to *A. eureka* resulted in dehydration, methyl migration, and ring expansion modifications in generation of 21 new products (Figure 84B).[688] Most products exhibited promising neuroprotective activity and antioxidative activity. Biotransformation of astragenol using the endophytic fungi *Syncephalastrum racemosum* AS 3.264 and *Alternaria alternata* AS 3.4578 resulted in acetylation, cycloetherifications, and ring expansion modifications.[686]

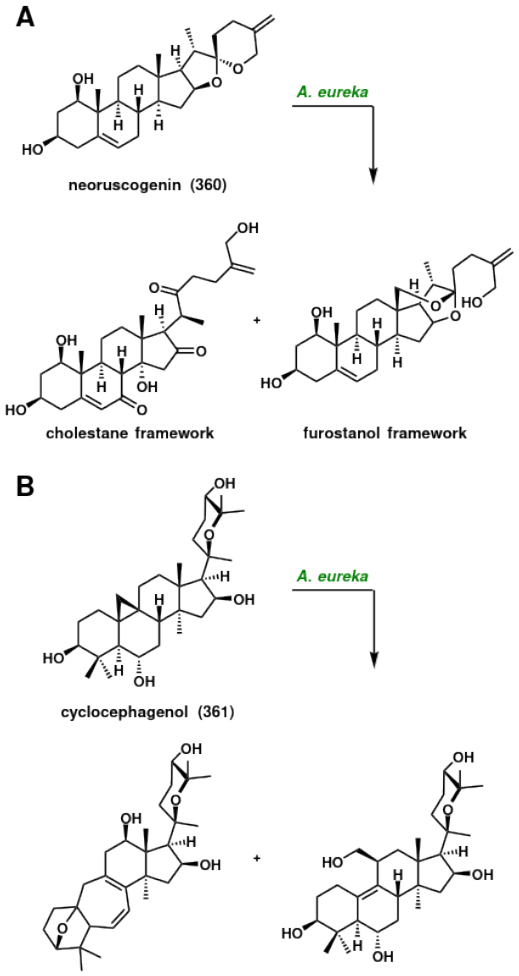


Figure 84. Biotransformation of (A) **360** and (B) **361** using *A. eureka* as a host.

Research groups continue to search for new microorganisms to be applied as hosts for biotransformation.[686,689] After screening 28 endophytic fungi for the biotransformation of β-mangostin (**362**) isolated from *Garcinia mangostana*, Arunrattiyakorn et al. found that feeding to cultures of *Xylaria feejeensis* GM06 resulted in generation of ring-fused xanthines, such as the (-)-mangostafeejin B (**363**), with an unprecedented hexacyclic heterocyclic skeleton (Figure 85). Feeding of curcumenol (**364**) to *Mucor polymorphosposporus* resulted in a series of rearrangement and ring contraction products (Figure 86).[690] Other recent applications of bioconversion include feeding of licochalcone A (**365**)[691] to *Aquilegia coerulea* and quercetin (**17**) to *Isaria fumosorosea*,[692] after which sulfonation and glycosylation of the substrates were observed, respectively. While most biotransformation applications have focused on using wild-type organisms as hosts, a recent example has shown the potential in using engineered hosts that express oxidative enzymes.[693] In an interesting example of biotransformation, Gliedar and coworkers engineered *Pichia pastoris* to overexpress the human P450 3A4,

enabling the diversification of papaverine (**366**) and strychnine (**344**) to novel products with net carbon- and nitrogen-oxidations. As more organisms and enzymes are probed for their biocatalytic utility, the range of chemical modifications to complex natural products achievable by biotransformation will certainly expand.

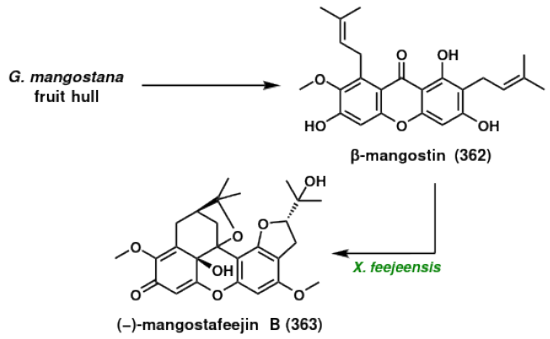


Figure 85. Feeding of **362** to X. feejeensis enabled the generation of **363** with an unprecedented chemical scaffold.

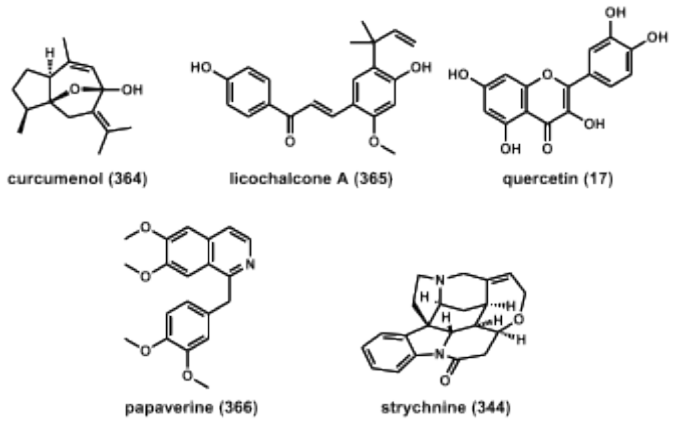


Figure 86. Examples of natural products that have undergone structural diversification via biotransformation.

4.2.4 Examples of Redirecting Polyketide Biosynthesis

4.2.4.1 PDB with Assembly-Line PKSs

Pioneering work in applying PDB toward polyketide structural diversification was performed with bacterial type I PKSs, specifically the megasynthases responsible for erythromycin (**1**) and pikromycin biosynthesis. The three PKS enzymes (DEBS1, DEBS2 and DEBS3) that synthesize 6-deoxyerythronolide B (**99**) are the prototypical polyketide assembly lines. Unnatural acyl-CoA and synthetic diketide *S*-*N*-acylcysteamine (SNAC) thioesters were used to probe the starter unit

specificity of the loading module and the KS domain of module 2 (traditional definition), respectively (Figure 87A).[694,695] Acyl-SNAC thioesters mimic the acyl-*S*-phosphopantethenoates of acyl-CoA and acyl-ACP.[696,697] To eliminate competition of propionate incorporation by the loading module, the loading AT domain was inactivated, followed by feeding of monoketide-SNAC to prime the assembly line. Similarly, inactivation of the KS domain in module 1 allowed exclusive priming of module 2 with diketide-SNAC.[651,698] In one example, feeding acyl-SNAC **367** as a precursor led to the production of phenyl 6-deoxyerythronolide B (**99b**), which was then biotransformed by *S. erythrea* A34 to the corresponding phenyl erythromycin analog (**1b**) (Figure 87B).

Figure 87. PDB of type I PKS using acyl-SNACs as substrate mimics. (A) Scheme of intercepting type I PKS with unnatural starter and extender units; (B) PDB to produce **1b**.

Feeding of acyl-SNACSs has also been applied to iterative fungal PKSs, especially to confirm if a proposed acyl-chain is indeed an intermediate. This is more challenging than the modular PKSs, as only a single set of domains is present and inactivation of either the KS or AT domain abolishes all enzyme activities. Instead, Vederas and coworkers synthesized and fed ^{13}C-labeled acyl-SNACs as precursors to the dehydrocurvularin and hypothemycin PKSs to compete with natural acyl

intermediates, and used ^{13}C NMR of the resulting products to verify incorporation of the advanced acyl intermediates.[699,700] In a classic example of acyl-SNAC feeding, feeding of the triene-containing, hexaketide acyl-SNAC to the lovastatin nonaketide synthase, followed by NMR characterization verified that the Diels-Alder reaction is enzyme catalyzed.[701] Dual PKSs pathways in fungi produce compounds such as resorcylic acid lactone through a distribution of labor between a highly reducing PKS and a nonreducing PKS. To diversify the products, the highly reducing PKS can be removed, followed by feeding of acyl-SNAC substrates that can directly prime the KS domain of the nonreducing PKSs.[702] Similarly for aromatic fungal PKSs, the biosynthesis of numerous naphthopyrone analogs was observed following feeding of linear, branched, and functionalized (such as nitrile-containing, ether-containing, thioether-containing, halogenated) fatty acyl-SNACs as starter-units to the iterative PKS PksA.[703]

Different nucleophiles in the PKS assembly line have been shown to intercept the acyl-SNAC during PDB, leading to a distribution of products. This was demonstrated by Sherman and coworkers in the feeding of advanced hexaketide acyl-SNAC **368a** to the final PikAIV module of narbonolide synthase. Analysis of products showed a 4:1 ratio between the directly-cyclized 12-membered macrocycle 10-deoxymethynolide (**369**) versus the chain-extended and cyclized 14-membered macrocycle narbonolide (**370**) (Figure 88A).[704] In contrast, feeding of the native PikAIII pentaketide acyl-SNAC (**371**) to PikAIII together with PikAIV resulted in the desired **370** as the major product (Figure 88B). The formation of **369** with PikAIV was attributed to the capturing of **359a** by the active site serine of the TE domain in favor of the KS domain cysteine, while the shorter **371** is not recognized by the TE. By modifying the acyl-SNAC carrier to thiophenol as in **368b** to deter TE recognition, the product profile was shifted to **370** as the major product (Figure 88C). With this knowledge in hand, Sherman and coworkers supplied pentaketide thiophenol thioesters mimicking the natural substrate to PikAIII-TE. This resulted in the production of four desmethyl analogs of **369**.[705] In subsequent work, the group also demonstrated the TE could be engineered via a single active site mutation to cyclize an epimer of the native hexaketide substrate.[706] This gain-of-function macrocyclization was rationalized through quantum mechanical modeling to occur through a concerted acyl substitution rather than a stepwise addition-elimination mechanism associated with the wild-type TE domain. Highlighting the utility of thiophenol thioesters as an alternate to acyl-SNACs, Sherman and coworkers synthesized the hexaketide thiophenol thioester to prime module 5 of the juvenimyicin PKS in their efforts to chemoenzymatically synthesize juvenimicin and rosamicin macrolides.[707]

Figure 88. PDB with narbonolide PKS with advanced acyl thioesters.

4.2.4.2 Engineering Extender Unit Supply

The promiscuity of the AT domains in *cis*-AT PKSs has been explored for analog generation at targeted positions in a polyketide product.[708,709] The unnatural extender units, however, must be in the form of activated α-carboxyacyl-CoA thioesters. Feeding of acyl-CoA derivatives to the producing organisms is unfeasible because of poor membrane permeability and synthetic challenges. To overcome the extender unit supply issue, three routes of direct enzymatic

production of the α-carboxyacyl-CoA have been demonstrated.[710] Each route requires coexpression of additional enzymes and feeding of precursors, essentially extending the PDB one step prior to the assembly line.

One route is the ATP-dependent thioesterification of malonate derivatives using malonyl-CoA synthetase (MatB), which has received considerable attention due to straightforward synthesis of malonic acid derivatives.[635,697,711–713] For instance, MatB_Sc from the monensin (**372**)-producing organism *Streptomyces cinnamonensis* was found to be promiscuous in the activation of substituted malonates (Figure 89A). [714] Using this approach, malonate diesters are typically fed to the cultures due to stability and nontoxicity. The diesters are hydrolyzed to malonate upon cell entry for MatB_Sc catalyzed CoA-thioesterification. Schulz and coworkers supplied malonate derivatives with varied chain lengths, degrees of unsaturation, and heteroatom substitutions to fermentation cultures of *S. cinnamonensis*, and detected accumulation of new derivatives of **372** via incorporation by the promiscuous module 5 AT of the PKS. In a work by Williams and coworkers, a Thr207Gly/Met306Ile double mutant of MatB from *Rhizobium trifolii* was engineered with enhanced selectivity for nonnatural extender units, including isopropyl-, butyl-, phenyl-, and azido-malonate analogs.[712] This mutant has since been used as a supplier of unnatural extender units to probe PKS substrate specificity, and to generate unnatural polyketides, such as analogs of kirromycin and 6-deoxyerythronolide B (**99**).[635,715–719] The promiscuous MatB from *S. coelicolor*[720] has also been used for these purposes.[721–723]

The reductive carboxylation of α,β-unsaturated acyl-CoAs by crotonyl-CoA reductase/carboxylase (CCR)[724] has also been utilized in the generation of unnatural extender units, as prominently featured in the antimycin example in Figure 81. Several BGCs encode a pathway-specific CCR enzyme to supply a rare extender unit to the PKS core enzyme, such as ethylmalonyl-CoA,[725] allylmalonyl-CoA,[726] and chloroethylmalonyl-CoA.[727] In an effort to further expand the substrate scope of AntE to synthesize aromatic extender units *in vivo*, Abe and coworkers used structure-guided engineering to enlarge the binding pocket.[728] AntE Val350Gly mutants readily accepted 3-(2-thienyl)acrylic acid and *para*-methylcinnamic acid to be subsequently incorporated into antimycin analogs. *In vitro* assays indicated that mutants could also act upon *para*-coumaric acid and 3-indoleacrylic acid, although the corresponding antimycins were not generated *in vivo* due to substrate intolerance of the antimycin PKS. KntE is a CCR involved in the biosynthesis of kitaniitetronin A (**373**) and converts 2-hexenoyl-CoA to butylmalonyl-CoA as an extender unit. The *knt* BGC for kitaniitetronin biosynthesis was heterologously expressed in a *S. albus* Δ*kntE* mutant that expresses the promiscuous AntE CCR instead.[710] The enhanced promiscuity of

AntE enabled the synthesis and incorporation of cinnamic acid, and a variety of other six-carbon α,β-unsaturated fatty acids to generate seven new tetronate compounds (Figure 89B).

The third and less utilized route is α-carboxylation of acyl-CoA by acyl-CoA carboxylases (ACCase).[729] In ACCases, the α- and β-subunits catalyze the carboxylation of activated biotin and the subsequent carboxylation of acyl-CoA, respectively.[730] The β-subunit encoded by *arm13* from the armeniaspirols biosynthetic pathway is unusual in its expanded substrate scope towards medium-chain acyl-CoAs.[731] Arm13 forms a catalytical complex with saAccA2a (ACC α-subunit) from *Streptomyces armeniacus* to catalyze the carboxylation of a wide variety of aliphatic and aromatic acyl-CoAs (Figure 89C). Feeding the methyl ester precursors to the fermentation broth led to α-carboxylation, and subsequent incorporation to produce several new analogs of armeniaspirol A (**374**).

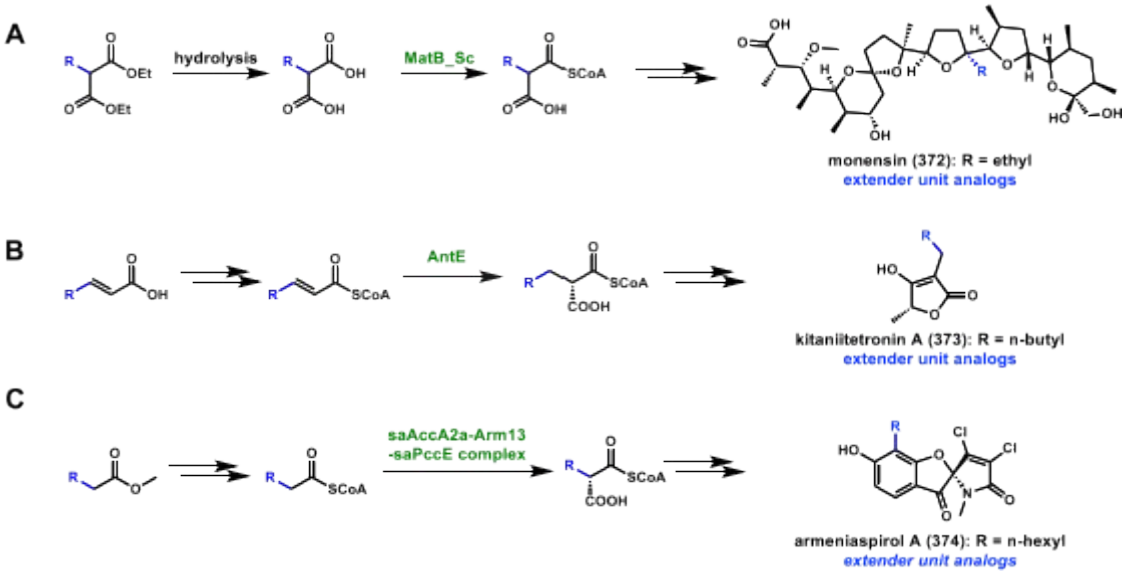


Figure 89. Unnatural polyketide extender unit can be generated via feeding of unnatural carboxylic acids and esters to strains expressing promiscuous enzymes including (A) malonate-CoA synthetase; (B) crotonyl-CoA reductase/carboxylase; or (C) acyl-CoA carboxylase.

The incorporation of orthogonal reactive handles into natural products using PDB can facilitate "click" conjugation with various functional groups for structural diversification. The Zhang lab used the JamA-C enzymes from jamaicamide BGC to introduce extender units with a terminal alkyne for incorporation into polyketides.[732] Fatty acyl-CoA ligase JamA and fatty acid desaturase JamB catalyze the transfer of 5-hexenoic acid (**375**) to the ACP JamC, and the desaturation to terminal

alkyne to form 5-hexynoyl-JamC (**376**), respectively (Figure 90). Coexpression of *jamABC* with *antCDEFGM* from antimycin BGC in *E. coli* resulted in the production of alkyne-bearing 8-deacyl antimycin analog **377**. The ACP-bound **376** is proposed to be processed by endogenous fatty acid metabolism to 5-hexynoyl-CoA (**378**), prior to carboxylation by AntE and incorporation by AntD as an extender unit. **376** was also explored as an unnatural starting unit to prime a truncated LipPKS1 lacking the loading domain.[733] Enhancing the protein-protein interactions between LipPKS1 and JamC via docking domain manipulation and site-directed mutagenesis of JamC was required for successful priming by the unnatural starter unit.

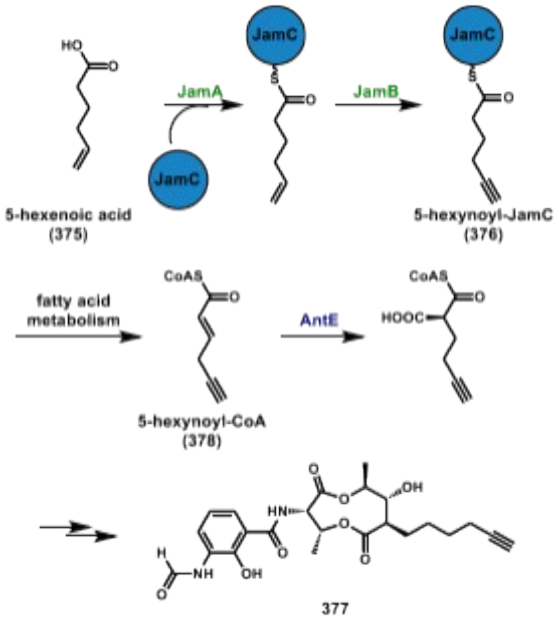


Figure 90. Incorporation of alkyne-containing extender unit into polyketide structures using the JamABC cassette.

4.2.4.3 Engineering Acyltransferase Domains

The AT domain of PKSs has been extensively explored through structural, computational and mutagenic studies towards the incorporation of unnatural extender units into polyketide structures. Several free-standing *trans*-ATs have shown considerable substrate promiscuity and have been used, including ZmaF[734] and KirCII.[717] Rittner and Sherman incorporated fluorine atoms into both 10-deoxymethynolide (**369**) and narbonolide (**370**) by swapping the native AT domains of DEBS module 6 with the substrate-tolerant MAT from metazoan type I FAS.[735] Feeding pentaketide and hexaketide thiophenol thioesters to this hybrid DEBS module 6, in the presence of fluoromalonyl-CoA and

fluoromethylmalonyl-CoA resulted in the generation of macrolactones with flouro and fluoromethyl substitution at the α-carbon. Adding isolated 2-fluoro-10-deoxymethynolide (**369b**) to the biotransformation host *Streptomyces venezuelae* YJ112 generated fluorinated derivatives 2-fluoro-YC-17 (**379**) and 2-fluoromethymycin (**380**) (Figure 91). Using a computational-guided approach, Englund et al. altered the extension substrate of DEBS M6+TE by exchanging the AT domain with those that can accept sterically bulkier extender substrates, such as phenylmalonyl-CoA, isopentylmalonyl-CoA, and hexylmalonyl-CoA.[635] Based on kinetic analysis, successful AT domain swapping requires the KS-AT linker (KAL) and post-AT linker 1 (PAL1) be maintained to preserve proper protein folding.

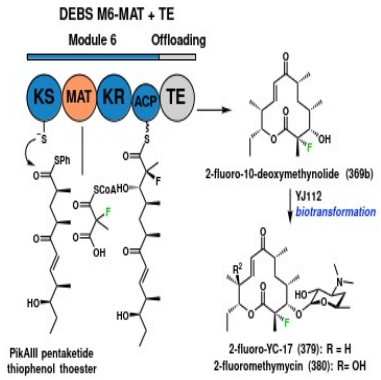


Figure 91. Promiscuous AT can be explored to generate fluorinated macrolides.

Efforts to engineer AT selectivity through mutagenesis have also been successful. To improve the recognition of AT domain of DEBS module 6 towards propargylmalonyl-CoA, Williams and coworkers generated saturation mutagenesis libraries of Tyr189 and Ser191 that are part of the specificity-determining **Y**A**S**H loop.[718] Screening of this library led to the identification of a Tyr189Arg mutant that can incorporate the unnatural substrate with twice the efficiency into 6-deoxyerythronolide B (**99**) over the native methylmalonyl-CoA. By using a mutant *trans*-AT from the disorazole biosynthetic pathway (DszAT) to complement an inactive *cis*-AT domain, Sirirungruang et al. successfully inserted fluorine into **99** via fluoromalonyl-CoA to generate 2-fluoro-2-desmethyl 6-deoxyerythronolide (**99c**) (Figure 92).[736] Crystal structure of DszAT revealed that Phe190 near the active site plays a key role in sterically enclosing the oxyanion hole and orienting malonyl-

CoA as a preferred substrate. They hypothesized that expanding the oxyanion hole would reduce the stringent steric requirement for malonyl-CoA and allow fluoromalonyl-CoA to be electronically preferred as an inductively activated extender unit. Based on these observations, the authors constructed and screened a saturation mutagenesis library at Phe190. The Phe190Val mutant showed high specificity towards fluoromalonyl-CoA and reduced specificity for malonyl-CoA, enabling site-specific incorporation of fluorine at positions 2- or 4 of **99** by inactivating the corresponding *cis*-AT domains.

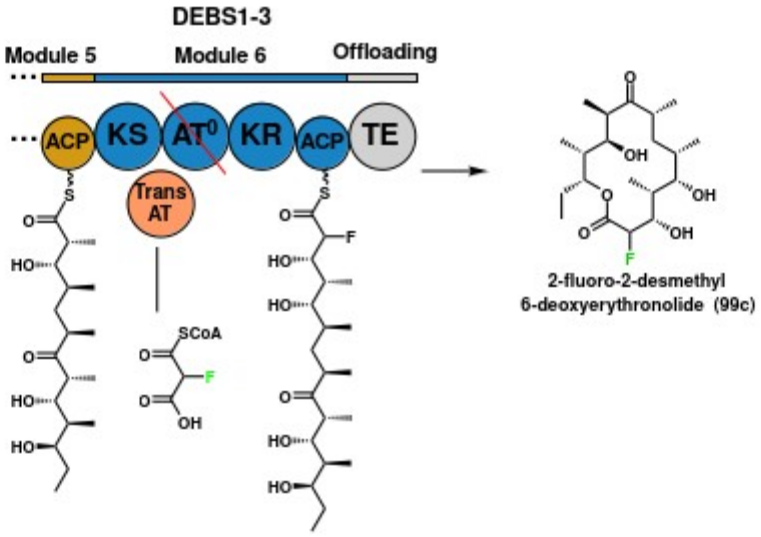


Figure 92. Using engineered *trans*-AT to alter extender unit specificity.

Similarly, using the AT domain of DEBS module 6 as a model system, Williams and coworkers performed molecular dynamics (MD) simulations to identify mutation sites for generating analogs of **99**.[719] MD simulations led to mapping of amino acid residues proximal to the YASH loop or the VDVVQP motif that can be mutated to probe extender unit specificity, many of which were not previously associated with AT specificity. Thirteen out of thiry-two mutants were found to confer specificity towards alternate malonyl-CoA extender units based on competition assays with methyl, ethyl, propargyl, isopropyl, butylmalonyl-CoAs. The Tyr74Arg substitution, for instance, resulted in a two-fold increase in selectivity for

propargylmalonyl-CoA relative to the wild-type. MD results and *in vitro* assays also implicated two uncharacterized, conserved regions in the AT domain residues 739-747, and 611-617, in playing substantial roles in extender unit specificity. Swapping these two motifs with the analogous motifs from the hexylmalonyl-CoA selective AT from module 1 of cinnabaramide PKS, or the butylmalonyl-CoA-selective AT from module 13 of thailandin PKS, resulted in chimeric AT domains that were more specific for these larger extender units.

4.2.5 Examples of Redirecting NRP Biosynthesis

4.2.5.1 PDB and Mutasynthesis of Assembly-line NRPSs

The most common PDB of nonribosomal peptides is through feeding of unnatural amino acids or other noncognate acids to the producing organism, in anticipation that the assembly lines will be able to accept and elongate these residues into the final product. A classic example of this strategy, in addition to that discussed for penicillin (Figure 80), is the feeding of allylglycine, β-cyclohexylalanine, and D-serine to *Tolypocladium inflatum* to generate cyclosporine analogs.[737] These nonnative amino acids replaced α-aminobutyric acid, *N*-methyl-4-(2-butenyl)-4-methyl threonine, and D-alanine residues, respectively, in the final products. Given the substrate scope and promiscuity of A domains can vary significantly, the success rate of PDB also vary greatly between different NRPSs, and different modules within a NPRS.[657,738–741] If the A domain is promiscuous, as found in cyanobacterial NRPS that produces microcystins, PDB can be applied to incorporate more diverse analogs.[742] For example, Dittmann and coworkers biosynthesized numerous "clickable" cyclic heptapeptide microcystins through supplementing azido and alkynyl amino acid derivatives to the cyanobacteria producers.[743] Conjugating one of the analogs with a fluorescent 7-hydroxycoumarin allowed the authors to visualize microcystins intake into HEK293 cell lines expressing the human organic anion transporting polypeptides 1B1 or 1B3 using time-lapse microscopy.[742]

NRPS-containing BGCs commonly encode dedicated enzyme(s) to synthesize nonproteinogenic amino acids for incorporation into the final product. Inactivation of these dedicated enzymes would shut down the NRPS pathway, unless the amino acid substrates are chemically complemented through feeding. Mutasynthesis strategies can therefore take advantage of this through feeding of unnatural substrates to blocked mutants. Hitachimycin (**381**) from *Streptomyces scabrisporus* contains an (*S*)-β-phenylalanine (**382**) residue that is generated by the clustered phenylalanine-2,3-

aminomutase HitA (Figure 93A).[744] Inactivation of HitA followed by feeding of (*S*)-β-phenylalanine analogs led to the synthesis of several derivatives of the macrolactam antibiotic. Several analogs exhibited more potent cytotoxicity against HeLa cells compared to **381,** including those containing *meta*-halogenation in the supplied **382** mimics. A similar approach was to inactivate the four-genes pathway responsible for 4-methyl-3-hydroxyanthranilic acid (**383**) biosynthesis in *Scutellaria costaricana* SCSIO ZS0073 (Figure 93B). Feeding the mutant with halogenated analogs resulted in a series of actinomycin D $X_{o\beta}$ (**384**) analogs with modifications on the chromophore.[745] Similarly, inactivation of the IlaM and IlaN involved in the biosynthesis of 3-nitrotyrosine (**385**) from *Streptomyces atratus* with subsequent feeding with 3-halogenated tyrosines and 3,5-dinitrotyrosine enabled the preparative isolation of 16 analogs of ilamycin F (**386a**) (Figure 93C).[746] The halogenated analogs generally demonstrated potent antitubercular activity, with **386b** exhibiting a MIC of 0.29 μM against *Mycobacterium tuberculosis* H37Rv. In contrast, direct PDB with the wild-type *S. atratus* was not successful due to preference of the corresponding NRPS A domain for **385** over unnatural amino acids.

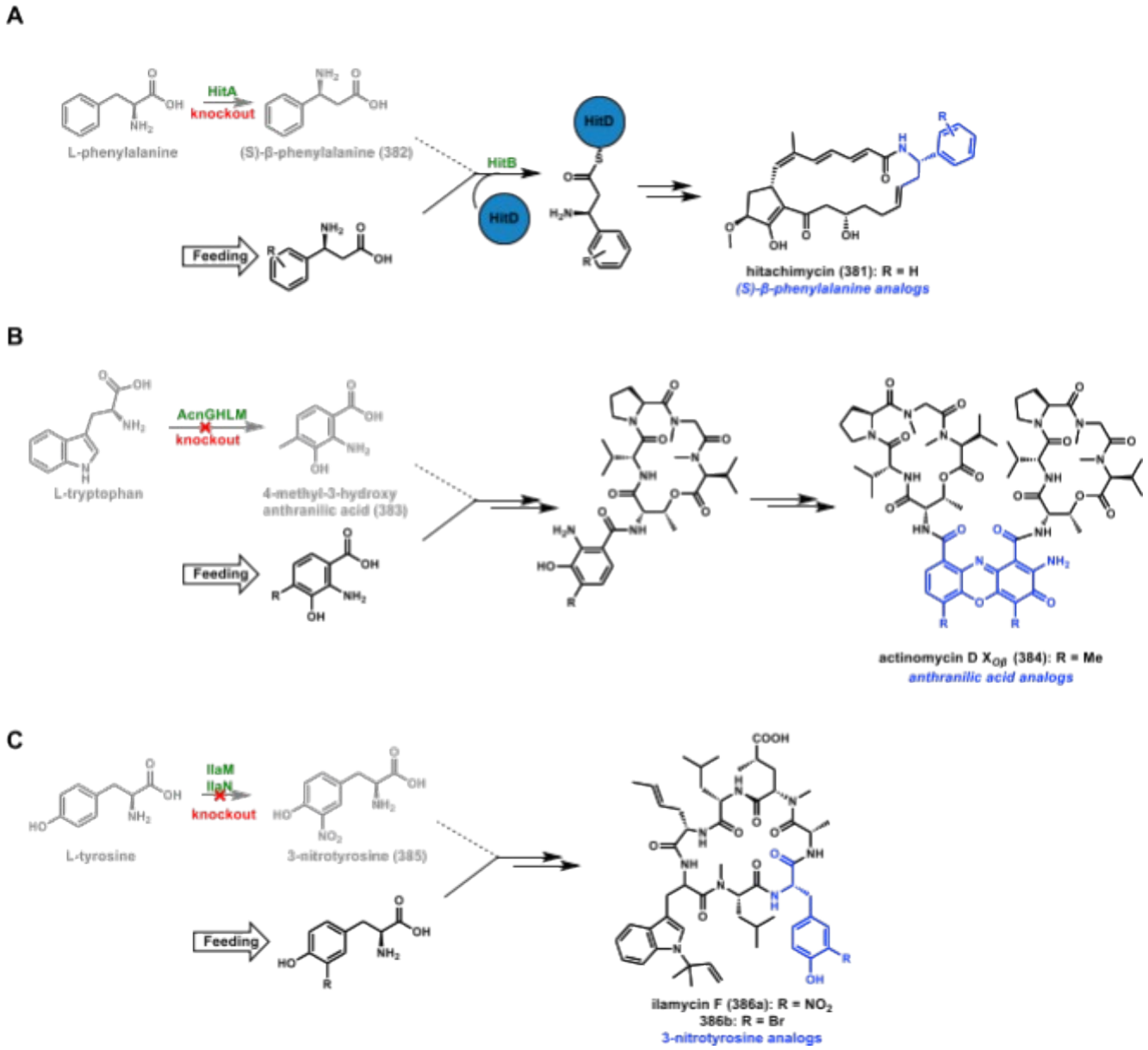


Figure 93. Mutasynthesis of NRP through inactivation of pathways for nonproteingenic amino acid biosynthesis.

One application of repurposing NRPS enzymes/domain towards generation of non-natural products is the chemoenzymatic cyclization of linear peptides by the macrocyclization TE domains.[747] While this is a challenging reaction to perform synthetically, the TE domains can achieve excellent regioselectivity in macrocyclization using linear substrates acylated to the upstream thiolation domain.[748–750] Penicillin-binding protein (PBP)-type thioesterases (TE) have been mined and applied towards the generation of large libraries of macrocycles. Using one such enzyme SurE from surugamide biosynthesis in *Streptomyces albidoflavus*, ethylene-glycol esters of linear peptides (6mer, 8mer and 11mer) synthesized

from solid-phase peptides synthesis can be readily cyclized.[751] Functional groups such as alkyne and azide were tolerated by the PBP-TE, which were subjected to copper-catalyzed azide–alkyne cycloaddition reaction to yield novel bicyclic macrocycles. The PBP-TE Ulm6 from *S. albus*, was shown to readily cyclize *S*-butyl-3-mercaptopropionate (SBMP) thioester of hexa-, penta- and tetrapeptides.[752] Nearly all tested peptides with an *N*-terminal L-amino acid and a *C*-terminal D-amino acids were accepted.

4.2.5.2 Engineering Adenylation Domains

The substrate selectivity of A domains has been frequently engineered through mutagenesis. In bacterial A domains, substrate specificity is largely determined by ten residues in the substrate-binding pocket, referred to as the specificity code.[753] The residues that control specificity of fungal A domains are not as well-defined, although new rules from bioinformatics analysis are emerging.[754] Micklefield and coworkers performed a single Lys278Gln mutation in the A domain of module 10 of the NRPS CdaPS3 involved in CDA4a biosynthesis to change the amino acid specificity from (2*S*,3*R*)-3-methyl glutamic acid (mGlu) to (2*S*,3*R*)-3-methyl glutamine, which was incorporated to generate a CDA4a analog.[755] In a study by Eguchi and coworkers as a follow up to the mutasynthesis of hitachimycin (**381**) (Figure 93A), the authors noted Phe328 and Thr293 are in proximity to the *meta* and *ortho* positions of (*S*)-β-phenylalanine (**382**), respectively. Large to small mutations at the two positions allowed the incorporation of substituted **382**, including, (*S*)-*meta*-nitro-β-phenylalanine, (*S*)-*meta*-cyano-β-phenylalanine, (*S*)-*ortho*-methyl-β-phenylalanine and (*S*)-*ortho*-bromo-β-phenylalanine.[756] In efforts to engineer the substrate specificity of the aryl acid A-domain DhbE, which natively accepts 2,3-dihydroxybenzoic acid (DHB) for bacillibactin biosynthesis, a yeast cell surface display platform with fluorescence-activated cell sorting (FACS) was implemented by Yin and coworkers.[757] Two mutants had over 200-fold reprogrammed selectivity for 3-hydroxybenzoic acid and 2-aminobenzoic acid relative to the native substrate. Tanabe and coworkers expanded the substrate scope of a homologous DHB-activating enzyme EntE via a single Asn235Gly mutation, which enhanced loading of 3-nitrobenzoic acid and 3-cyanobenzoic acid relative to the wild-type. [758]

Messenger et al. undertook an interesting metagenomic approach for high-throughput, large-scale A domain diversification.[759] Next-generation sequencing of degenerately amplified A domains showed that a single sample of soil metagenomic DNA can harbor thousands of unique A domains.[760,761] Conserved motifs within NRPS sequences served as

viable recombination sites for large-scale domain substitutions and enabled 119 functional domain substitutions into the pyoverdine NRPS to generate 16 novel analogs. This was a significant increase from the largest previous report of performing substitutions with 12 unique domains. In this case, screening was facilitated by UV-fluorescence of the dihydroxyquinoline moiety of pyoverdine. While conceptually different than active site mutagenesis, metagenomic domain substitution offers an alternative approach to generate protein sequence diversity.[641] Working with the massive 2.0 MDa enduracidin NRPS that produces enduracidin A (**387a**), Thong et al. used CRISPR-Cas9 mediated strategies to exchange A domains both within the enduracidin NRPS, as well as introducing domains from diverse NRPSs (Figure 94).[637] These efforts afforded ten new lipopeptide variants, such as **387b-e,** that were produced in titers up to 82% relative to the wild-type in *Streptomyces fungicidicus*.

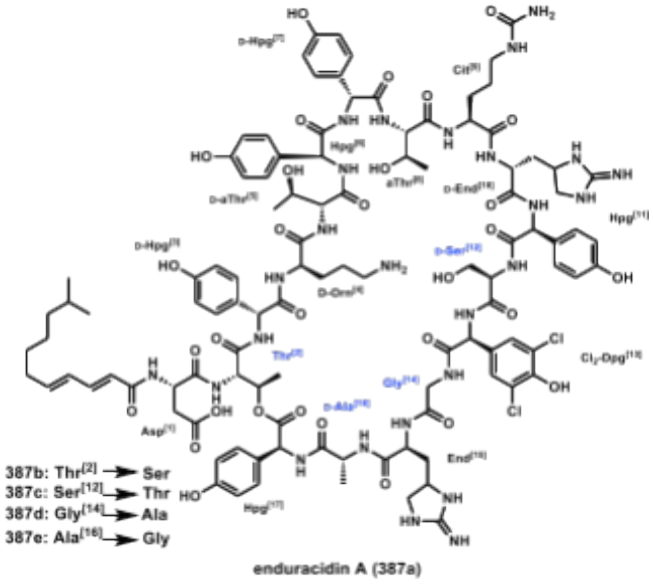


Figure 94. Enduracidin A analogs **(387b**-**e**) were generated by swapping A domains in its NRPS assembly-line.

4.2.5.3 Engineering Condensation Domains

Whereas the clear substrate gatekeeping domain in an NRPSs is the A domain, engineering efforts can be further hindered if other domains are unable to process nonnative substrates. It has been demonstrated that this issue is especially prevalent in the case of C domains,[762–764] which is reflected in varied selectivity during the condensation reactions. Often, C domains exhibit a high degree of selectivity with respect to the acceptor substrate,[765–772] but can also be selective

for changes in the donor substrate.[326,765,773,774] However, in contrast to A domains where a general guideline of specificity rules has been established for increased engineering success,[764,775] no such rules currently exist for engineering C domains. Folger et al. developed a yeast display platform[776] for the rapid analysis and engineering for C-domain substrate specificity.[777] In this strategy, full-length functional NRPS modules were displayed on the surface of yeast to link the condensation activity of the displayed C domain to a fluorescent output. FACS was then used for screening variant libraries and selection. This workflow was used on the model surfactin synthetase C domain (SrfA-C) to reprogram the substrate selectivity from amino acids to fatty acid substrates.

Condensation domains responsible for initiation and termination of lipopeptide biosynthesis have been explored as domain swapping candidates to generate lipopeptide analogs. Incorporation of the lipid starter unit for lipopeptides such as daptomycin, a process known as lipoinitiation, is gatekept by a starter condensation ($C_{starter}$ or C_s) domain. The acyl moiety is delivered to the C_s domain either as an acyl-ACP or as a free-standing acyl-CoA followed by condensation with the first amino acid. To explore C_s domain promiscuity, Zhong et al. identified two NRPS-derived peptides, the linear holrhizin A (**388a**) and the cyclized rhizomide A (**389a**), which differ in the length of the starter unit alkyl chain (C8 for **388a** and C2 for **389a**).[326] The NRPSs involved in biosynthesis of **388a** and **389a** were identified, and the two C_S domains were shown to have high sequence identity, despite the different acyl chain length preference. Switching the C_S domain between the two NRPSs lead to the corresponding switch in starter unit size and afford the C2-holrhizin A (**388b**) and C8-rhizomide A (**389b**) (Figure 95). Structural-guided mutagenesis mapped several key residues that are determinants of acyl chain length specificity. Mutant C_S domains that can accept chain length from C2 to C18 for lipoinitiation were identified, which can be useful domains for lipopeptide engineering.

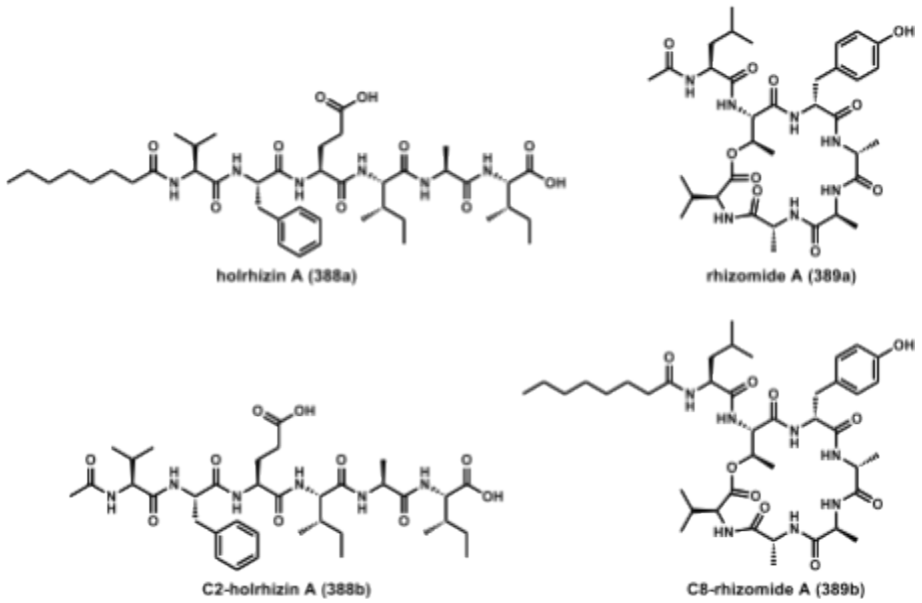


Figure 95. Swapping of NRPS starter condensation domains (C_S) enabled generation of new lipopeptides.

The amidation of the *C*-terminal carboxylate with putrescine (**205**) in the glidonin family of lipopeptides (such as **390**) was shown to be catalyzed by the terminal C domain in the NRPS. The putrescine amide is important not only for the bioactivity, but also for enhancing the water solubility of the lipopeptide, suggesting that the introduction of the putrescine amide as a C-terminal cap could be used to improve chemical and bioactivities of other lipopeptides.[778] The C domain-mediated incorporation of putrescine is promiscuous with respect to both the NRPS substrate and the amine. C domain swapping with the rhizomide A and holrhizin A NRPSs enabled the successful incorporation of putrescine into both **388a** and the linear form of **389b**, affording the derivatives **388c** and **389c**, respectively (Figure 96). Alternative diamines 1,7-diaminoheptane, 1,6-diaminohexane, and 1,5-diaminopentane could all be used as the amine nucleophile for *C*-terminal amidation.

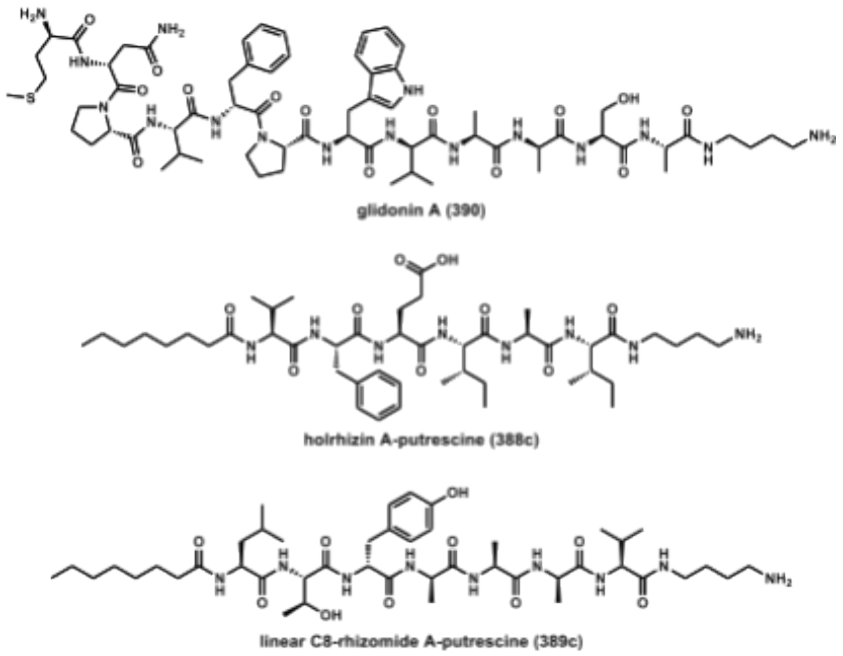


Figure 96. Promiscuity of C-terminal condensation domain allows the selective incorporation of putrescine into nonribosomal peptides.

4.2.6 Examples of Redirecting RiPP Biosynthesis

Many reports in the diversification of RiPP structures have explored the promiscuity of PTM enzymes toward unnatural precursors. Enzymes catalyzing head-to-tail macrocyclization, pyridine formation,[779] radical-based oxidative crosslinking,[780–784] cyclodehydration,[785–788] epimerization,[789] and methylation[790,791] have become popular candidates for generating modified RiPPs. Precursor peptides can be generated using *in vitro* or *in vivo* translational machinery[792–795] and the flexizyme technology,[796] as well as via synthetic means such as solid phase peptide synthesis.[797] The ribosomal approach, whether in cells or using cell-free extracts, allows mutagenesis of the DNA template encoding the core peptide to be incorporated as part of the workflow. In addition, longer precursor peptides that contain the recognition elements such as RRE for interactions with the tailoring enzymes can also be translated from the ribosome.[798] On the other hand, using synthetic approaches allows nonproteinogenic amino acids to be readily introduced into the core peptide. Schmidt and Schultz showed that combining ribosome precursor peptide synthesis and suppressor-tRNA based unnatural amino acid incorporation can diversify RiPPs sequences beyond the proteinogenic amino acids.[799]

A dedicated prolyl oligopeptidase POPB implicated in the macrocyclization and leader peptide removal during the maturation of cycloamanides was used to generate macrocyclic peptides, without requiring the leader sequence.[800] To generate a large library of cyclic peptides using this enzyme, *E. coli* mutants expressing variations in the amanitin precursor peptides were grown in polycultures. 100 out of 127 precursor peptides that range from seven to 16 residues in length, were accepted by the POPB to produce macrocyclic peptides. In another application of PDB, precursor peptides analogs of presegetalin B1 were fed to PCY1, a segetalin macrocyclase for macrocyclization, to afford segetalin B1 analogs. An assessment of the length of the C-terminus sequence indicated that only three residues were required for substrate recognition. Subsequently, PCY1 was shown to accept a wide array of substrates containing unnatural and non-amino acids with catalytic efficiency greater than the previously characterized POPB macrocyclase.[801,802] Several representative products (**391a**-**e**) of PCY1 are shown in Figure 97.

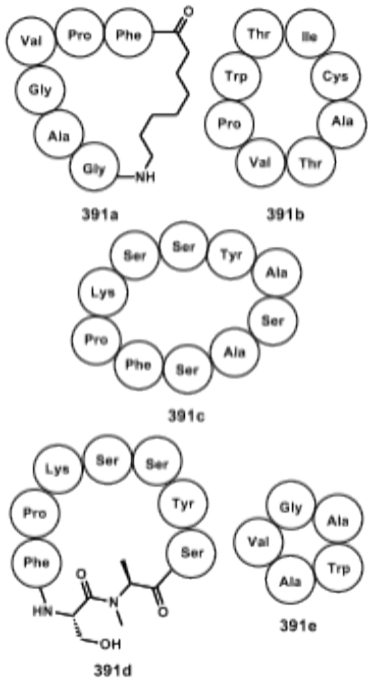


Figure 97. PCY1 is a promiscuous head-to-tail peptide macrocyclase.

Lanthipeptides are a large class of RiPPs defined by the presence of lanthionine and methyllanthionines residues, which are generated by the dehydration of serine and threonine residues, respectively, followed by cyclization from a cysteine residue in the peptide. A DNA library of 6,000 lanthipeptide precursors fused to the leader sequence of nisin A (**21**)

was generated via combination of units of rotationally constrained fragments from 12 natural lanthipeptides.[803] Following modification by PTM enzymes, a bioactivity-based assay was used to identify over 100 hits with enhanced activity against pathogenic bacteria. From this large-scale effort, it was found that residues flanking the serine, threonine, and cysteine involved in thioetherification are less flexible to change due to recognition by enzymes that process the lanthipeptide precursors. In a study aimed to overcome such limitations, numerous different precursor peptides were found to be colocalized with a class II lanthipeptide synthetase such as SyncM in different organisms. [787] Using PDB with different precursor peptides, SyncM was shown be the most promiscuous lanthipeptide-generating enzyme to date and can function independently of any specified residues near the crosslinking sites. Macrocycles up to 14 residues with diverse topologies, such as **392a**-**c**, were efficiently cyclized in both directions along the peptide (Figure 98).

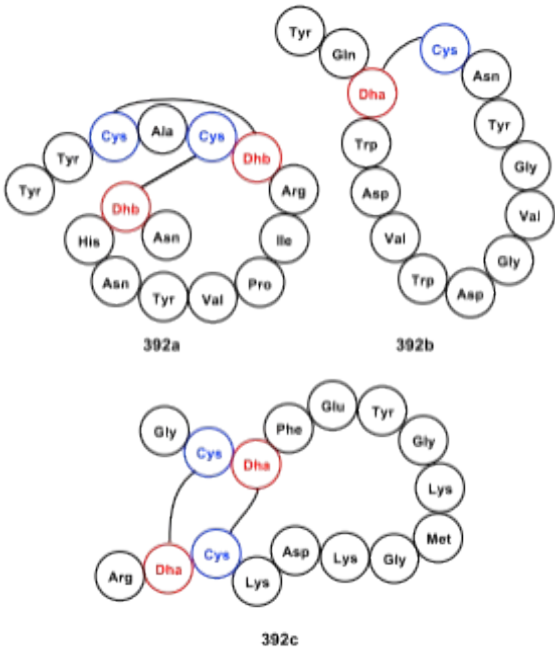


Figure 98. SyncM performs lanthipeptide cyclization with diverse substrates. The lines indicate lanthionine and methyllanthionines linkages.

The radical SAM enzyme PapB is a lanthipeptide-maturation enzyme involved in the formation of six thioether crosslinks between cysteines and the β and γ carbons of aspartic/glutamic acid residues in the precursor peptide PapA.[804] Substrate scope assays indicated that while the leader peptide was crucial for acceptance, the residues between cysteine and the thioether crosslinking site could vary from three to six residues, indicating that specificity is based on side-chain

recognition rather than primary sequence. Using PapB, a thioether analog (in place of its disulfide) of the FDA-approved therapeutic agent octreotide was generated.[805]

The two intramolecular crosslinks involving the two tryptophan residues in the precursor peptide DarA (**393**) are catalyzed by a single radical SAM enzyme DarE, followed by proteolysis to give darobactin (**167**). Müller and coworkers modified DarA to include different residues at all positions except the tryptophan residues at positions one and three, which resulted in the generation of 17 new analogs of **167** when fed to *E. coli* expressing DarE (Figure 99).[806] Interestingly, alanine and arginine were accepted at position 5 in place of lysine and can be crosslinked with tryptophan at position 3. While most analogs could not be isolated and assayed due to low titers, the isolated derivative $W_1N_2W_3S_4K_5S_6W_7$ (**167b**) was more potent compared to **167** when tested in antibacterial assays. A subsequent campaign by Mitchell and coworkers probed acceptance of DarA variants with different residues at crosslinking sites (i.e. positions 1, 3, and 5).[807] Of the 51 different variants modified by DarE, most scaffolds were desaturated or singly crosslinked, and only 10 products retained both intrapeptide crosslinks present in **167**. From this study, the authors concluded that DarE preferentially forms C-O ethereal crosslinks between two aromatic side chains bewteen residues 1 and 3, and C-C crosslinks between aromatic and nonaromatic side chains between residues 3 and 5.

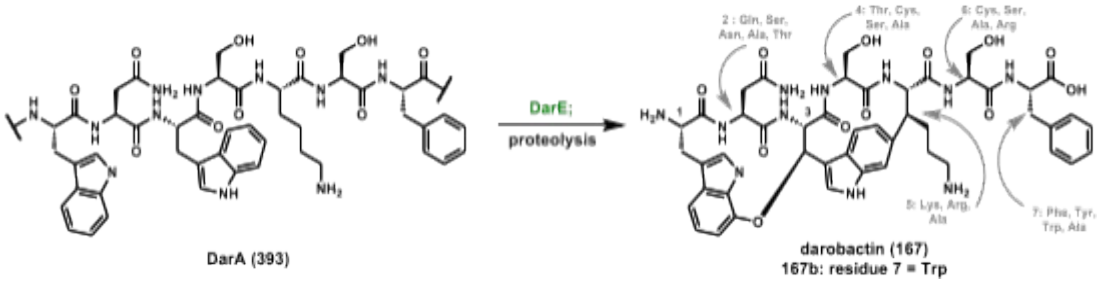


Figure 99. DarE is an oxidative crosslinking maturation enzyme. Residue substitutions tolerated by DarE are noted on the right.

The newly discovered spliceotide family of RiPPs are featured by their characteristic α-keto-β-amino acids-containing peptide backbone, which are formed by radical-SAM-dependent spliceases via tyramine excision.[808,809] The first discovered instance of this novel maturation enzymes was by the Piel group in their characterization of the radical-SAM enzyme PlpX and its auxiliary protein PlpY. The two enzymes act in tandem to modify the Nif11-like precursor peptides

PlpA2 and PlpA3 encoded within the same biosynthetic locus (Figure 100A).[808] Precursor-peptide variants of PlpA3 were assayed for acceptance by PlpX and PlpY. The study showed that the enzymes are well-tolerant of mutations to the core peptide and only require a minimal 11-residue sequence with a conserved XYG motif, in which X can be seven different residues and Y is the tyrosine to be excised. Using the enzyme as a biocatalyst, Piel and coworkers demonstrated bioorthogonal conjugation of the ketoamide-containing peptides with a fluorogenic dye.[808] Further, the group used PlpX and PlpY to generate **394**, which is a close homolog of the α-keto-β-amino amide containing CVS 4453 (**395**), a compound that inspired the clinically approved protease inhibitor boceprevir (**396**). Subsequent engineering studies indicated that incorporation of a minimal 10-residue splice tag into various proteins, such as maltose binding protein and glutathione *S*-methyltransferase, enabled site-selective ketoamide formation in the proteins expressed from *E. coli*. Tyramine splicing can take place whether the tag is fused at the *N*- or *C*-terminal, or inserted internally into a protein sequence, although greatest yields were observed for *N*-terminal tags.[810] These results indicate that the splice tag may be used as a bioorthogonal handle for protein labeling.

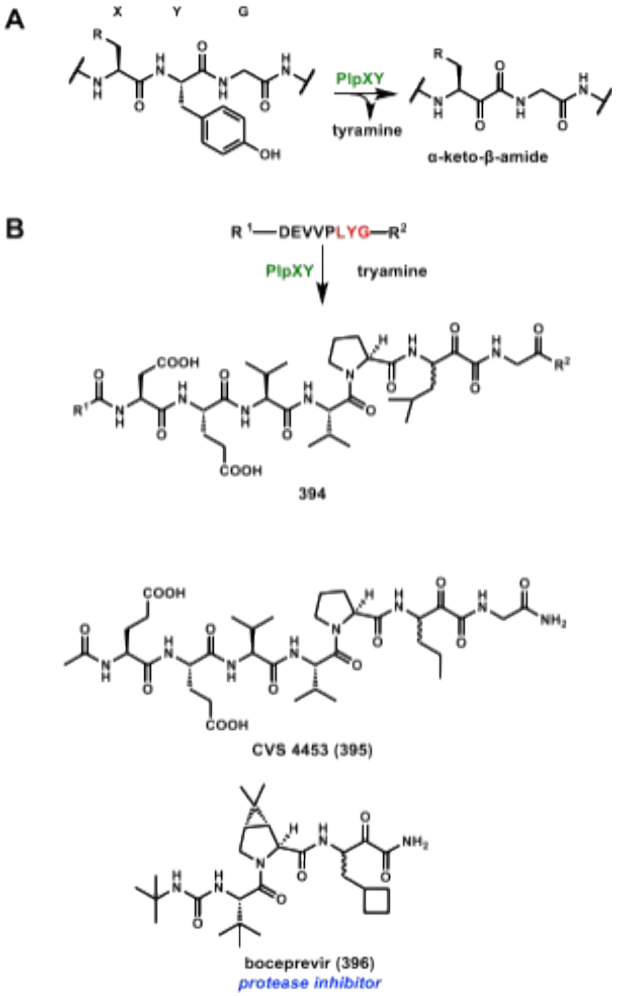


Figure 100. PlpX and PlpY excise tryptamine from precursor peptides. (A) Scheme of the reaction and (B) a precursor peptide variant was used to synthesize **394**.

4.2.7 Examples of Redirecting Terpene Biosynthesis

Redirecting terpene biosynthesis using unnatural precursors *in vivo* is much more challenging than polyketides and peptides. This is due to the difficulties associated with preparing analogs of isoprenyl pyrophosphates and the cell impermeability of the pyrophosphates. In addition, the native terpene precursors such as IPP, DMAPP, as well as FPP, GGPP and squalene, are essential for cellular function, and cannot be depleted for mutasynthesis. As a result, diversification of terpene biosynthesis is primarily focused on altering the cyclization through engineering of the terpene synthase (TS)/terpene cyclase (TC), and the combinatorial mix-and-matching of tailoring enzymes such as P450 monooxygenases (see section 4.3.2).

Unnatural terpene precursors have, however, been used as tools to probe TC mechanisms *in vitro*, whether through enzyme inhibition for crystallographic analysis or redirecting the carbocation-driven cyclization outcomes.[811,812] Many analogs of the linear substrates can be accepted by TCs, including those containing halogens, heteroatoms,[813–815] cyclopropanes,[816] altered branching methyl regioisomers,[817] or altered double bond positions.[818,819] While some terpene analogs can follow the same cyclization steps as the natural substrate,[817] in most cases variations in cyclization are observed due to both steric and electronic effects. One such example involves the TC Bot2 from *Botrytis cinerea* as shown in Figure 101. Whereas the sesquiterpene cyclase normally synthesizes presilphiperfolan-8-β-ol (**397**) from FPP, analogs of FPP (**398a-d**) were cyclized into products **399a-d** that are nearly unrecognizable compared to **397**. To generate unnatural GGPP analogs that are more challenging to prepare synthetically, Kirschning and coworkers supplied unnatural FPP building blocks to GGPP synthase from *Streptomyces cyaneofuscatus*.[820] Resulting GGPP analogs were then supplied as precursors to spata-13,17-diene synthase to generate three cyclododecane compounds and a macrocyclic ether.

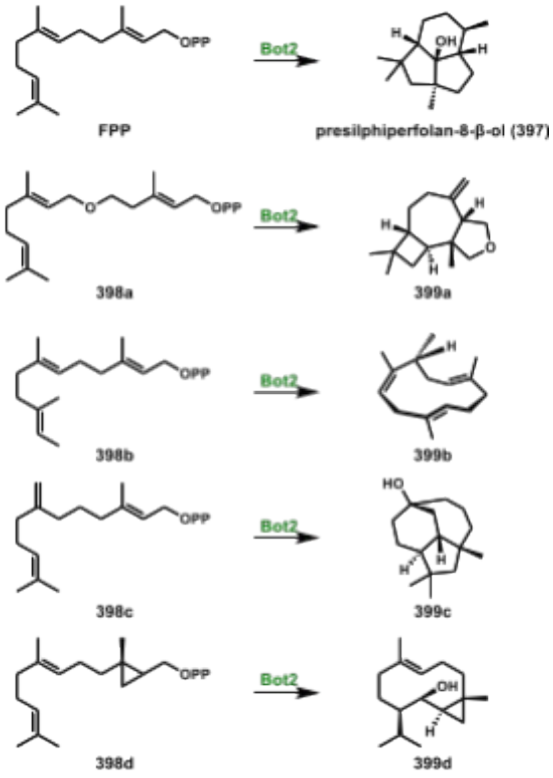


Figure 101. Cyclization of FPP and analogs by the sesquiterpene cyclase Bot2.

Several approaches have emerged in recent years to metabolically generate nonnative pyrophosphate precursors. The first is the isopentenol utilization pathway (IUP) already discussed in section 3.2.3.4. This strategy can be considered a form of PDB to synthesize isoprenyl pyrophosphate analogs from the corresponding alcohols. In addition, heterologous expression of *S*-adenosylmethionine (SAM)-dependent IPP methyltransferases from *Streptomyces monomycini* led to *in vivo* methylation of IPP to give C6 (**400**, **401**) and C7 analogs (**402**, **403**) of IPP and DMAPP in *E. coli* (Figure 102A).[821] These substrates can be accepted by GPP and FPP synthases to afford C11, C12, C16, and C17 prenyl pyrophosphates. Subsequently, these methylated linear precursors can be cyclized by terpene cyclases to form atypical terpenes, such as polymethylated zeaxanthin analogs including **239b**. Major products contained one to three SAM-derived methyl groups on the carotenoid backbones.

The homomevalonate pathway discovered in *Lepidoptera* uses propionyl-CoA instead of acetyl-CoA to synthesize C6 precursors such as homoisopentenyl pyrophosphate and homodimethylallyl pyrophosphate, which are readily incorporated into C16-C18 analogs of FPP.[822] Heterologous expression of this homomevalonate pathway, a propionyl-CoA ligase, and terpene cyclases in *E. coli* afforded production of several new C16 terpenes detected via mass analysis of crude extracts.[823] However, due to close structural resemblance between these methylated analogs and the natural C15 products, isolation and characterization of the new compounds was challenging. Another terpene methylation strategy uses *Sp*SodMT from sodorifen biosynthesis.[824] This unusual enzyme methylates and cyclizes FPP to the 16-carbon presodorifenyl diphosphate (PSPP, **404**), which can be converted into a variety of C16 terpenoid scaffolds such as **405-407** (Figure 102B). *Sp*SodMT was engineered to produce ten C16 pyrophosphate building blocks, which would then be processed into 24 different primary and tertiary alcohols, and olefins by endogenous yeast metabolism. A novel pathway to biosynthesize noncanonical C16 terpenoids was recently identified by Reuter, Barra, and coworkers, through the identification of a unique family of methyltransferases that catalyze the formation of (*2E,7E*)-6-methyl-farnesyl diphosphate from FPP, followed by various cyclization reactions catalyzed by TC.[825] Five new C16 terpenoid natural products were identified.

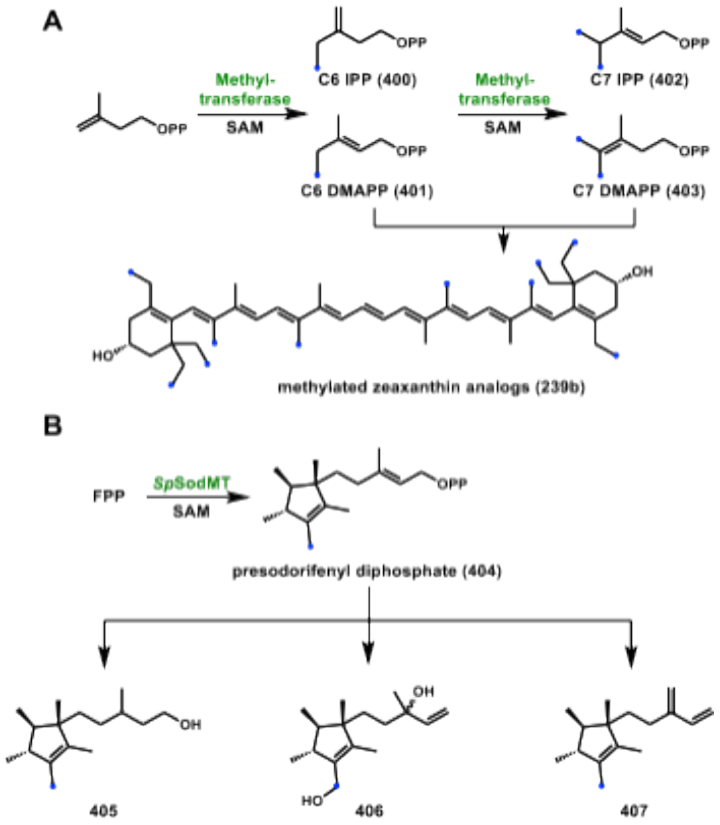


Figure 102. Methylated isoprenyl pyrophosphates as building blocks for unnatural terpene synthesis. (A) IPP methyltransferases facilitates synthesis of methylated zeaxanthin analogs; (B) *Sp*SodMT and engineered mutants facilitate synthesis of C16 terpenoids.

4.2.8 Examples of Redirecting Alkaloid Biosynthesis

4.2.8.1 PDB and Mutasynthesis using Tryptophan Analogs

L-Tryptophan is a key building block to many classes of alkaloids with its reactive indole ring remaining largely intact through the various biosynthetic pathways. One such example is the class of MIAs, which are biosynthesized through the universal precursor strictosidine (**50**).[422] For this reason, the feeding of halogenated or substituted L-tryptophan and tryptamine analogs has been extensively applied to generate alkaloid derivatives. For example, Clark and coworkers supplied fluorinated L-tryptophan analogs to cultures of *Bacillus atrophaeus* IMG-11 and isolated analogs of the tryptamine-containing metabolites bacillamide A and C.[826] The 6-fluoro substitution improved antialgal activity of derivatives. One utility of using halogenated L-tryptophan precursors (Cl, Br and I) is the resulting halogenated aromatic ring can serve as a

reaction handle in metal-based cross-coupling chemistry for further semisynthetic modifications.[826–829] This was demonstrated with the PDB of 5′-bromo-violacein (**150b**) and 5′-bromo-deoxyviolacein derivatives in *E. coli* (Figure 103).[830] Suzuki-Miyaura cross-coupling performed directly in cell cultures extracts with a series of aryl boronic acids enabled the preparative synthesis of 5′-aryl-violacein products (**150c**).

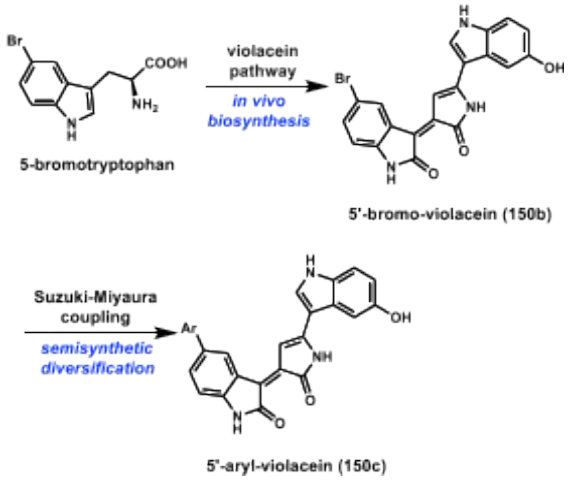


Figure 103. Cross-coupling of halogenated violacein derived from PDB using tryptophan analog.

Caputi and O'Connor generated plant-derived MIA derivatives by incubating various halogen- and methoxy-decorated tryptamine analogs and secologanin with strictosidine synthase (STR) from *Catharanthus roseus.*[831] This provided a substrate scope assessment of STR towards unnatural tryptamine analogs. Biotransformation of the modified strictosidine (**50**) was then performed in *N. benthamiana* engineered to heterologously express late-stage enzymes, resulting in the synthesis of analogs of downstream MIAs such as alstonine (**213**) and stemmadenine acetate (**408**) (Figure 104). While STR was tolerant of tryptamine analogs, the downstream enzymes required to generate **213** and **408** only accepted the fluorinated analogs of **50**. During efforts to engineer the *de novo* production of MIAs from *S. cerevisiae*, 7-halotryptamine analogs were supplied to the yeast strains to result in the production of a fluorinated serpentine (**212**), as well as chlorinated and fluorinated analogs of corynantheidine (**409**).[832]

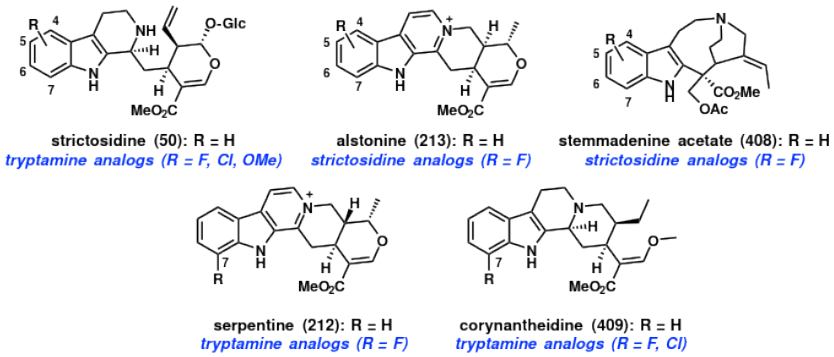


Figure 104. Biosynthesis of halogenated MIAs through feeding of halogenated tryptophan or tryptamine. Numbers on ring labeled represent sites of diversification.

4.2.8.2 Engineering Supply of Tryptophan Analogs

Several enzymes that can halogenate tryptophan at the 5, 6 or 7 positions of the indole ring have been reported, including PyrH[833,] and *Xszen*FHal (XsHal)[834] at C_5; SatH[835] and Thal at C_6;[836] and PrnA[837] and RebH[838] at C_7. The panel of L-tryptophan halogenases with different regioselectivities has been used to access diverse indole-derived building blocks for synthetic chemistry,[839] as well as natural product structural diversification through PDB.[840] One example is work by Salas and coworkers in the synthesis of chromopyrrolic acid analogs in *S. albus* (Figure 105).[841] RebH, which is associated with the biosynthesis of rebeccamycin (**410**), is involved in chlorination of L-tryptophan to give 7-chlorotryptophan, which is then incorporated into 11,11'-dichlorochromopyrrolic acid (**411**). By swapping RebH with ThaI or PyrH, the chlorination position around the indole ring in chromopyrrolic acid was changed to C_{10} or C_9, respectively.

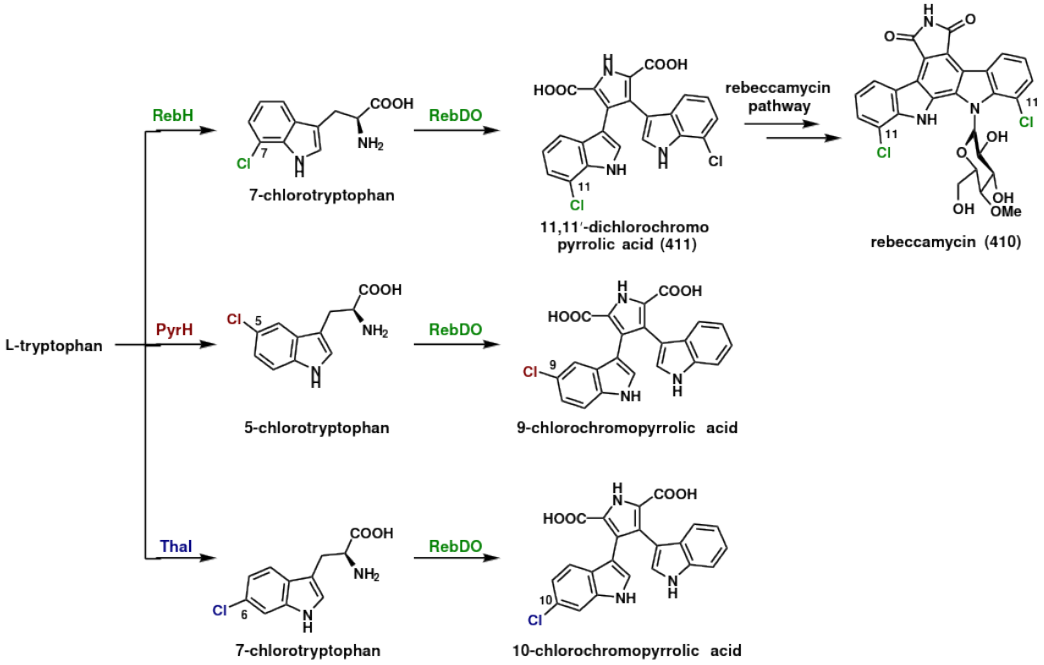


Figure 105. Swapping tryptophan halogenases led to the production of **411** analogs.

An application of this regioselective halogenation was demonstrated with the biosynthesis of halogenated brassinin (**412a**) analogs as antifungal pesticides (Figure 106).[842] The enzymes associated with the synthesis of brassinin were characterized to be relatively promiscuous for L-tryptophan analogs.[843] Coexpression of PyrH and a flavin reductase partner RebF with the brassinin pathway in *N. benthamiana* resulted in accumulation of 5-chlorobrassinin (**412b**) in leaves. Expressing RebH alongside RebF resulted in accumulation of 7-chlorobrassinin (**412c**). Supplying the hairy roots of *N. benthamiana* expressing RebH/RebF pair with potassium bromide was required for detection of 7-bromobrassinin (**412d**), because of low bromide levels in plants. **412b** was found to not only be more potent than **412a** against the generalist pathogen *B. cinerea,* but also be comparable to pyrimethanil which is an active component of commercial pesticides.

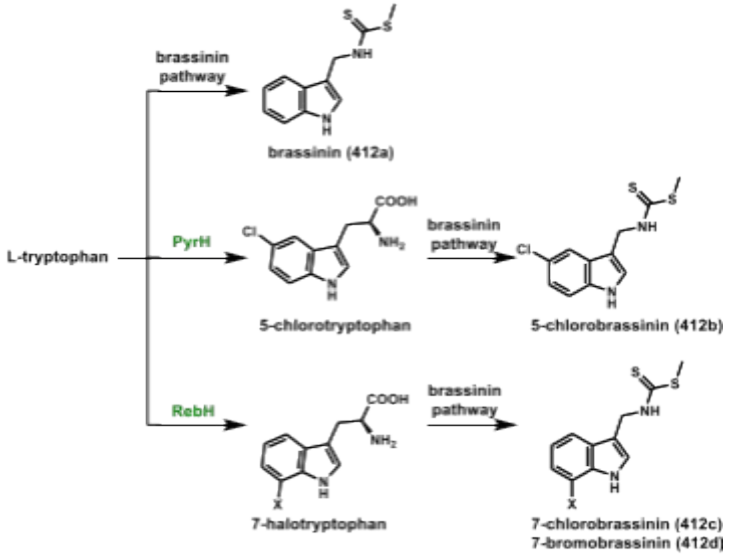


Figure 106. Using different L-tryptophan halogenases led to the production of **412b-d**.

The O'Connor group integrated the RebH/RebF and PyrH/RebF pairs into *C. roseus.*[840] In addition, a mutant STR was introduced to enhance incorporation of the halo-L-tryptamines generated from the native tryptophan decarboxylase (TDC). The RebH/RebF gain-of-function plants were found to accumulate 12-chloro-19,20-dihydroakuammicine in hairy roots. Given that 19,20-dihydroakuammicine is not a major alkaloid in wild-type plants, it is believed that downstream tailoring enzymes are unable to process this 12-chloro analog. Similarly, PyrH/RebF gain-of-function plants were found to accumulate a variety of MIA analogs, such as 10-chloroajmalicine and 15-chlorotabersonine. As an alternative to expressing halogenases for MIA diversification, Zhang and Jensen supplied fluorinated, chlorinated, and brominated indoles to a yeast strain that produced MIAs, and relied on the yeast tryptophan synthase (Trp5), a pyridoxial-5'-phosphate (PLP)-dependent enzyme that synthesizes L-tryptophan from indole and serine, to synthesize the halogenated L-tryptophan precursors (Figure 107).[429] The authors noted the production of nineteen analogs of serpentine (**212**) and alstonine (**213**).

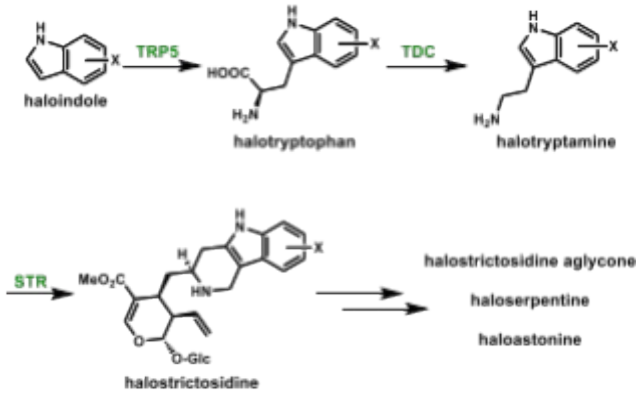


Figure 107. Biocatalytic access to MIA analogs via unnatural supply of haloindoles as a precursor to halotryptophan.

Using these halogenases, Alper and coworkers optimized *E. coli* to produce chlorinated and brominated L-tryptophans.[844] These precursors were further modified to analogs of 1-acetyl-3-carboxy-β-carboline. Goss and coworkers developed the GenoChemetic platform to combine halogenated natural products derived from L-tryptophan halogenases with an array of cross-coupling methodologies.[827] Brominated analogs of cystargamide, barrettin, and pacidamycin D were semisynthesized by Sonagashira,[845845] Buchwald–Hartwig,[846] and Heck[847] reactions, respectively.

4.2.9 Examples of Redirecting Tailoring Enzymes

While PDB is typically associated with providing substrate analogs for the core biosynthetic enzymes, exploring the cofactor/cosubstrate specificity of tailoring enzymes can also lead to alteration to the natural product structures. These include the electrophilic SAM used by methyltransferases, UDP-sugars for glycosyltransferases, activated acyl groups such as thioesters used by acyltransferases, etc.

4.2.9.1 SAM Analogs for Methyltransferases

For the vast majority of methyltransferases in metabolism, SAM is used as the electrophilic methyl donor. Nucleophiles such as alcohols, amines, thiols and electron-rich carbons can be regioselectively methylated by dedicated methyltransferases. SAM is naturally synthesized from methionine and ATP by SAM synthetase.[848] To explore methyltransferases for unnatural alkylation reactions, it is necessary to provide SAM analogs that are modified at the methylsulfonium position. In an early example, a 5′-*N*-mustard variant of adenosine that is known as a DNA alkylating

reagent, was accepted by the methyltransferase RebM to generate new rebeccamycin analogs.[849] More recent examples have focused on the enzymatic generation of SAM analogs followed by alkyl transfer to nucleophiles. For example, fluorinated SAM analogs have been generated towards site-selective fluoroalkylation. Seebach and coworkers used halide methyltransferases to prepare fluorinated SAM (F-SAM, **413**) starting with *S*-adenosylhomocysteine (SAH, **414**) and fluoromethyliodide (**415**) (Figure 108A).[850] **413** generated *in situ* was used by methyltransferases to fluoromethylate various nucleophiles. To limit the degradation of **413** to homoserine lactone, Dong and coworkers biocatalytically synthesized fluoro-decarboxy-SAM (F-dcSAM, **416**) starting with decarboxy-SAH (dcSAH, **417**) (Figure 108B). **416** was readily accepted by different methyltransferases to fluoromethylate natural product scaffolds, such as caffeic acid (**228**) to fluoroferulic acid (**229b**), and carminomycin (**418**) to fluorodaunorubicin (**419**).[851] Kinetic analysis indicated that several enzymes accepted **407** with higher efficiency than SAM. Using a different approach to circumvent degradation of **413**, Booker and coworkers chemoenzymatically generated Te-adenosyl-L-(fluoromethyl) homotellurocysteine, which is more resistant to epimerization and homolactone formation at 37 °C in the pH range of 2–12.[852] The unusual cofactor enabled regioselective production of fluorooxaline and **419**. Due to the larger size of tellurium compared to sulfur, the K_M is notably higher compared to that of SAM.

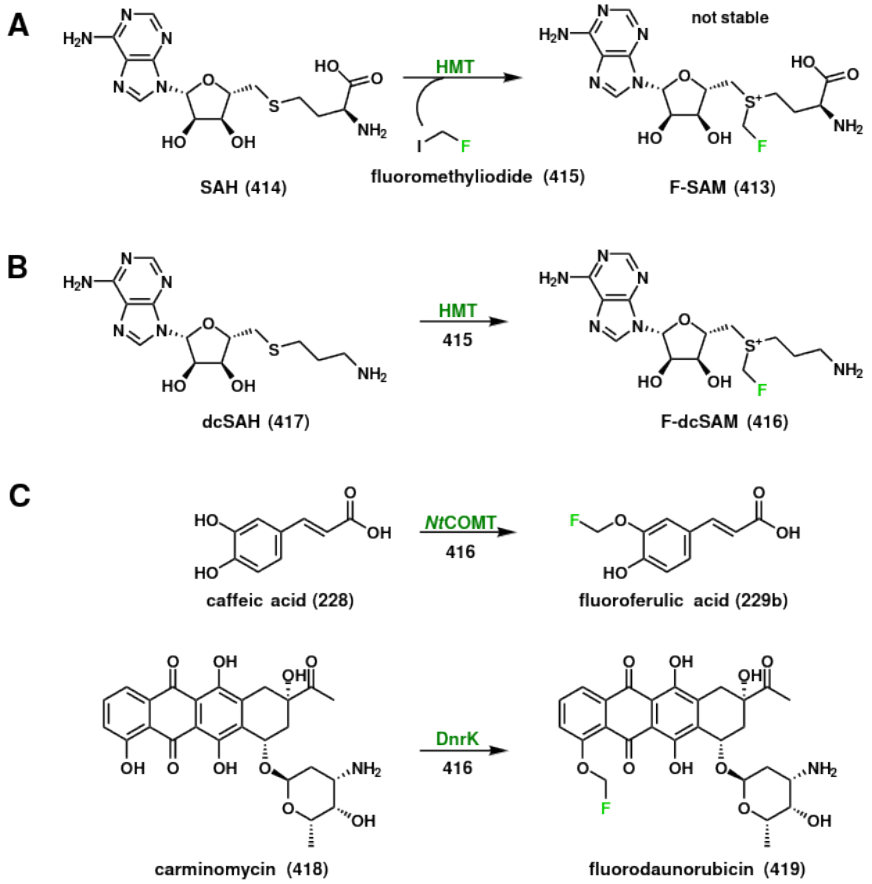


Figure 108. The use of fluorinated SAM analogs for fluoromethylation of natural products. Biocatalytic production of (A) **413** and (B) **416**. (C) Fluoromethylation of natural products.

SAM cofactors with ethyl,[853,854] allyl, crotyl, propargyl and benzyl[855] sulfonium have been used to alkylate nucleophilic centers in the presence of natural or engineered methyltransferases. NovO and CouO have been used to transfer ethyl and fluoroethyl groups to coumarin scaffolds.[856] Human methionine adenosyltransferase (hMAT2A) has been applied to generate unnatural SAM analogs *in vivo* using synthetic methionine derivatives. An engineered hMAT2A was used together with the *N*-methyl glycopeptide methyltransferase MtfA from the chloroeremomycin biosynthetic pathway, to generate several vancomycin (**3**) analogs, such as 4-fluorobenzyl and 3,3-difluoroallyl variants.[857] Similarly, regioselective *O*-allyl and *O*-ethyl analogs of rapamycin (**420**) were generated using the hMAT2A Ile332Val mutant and the methyltransferase RapM (Figure 109).[858] Carboxy-*S*-adenosylmethionine (carboxy-SAM) is a naturally occurring carboxymethylation cofactor. Micklefield and coworkers used the synthase CmoA to generate carboxy-SAM and carboxy-*S*-

adenosylethionine,[859] which were accepted by an engineered *N*-methyltransferase to afford various *N*-carboxymethylated and *N*-carboxyethylated THIQ compounds.

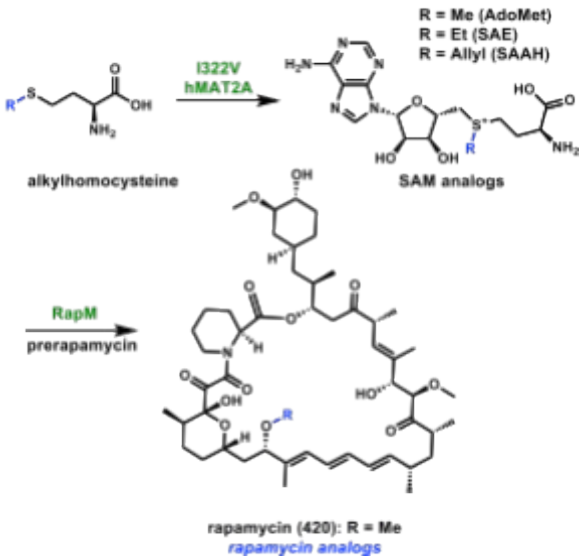


Figure 109. Generation of **420** analogs using engineered SAM synthetase and RapM.

4.2.9.2 Unnatural Sugars for Glycosyltransferases

The sugars that decorate natural product aglycones dramatically alter the physical properties, and are often indispensable in endowing biological activities.[860] Natural product can be glycosylated with monosaccharides as in staurosporine[861] and doxorubicin,[862] with disaccharides as in avermectin[223] and vancomycin (**3**),[863] and with oligosaccharides such as avilamycin[864] and saccharomicin.[865] While common sugars such as D-glucose, D-mannose, etc. are used to modify natural products, BGCs often encode dedicated biosynthetic enzymes to generate unusual deoxysugars for subsequent glycosylation. Dedicated glycosyltransferases transfer the sugars, in their thermodynamically activated form of NDP-sugars, to nucleophiles present in natural products, most commonly alcohols. Much knowledge on the biosynthesis of unnatural sugars have been accumulated in the last twenty years and are reviewed by Liu and coworkers.[866] With such information in hand, mutasynthesis to changing the sugars attached to natural products can be performed by inactivating the genes involved in biosynthesis of the natural sugar, and expressing enzymes involved in alternative sugar biosynthesis. Examples of such efforts include the biosynthesis of analogs of doxorubicin (**421**),[867] gilvocarcin V (**422**),[868] YC-17, etc. (Figure 110).[869]

This approach is particular useful when the glycosyltransferase that modifies the natural product is promiscuous towards different sugars. This was the case with the glycosyltransferase ElmGT from the elloramycin A (**423**) biosynthetic pathway, which glycosylates 8-demethyltetracenomycin C aglycone with TDP-L-rhamnose.[870] Salas and coworkers integrated plasmids bearing unnatural sugar biosynthesis genes into *S. albus* to produce L-olivosyl-, L-rhodinosyl, and D-olivosyl-elloramycin analogs.[871] Since then, a wide variety of **423** analogs containing unnatural sugars have been reported using this method.[872–874] Shabaan and coworkers metabolically engineered *S. coelicolor* M1146 to express genes involved in 8-demethyltetracenomycin C biosynthesis alongside ElmGT, and introduced sugar biosynthetic pathways to generate 4′-keto-D-digitoxosyl-, D-fucosyl-, D-allosyl-, and D-quinovosyl- analogs of **423**.[875]

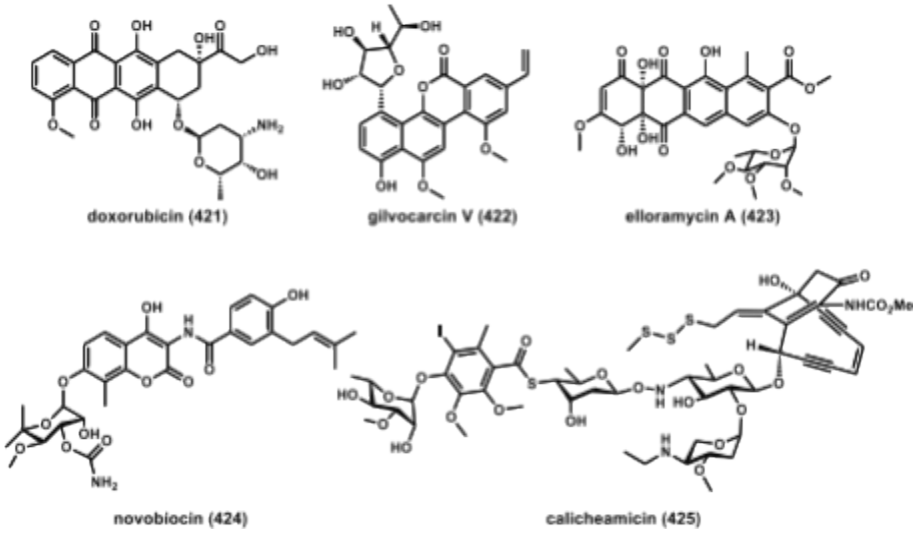


Figure 110. Select natural products targeted for sugar diversification.

One powerful approach to combinatorially vary the sugars attached to natural products is to take advantage of the substrate promiscuity of glycosyltransferases—a process termed glycorandomization or chemoenzymatic glycodiversifiication.[876] The Thorson and Liu groups have applied glycorandomization in diversification of several natural products, such as novobiocin (**424**),[877] calicheamicin (**425**),[878] methylumbelliferone,[879] avermectin,[880] pikromycin,[881] and methymycin.[882] In a classic example, Thorson and coworkers generated a large library of vancomycin (**3**) derivatives.[883] The authors first performed enzyme engineering to expand the substrate scopes of *E. coli* galactokinase (GalK)[884] and *Salmonella enterica* LT2 α-D-glucopyranosyl phosphate thymidylyltransferase (E_p)[885] to access a panel of NDP-sugars.[886] Then, taking advantage of the substrate promiscuity of the vancomycin glycosyltransferase GtfE, 31 glycosylated analogs

of **3** were generated. For instance, a 6-azido-6-deoxy-D-glucosylvancomycin (**3b**) derivative was generated via feeding of the corresponding 6-azido-6-deoxy-D-glucose and could be further elaborated via "click" chemistry to 17 additional analogs of **3**, some of which (such as **3c**) were more potent than **3** (Figure 111).

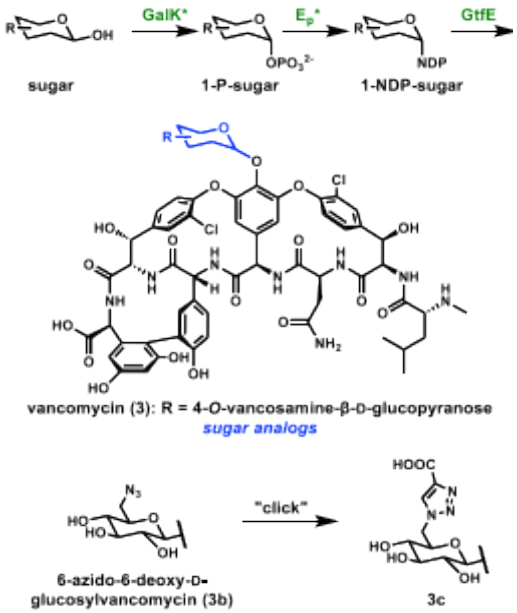


Figure 111. Biocatalytic prduction of NDP-sugars for vancomycin analog generation.

Sipanmycin A (**426a**) from *Streptomyces sp.* strain CS149 displays potent antibacterial and cytotoxic bioactivity; the β-D-xylosamine-β-D-sipanose disaccharide *O*-glycosylated to the macrolactam aglycone is essential for the bioactivity (Figure 112A). [887] To probe the substrate promiscuity of the glycosyltransferases, Salas and Olano introduced genes encoding enzymes that can synthesize TDP-D-olivose, TDP-L-amicetose, or TDP-L-digitoxose into the native strain. The new TDP-sugars replaced the β-D-sipanose to generate four analogs (**426b-e**) without the need to delete any β-D-sipanose biosynthetic genes. β-D-olivomycose and β-D-forosamine were unexpectedly incorporated upon the integration of TDP-olivose and TDP-amicetose biosynthetic pathways, respectively, and are believed to be a result of β-D-sipanose biosynthetic enzymes acting upon intermediates of the newly integrated sugar pathways. In the case with **426c** and **426e**, the new compounds were biosynthesized at levels greater than that of the native compound. The authors also attempted to replace β-D-xylosamine using mutasynthesis by deleting three genes involved in UDP-xylosamine biosynthesis, but this did not result in any new compounds likely due to the more stringent substrate scope of the first glycosyltransferase.

Along these lines, glycosyltransferases have been exchanged to alter sugar incorporation and glycosyltransfer regioselectivity. For example, GT723 is the unclustered glycosyltransferase that installs α-D-glucose onto the actinopyrone PM050463 (**427**) to give PM050511 (**428a**) (Figure 112B).[888] Using a glycosyltransferase homolog GT1507 from piercidin pathway with 65% identity, a variety of UDP-hexoses were transferred to PM05463, to result in the formation of the *N*-acetyl-D-glucosamine, *N*-azidoacetyl-D-glucosamine, or D-xylose analogs of **428a**, which are **428b, 428c.** or **428d**, respectively. GT1507 had additional substrate promiscuity to glycosylate actinopyrone metabolic intermediates preceding **427**, which led to the formation of 14 non-natural actinopyrone analogs.

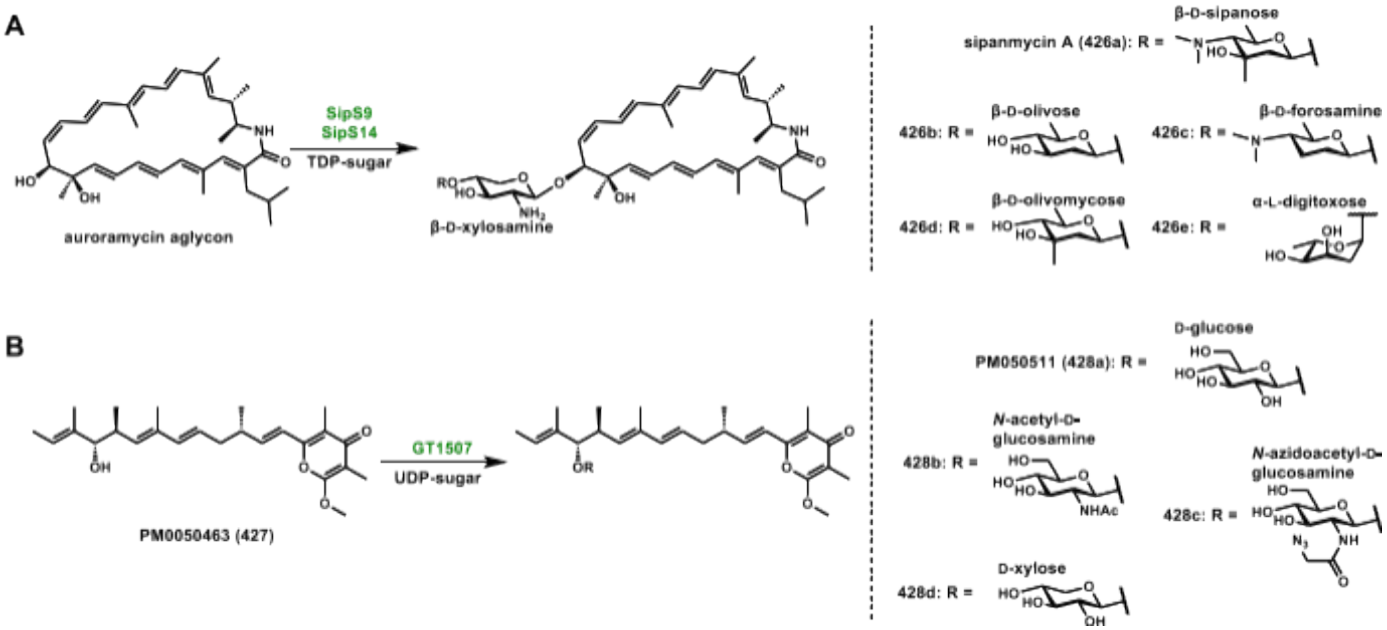


Figure 112. Modifying the sugar groups on polyketide scaffolds. (A) **426a** as an example; (B) **428a** as an example.

By exchanging glycosyltransferases from homologous pathways, Liu and coworkers generated a polyether compound with potent activity against bladder cancer.[889] Lenoremycin (**429**), endusamycin (**430**), and CP-80219 (**431**) are glycosylated polycyclic ethers with moderate bioactivity against bladder cancer. These compounds share identical aglycone structures, but each vary in the regioselectivity of methylation and glycosylation modifications (Figure 113A). From SAR studies, the distance between the sugar unit and the carboxylic acid at C_1 is correlated with potency, in the decreasing order of **429** (C_{11}), **430** (C_{15}), and **431** (C_{27}). **429**, however, possesses nonnegligible toxicity which is undesirable for use as a chemotherapy lead.[890] Aiming to keep the glycosylation at C_{11} in **429** while introducing the C_{10} methylation

present in **430**, the authors first inactivated the EndP3 in the BGC of **430**, which is responsible for C_{15} hydroxylation. The EndG5 glycosyltransferase was then swapped with the LenG5 glycosyltransferase from BGC of **429**, which resulted in the production of the hybrid compound End-16 (**432**) (Figure 113B). **432**, with a subtle C_{10} methyl group addition, was shown to be a potent inhibitor of BLCA cell proliferation with diminished side effects. Additionally, **432** resulted in very minor damage and histopathological changes to spleen and liver tissue in subcutaneous xenograft models compared to cisplatin, which has been historically associated with side effects as a first-line treatment of bladder cancer.

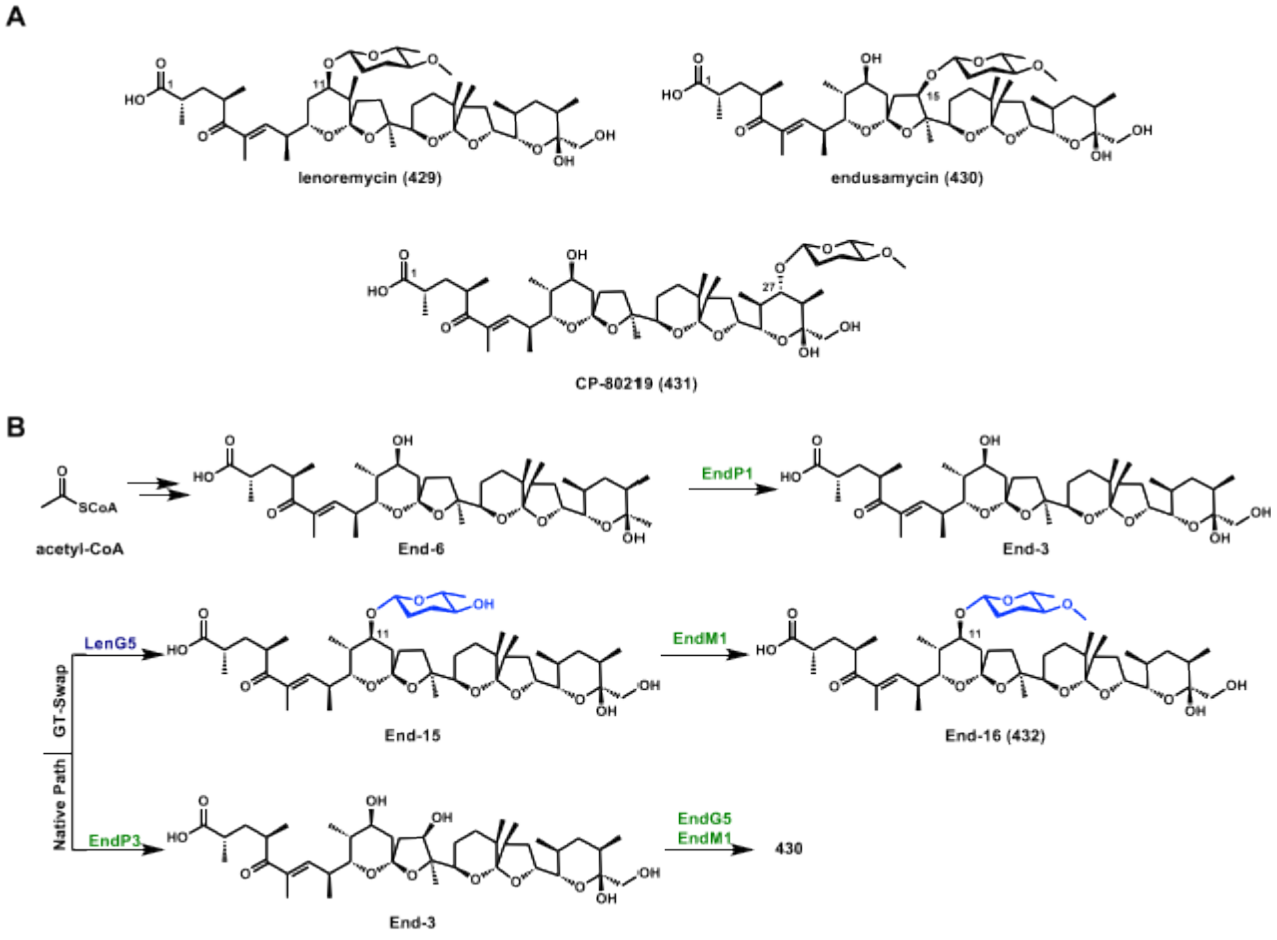


Figure 113. Glycosyltransferase swap between BGCs of lenoremycin and endusamycin enabled generation of **432**.

Glycodiversification was applied towards the generation of various aminoglycosides, a natural product family that features aminosugars linked to an aminocyclitol core.[891] Starting from paromamine (**433**), the gentamicin and kanamycin pathways differ in the sugar moiety that is transferred (Figure 114). In gentamicin biosynthesis, GenM2 catalyzes the transfer of xylose to generate gentamicin A2 (**434**), which is further modified into gentamicin C1 (**435**); while in kanamycin

biosynthesis, KanE transfers glucose to **433** to give 3''-dehydroxy kanamycin C (**436**) followed by additional modifications. Sun and coworkers swapped the native GenM2 with KanE in the gentamycin producing strain *Micromonospora echinospora* to generate six new hybrid aminoglycosides, including genkamicin C1 (**437**). Analogs were active against critical priority pathogens, such as *Acinetobacter baumanni* that is resistant to gentamicins and kanamycin B (**14**), while displaying low ototoxicity using zebrafish embryos as test organisms.

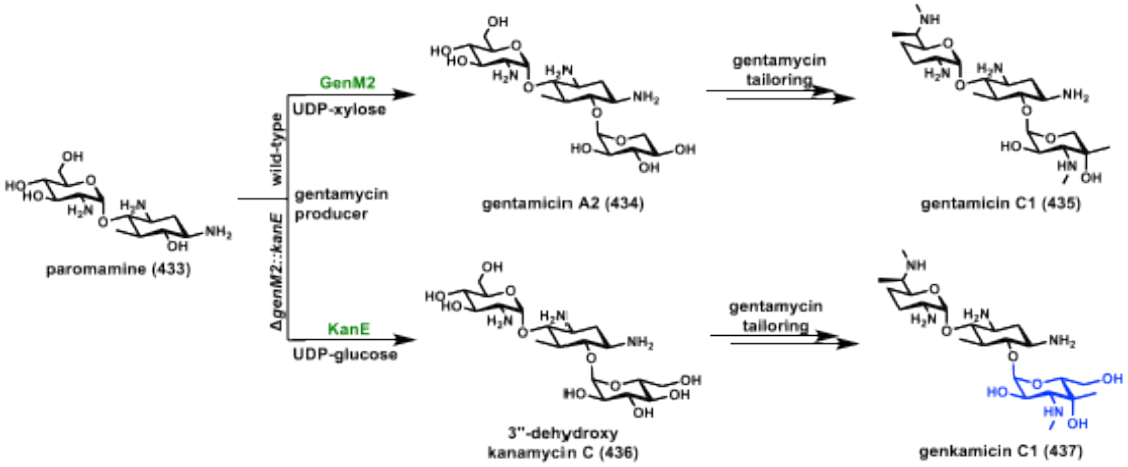


Figure 114. Genearation of aminoglycoside derivatives through exchange of glycosyltransferases.

4.2.9.3 Unnatural Acyl Donors for Acyltransferases

Acyltransferases are a large class of tailoring enzymes that transfer thermodynamically activated acyl groups, such as thioesters of CoA or ACP, onto nucleophilic oxygen or nitrogen centers in natural products. Acyltransferases can use a wide variety of acyl groups, ranging from short to long-chain fatty acyl thioesters, to more functionalized acyl groups specific to a natural product. Like glycosylation, acyl transfer is a critical step in the maturation of natural product structures and are often indispensable for the biological activity. One well-documented example is in the last step of lovastatin (**2**) biosynthesis, in which the acyltransferase LovD catalyzes the transfer of α-methylbutyryl group from the PKS LovF to monacolin J (**438**) (Figure 115). This reaction installs the "side chain" in **2** that is essential for the inhibition of HMGR. LovD was engineered to produce the semisynthetic cholesterol-lowering drug simvastatin (**439**) (Figure 115),[892] which contains the α,α-dimethylbutyrate side chain and is more potent than **2**. Directed evolution of LovD was performed to increases its catalytic efficiency towards a small molecule thioester donor, dimethylbutyryl-*S*-methyl mercaptopropionate

(DMB-SMMP, **440**), as well as performance under slightly basic conditions in which **439** is precipitated to drive the reaction to completion. Heterologous production of **439** from *S. cerevisiae* was accomplished by coexpressing the evolved mutant LovD9 with the *lov* biosynthetic enzymes (except LovF) and feeding with **440**.

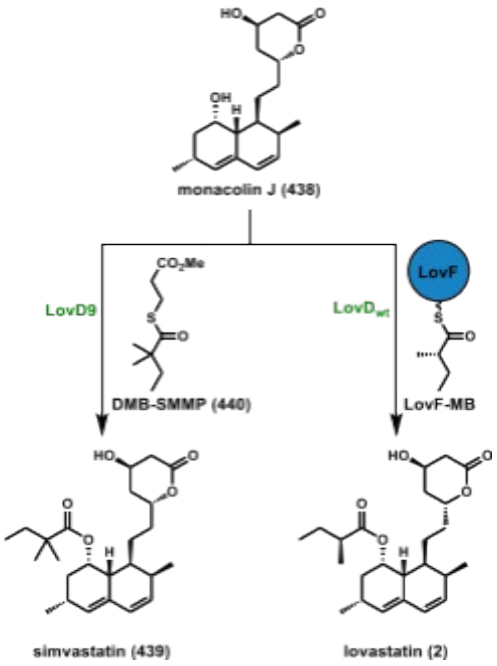


Figure 115. Biosynthesis of the semisynthetic **439** using engineered acyltransferase and small molecule acyl thioester **440**.

To alter the 1,8-dihydropyrrolo[2,3-b]indole-2-carboxylate moiety present in pyrroindomycin A (**441**), Liu and coworkers inactivated the *pyrK1* gene that is involved in the maturation of L-tryptophan to 1,8-dihydropyrrolo[2,3-b]indole-2-carboxylic acid in the producing strain *Streptomyces rugosporus* (Figure 116).[893] Subsequent supplementation of aromatic carboxylic acids resulted in six novel analogs of **441**. Similarly, Liu and coworkers reported the replacement of the 3-methyl-2-indolic acid unit (synthesized via NosL from L-tryptophan) tethered to Glu6 and Cys8 of nosiheptide (**442a**) with benzothiophene. In this RiPPs pathway, NosI catalyzes the transfer of 3-methyl-2-indolic acid onto the carrier protein NosJ, which is used as a substrate by the acyltransferase NosK to acylate Cys8 in the thiopeptide. These enzymes were tolerant of the unnatural 3-methyl-2-benzothiophenic acid, which was fed to a *ΔnosL Streptomyces actuosus* mutant and was incorporated to generate **442b**, the thiophene analog of **442a.**

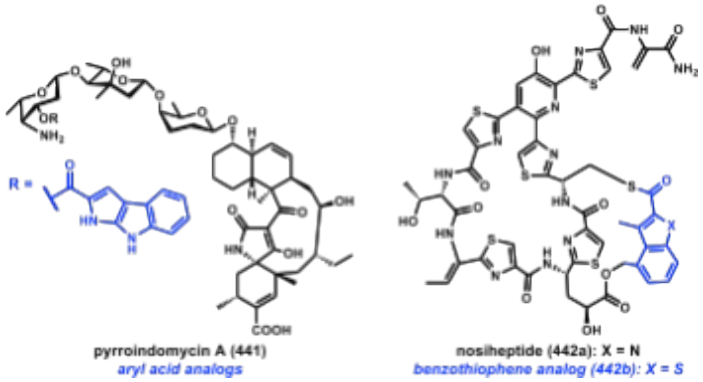


Figure 116. Targeting acyl transfer steps in biosynthesis for product diversification.

Exploring acyltransferase specificity for the generation of "clickable" analogs such as alkyne- and azide-containing fatty acids can lead to further synthetic modifications. Kirschning and Ott generated a magnetic-nanoparticle/ansamitocin conjugate (**443**) to probe a paramagnetic drug release system and direct ansamitocin to cancerous tissue.[894] A "clickable" ansamitocin alkyne analog (**444**) was generated by feeding hex-5-ynoic acid to the ansamitocin-producing *Actinosynnema pretiosum*. The terminal alkyne containing compouds was subsequently conjugated to a thermally sensitive oxanorbornadiene linkage and supermagnetic nanostructure particle (SPION) (Figure 117A). As demonstrated in xenograft mouse models, exposure of the nanoparticle to an external electric field at cancerous tissues resulted in a thermally activated retro-Diels-Alder reaction at the linker and release of the anticancer drug (Figure 117B).

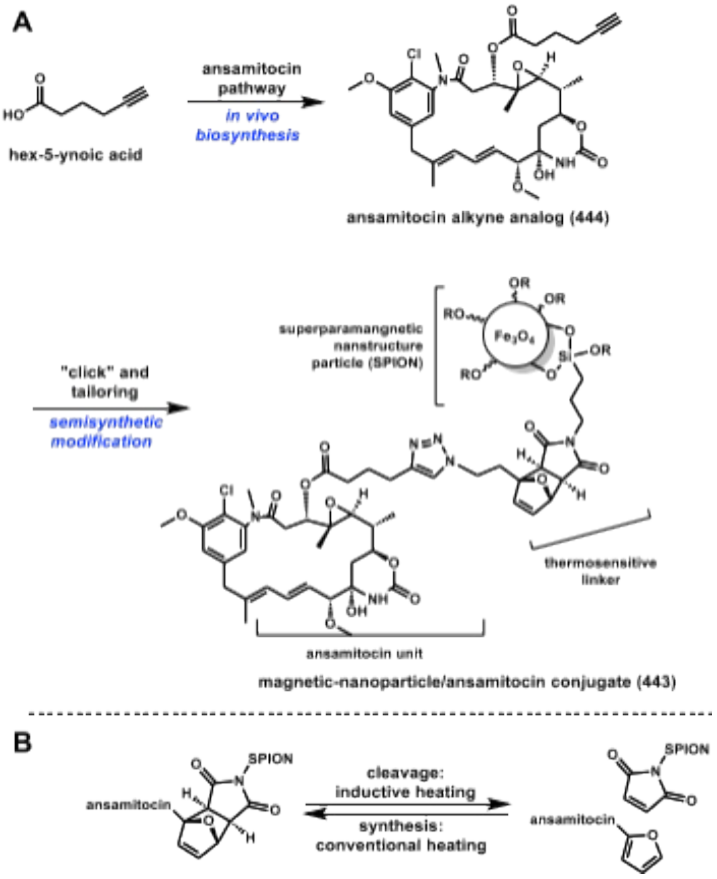


Figure 117. PDB was used to generate **444** that is the precursor to the magnetic-nanoparticle/ansamitocin conjugate **443**.

4.3. Combinatorial Biosynthesis

In the previous sections, we focused on natural product structural diversification by PDB, mutasynthesis or biotransformation. These approaches rely on the feeding of precursors and intermediates to biosynthetic pathways. During these discussions, we have shown examples in which biosynthetic genes from different pathways are introduced to aid in the *de novo* biosynthesis and incorporation of precursors, as well as modification of the intermediates. Broadly speaking, this approach can be classified as combinatorial biosynthesis, in which enzymes from different pathways are combined to biosynthesize a new compound, typically not found in Nature. While combinatorial biosynthesis can be as simple as switching a glycosyltransferase to diversify the sugar pendant on a natural product, this concept can also involve *de novo* pathway design to assembly otherwise completely unrelated enzymes into a single pathway. Some examples discussed in heterologous biosynthesis (section 3), in which enzymes from different kingdoms are sourced to complete a

functional pathway towards a final product, can thus be considered as ultimate examples of combinatorial biosynthesis. To close out this review, a few examples of combinatorial biosynthesis and pathway engineering will be summarized. Comprehensive coverage of this topic can be found in reviews by Goossens[895] and Leadlay.[896]

4.3.1 Modifying Biosynthetic Pathways to Generate New Compounds

The concept of combinatorial biosynthesis was first demonstrated in 1985, wherein Hopwood and coworkers explored the possibility of "hybrid" antibiotics via interchanging biosynthetic genes from related pathways found in different species.[37,897] The authors reported the novel antibiotics mederrhodin A (**445**) and dihydrogranatirhodin (**446**) (Figure 118). **445** features the medermycin (**447**) core scaffold with an additional phenolic alcohol found at the same position in actinorhodin (**103**), while **446** has structural features from both dihydrogranaticin (**448**) and **103**, differing at a single stereocenter. Although these changes are relatively minor, such early demonstrations laid the foundation to modern day combinatorial biosynthesis.

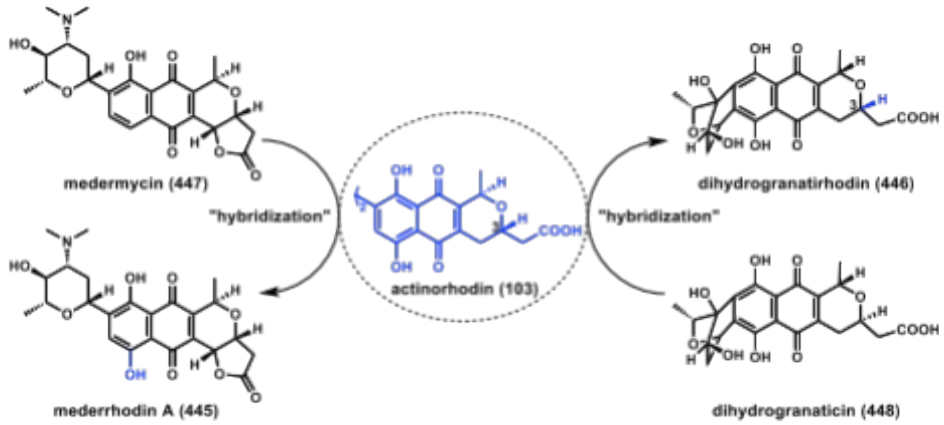


Figure 118. Early example of combinatorial biosynthesis by Hopwood and coworkers.

Several new polycyclic tetramate macrolactams (PoTeMs) compounds were generated via the addition of redox enzymes from related pathways. For example, the OX4 gene encoding a P450 from the heat stable-antifungal-factor (HSAF) biosynthetic pathway was expressed in the heterologous host *Streptomyces* sp. SR111 that is engineered to produce combamides A (**449**) and B (**450**). This hybrid pathway resulted in production of two new PoTeM derivatives, combamides E (**451**) and combamides F (**452**) with a 5,5,6 tricyclic system not observed among the nature combamides congeners (Figure

119).[898] Similarly, the substitution of CbmC with several cytochrome P450s from other PoTeMs resulted in derivatives with varying degrees of hydroxylation at different positions in the 5,5,6 tricyclic core when expressed in conjunction with OX4.[899] Along the same lines, Chen and coworkers rationally biosynthesized dimerized alboflavusins by introducing unrelated P450 biaryl coupling enzymes HmtS and ClpS to the biosynthetic pathway. The dimer products displayed enhanced antibacterial and antitumor activities compared to their monomeric forms.[900]

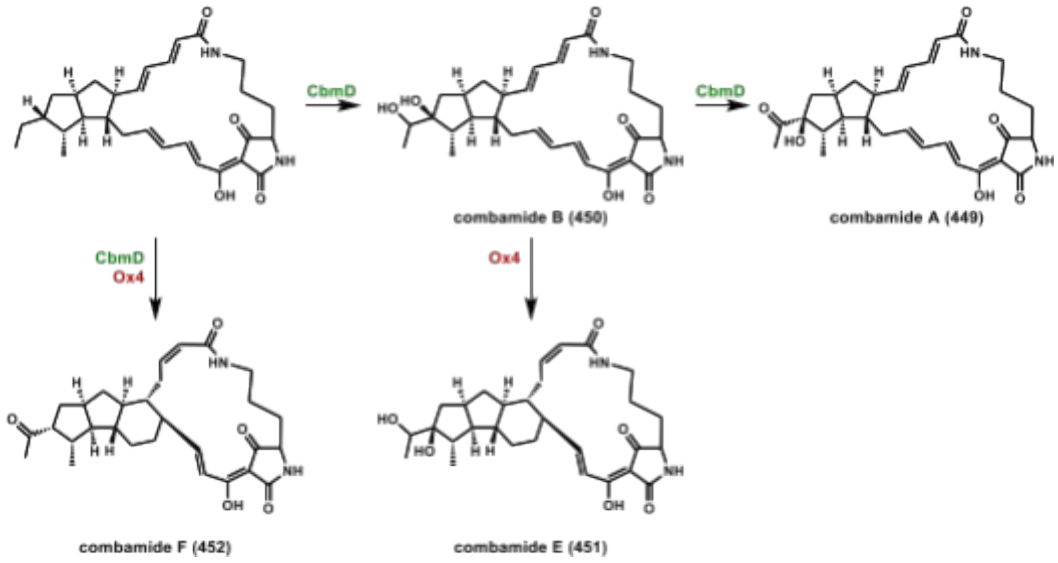


Figure 119. Coexpression of *ox4* with *cbmA-E* resulted in the generation of new hybrid PoTeMs **451** and **452.**

Cox and coworkers assembled *eup* and *as* genes involved in the biosynthesis of the tropolone sesquiterpenoids eupenifeldin (**288**) and xenovulene A (**38**), respectively, in *A. oryzae* to generate new compounds (Figure 120).[901] The two tropolone sesquiterpenoids derive from a common stipitaldehyde (**453**) intermediate but vary in the subsequent chemical transformations, including redox modification of **453**; the hetero-Diels Alder (hDA) reaction using either humulene (**39**) or epoxy-humulene (**454**); and the late-stage oxidative tailoring steps. By rationally mixing enzymes from the two pathways, seven new tropolone sesquiterpenoids were generated. To generate analogs of **288**, the pericyclase EupF from the BGC of **288**, which uses **454** to catalyze intermolecular hetero-Diels-Alder cyclization**,** was replaced with AsR5 from the BGC of **38**, which accepts **39**. Subsequent **288**-specific tailoring enzymes modified the unnatural intermediate to produce compounds such as **455** (Figure 120A). Similarly, to generate analogs of **38**, the **288**-specific tailoring enzymes such as EupR5 and EupR6 were used in place of **38**-specific tailoring enzymes AsL4 and AsL6, to modify the cycloaddition adduct xenovulene B (**456**)

(Figure 120B)**.** This combination of genes generated compounds such as **457** and **458.** The monobenzopyran moieties found in **455** and **457** were proposed to derive from oxidative ring contraction reactions not natively associated with any natural tropolone sesquiterpenoid and derivatives.

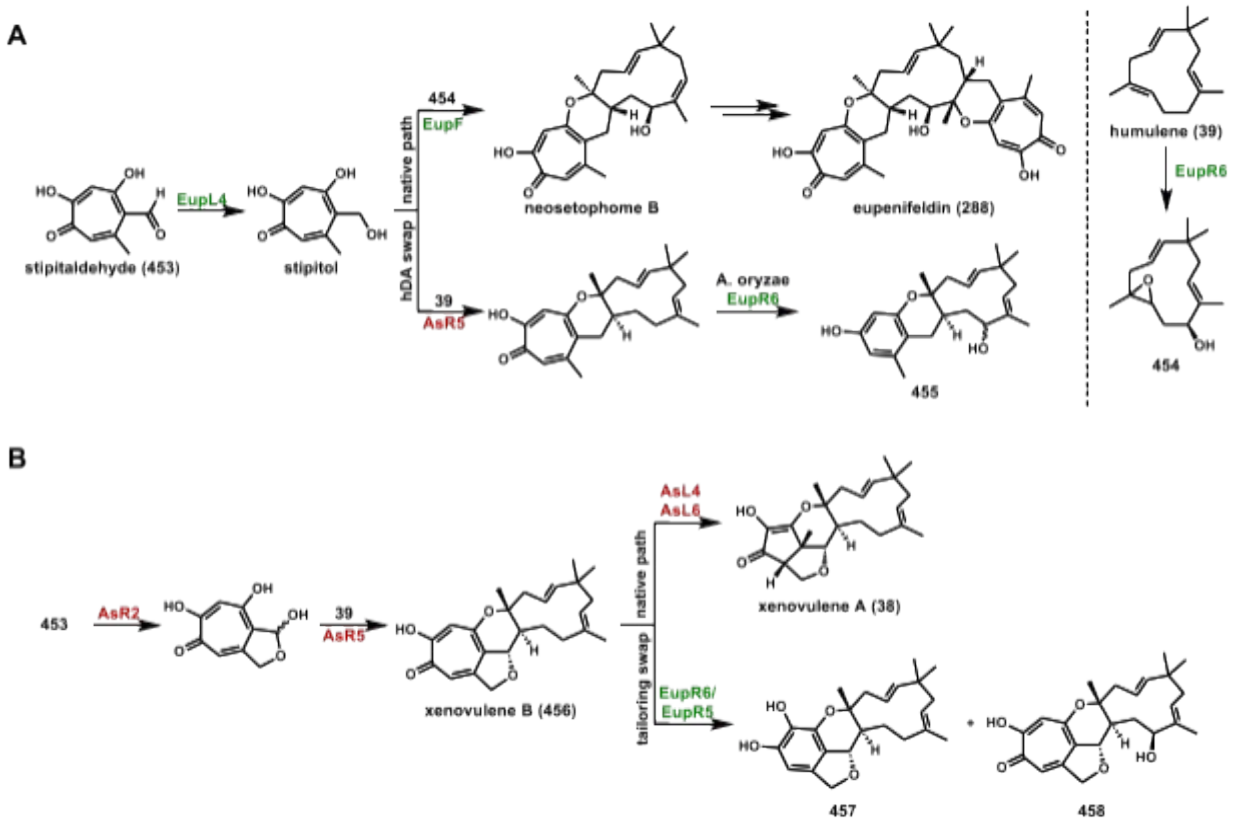


Figure 120. Access to new sesquiterpenoids using (A) the pericyclase from the BGC of **38** to modify the eupenifeldin core; or (B) tailoring enzymes from the BGC of **288** to modify **456**.

4.3.2 Mixing and Matching of Terpene Cyclase and Oxygenase Pairs

Combinatorial biosynthesis has been successful in generating new terpenoid compounds due to the promiscuity of oxidative tailoring enzymes from related pathways. Many productive examples of mixing TCs with oxidative tailoring enzymes from different terpene BGCs have been reported.[582,902–904] For example, in a high-throughput study, Voigt and coworkers built 1,088 pathways in *E. coli* from the pairing of 40 bacterial P450s and 17 TCs. Most of the P450s exhibited substrate promiscuity on multiple terpene scaffolds.[905] Similarly, the recently characterized SptF from the biosynthesis of fungal meroterpenoid emervaridones is a Fe(II)/α-ketoglutarate-dependent oxygenase with a broad substrate tolerance

towards diverse meroterpenoids.[906] This multifunctional enzyme catalyzes a range of oxidative manipulations, such as hydroxylations, desaturation, epoxidation, and skeletal rearrangements to furnish emervaridone B (**459**), emervaridone C (**460**), **461**, and **462** from andiconin D (**463**). Replacement of the homologous enzyme AndF responsible for desaturating andilesin C (**464**) to anditomin (**465**) with SptF resulted in production of the diastereomeric pair of **466** and **467** (Figure 121A). In an example of *in vitro* combinatorial biocatalysis, SptF was incubated with fungal meroterpenoids with varied core scaffolds (Figure 121B). Using preandiloid B (**468**) as a substrate, SptF catalyzed C_1 hydroxylation to give **469.** SptF also catalyzed oxidative ring contraction of tetronin J (**470**) and tetronin C (**471**) to form the cyclopropane-containing **472** and **473**, respectively. The same versatile enzyme was further found to catalyze hydroxylation at various positions in the sterol hormones androsterone, testosterone, and progesterone.

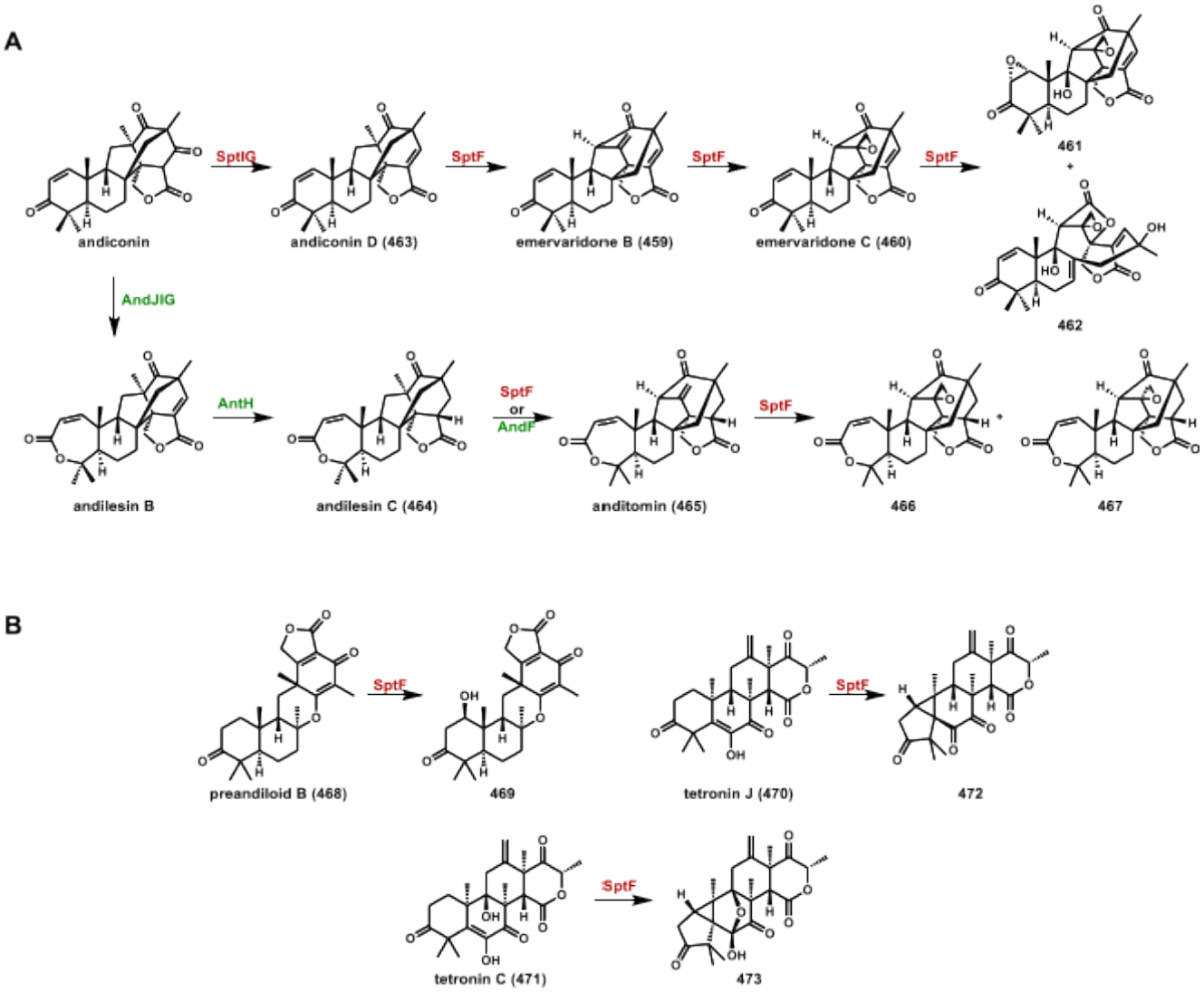


Figure 121. SptF is a promiscuous oxygenase that modifies meroterpenoids. (A) Modification of andilesin C (**464**) and (B) oxidation of other fungal meroterpenoids.

The Tissier group successfully reconstituted ten diterpene synthases and four cytochrome P450 oxygenases in various combinations in yeast. This effort led to the biosynthesis of over 200 distinct diterpenoids, 162 of which are new-to-nature diterpenoids.[907] Similarly, Tang and coworkers used yeast as a platform for the pairing of TCs and P450s from different bergamotene pathways to generate new sesquiterpene products.[908] Fma-TC and Fma-P450 are from the biosynthetic pathway of fumagillin in *Aspergillus fumigatus*, while Tv86-TC and Tv86-P450 are mined from *Trichoderma virens*. Fma-TC and Tv86-TC are atypical TCs that convert FPP into β-*trans*-bergamotene (**474**) and α-*trans*-bergamotene (**475**), respectively. Both P450s are multifunctional enzymes that act on their respective pathways to give dramatically modified compounds such as **476**-**479** as shown in Figure 122. Pairing the TC with the noncognate P450 partner led to structural diversification of sesquiterpenes **480**-**483**.[908] Similarly, pairing two fungal P450 enzymes BcBOT4 and PeniB, responsible for the maturation of phytotoxin botrydial and the insecticide penifulvin A, respectively, with seven TCs in yeast led to new polyquinane sesquiterpenoids containing modifications such as epoxidations, carboxylation, hydroxylation, etc.[902] Kampranis and coworkers developed a high throughput combinatorial biosynthesis workflow using yeast wherein combinations of terpene cyclases and cytochrome P450s were iteratively refined based on inhibitory activity against protein-tyrosine phosphatase 1B of yeast mutant extracts.[909] Amongst over 1000 assayed combinations, 20 triterpenoids exhibited potent activity against protein-tyrosine phosphatase 1B, four of which represent new compounds.

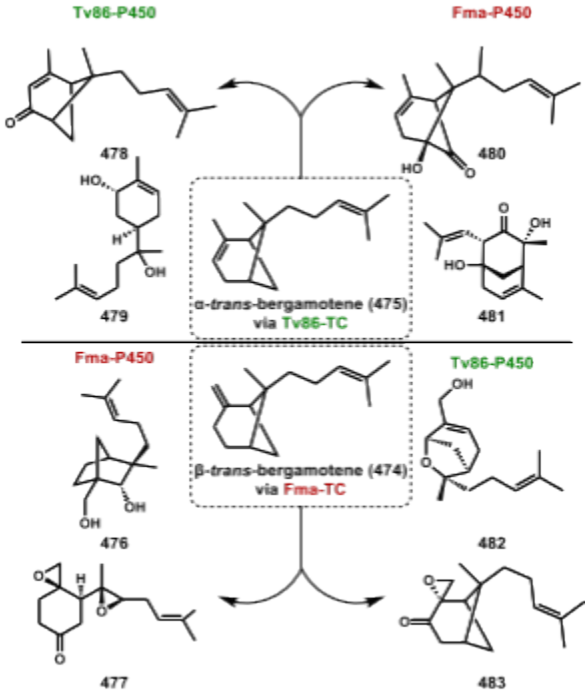


Figure 122. The mixing and matching of two sesquiterpene cyclases and their associated P450s enabled the generation of a wide variety sesquiterpenoids.

Asai and coworkers utilized heterologous expression to build unnatural biosynthetic pathways that produce decalin-containing diterpenoid pyrones.[910] Flavin-dependent monooxygenases and methyltransferases from five different BGCs were added or swapped between pathways, enabling the generation of ten previously unknown compounds with varied carbon skeletons and oxidative modifications. The Cox group demonstrated the utility of combinatorial biocatalysis to generate a web-like biosynthetic network of chimeric pathways containing oxidative enzymes acting upon the aristolochene (**484**) core in concert (Figure 123). [911] Through manipulation of six different genes in *A. oryzae*, twenty compounds could be derived. Further diversity was generated by the endogenous oxidative machinery of *A. oryzae,* producing on-pathway "shunt" compounds that could be further tailored by introduced P450 enzymes.

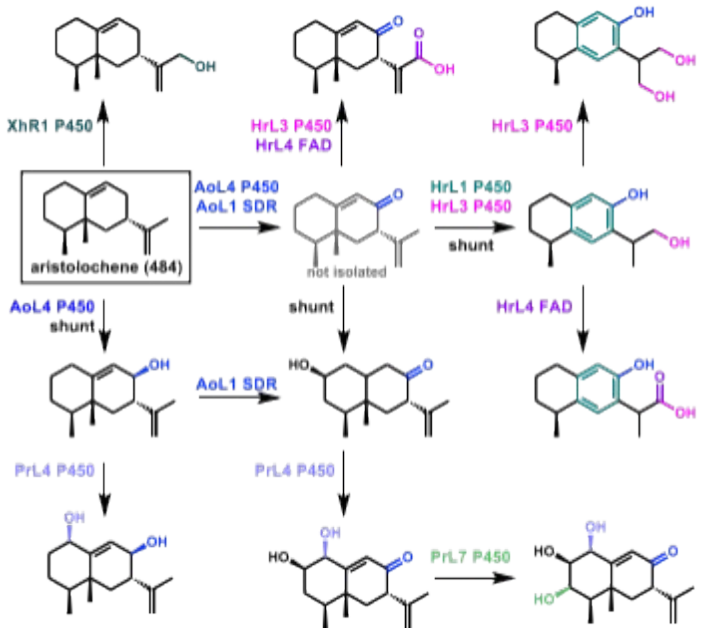


Figure 123. Combinatorial expression of oxidative enzymes that act on **484**.

4.3.3 *De Novo* Pathway Design

In an interesting application of pathway design via natural product "fusion," Keasling and Hartwig sought to integrate the newly discovered carbene-transfer reaction catalyzed by an engineered P450 enzyme into cellular metabolism and produce new-to-nature natural products.[912] In a proof-of-concept study, the authors integrated the biosynthetic pathway for α-diazoester azaserine (**160**) from *Salix fragilis* into *S. albus*, to produce the carbene precursor. In addition, the authors expressed the enzymes required to produce styrene (**485**) as the carbene acceptor. The engineered mutant of P450-T2, which natively accepts L-mimosine as a substrate, could catalyze the cyclopropanation of **485** using **160** to generate "fusion" product **486** (Figure 124). While the wild-type P450-T2 was active in catalyzing the cyclopropanation reaction with moderate diastereoselectivity, competing reaction catalyzed by free heme resulted in reduced diastereocontrol *in cellulo*. To enhance P450-T2 catalytic turnover, five rounds of directed evolution were conducted, yielding an engineered P450-T2 with dr >99% and 250-fold higher turnover rate than the wild-type. This example illustrates how pathway discovery (that of **160**), pathway engineering (that of **485**) and protein engineering (that of P450-T2) can be integrated to enable unprecedented chemistry in biology that can be harnessed for chemical synthesis.

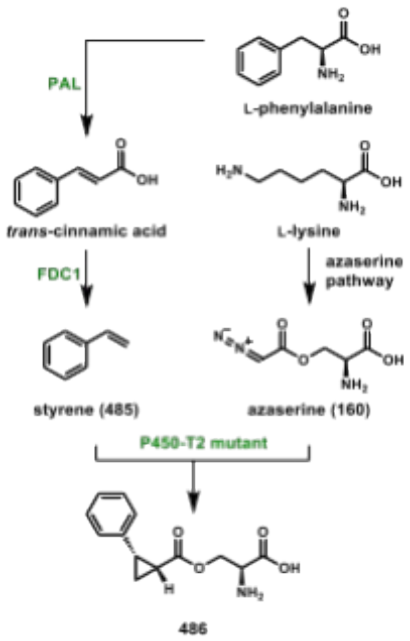


Figure 124. Engineering of a P450-T2 mutant enabled the enzymatic "hybridization" of **485** and **160** to **486** via cyclopropane formation.

5 Conclusions and Outlook

There is no denying that interest in natural product-based drug discovery in the pharmaceutical industry declined in recent decades, largely due to the decreases in hit rates from microbial and plant sources. In the post-genomics era, however, a resurgence of academic, entrepreneurial, and industrial efforts towards natural product discovery has taken place. Such explosion of creative efforts to discover, produce, and engineer natural products has transformed the field into one that is highly interdisciplinary, bringing together bioinformaticians, computer scientists, synthetic biologists, chemists, enzymologists, etc. Synthetic biology in particular has become a mainstay in the field as the emphasis shifts from organism-based discovery to genome-based discovery. Forty years after the initial discovery of a natural product biosynthetic gene cluster (BGC) by Sir David Hopwood, it is now routine, and arguably preferred, for a BGC to be completely characterized and the associated natural product to be identified without working with the native organism.

To emphasize the role of bioinformatics and machine learning on the field, Section 2 described advances in computational tools to find BGCs and predict associated natural products, as well as to elucidate and dereplicate natural product structures using analytical tools such as MS and NMR. These computational tools have become more sophisticated

and higher-resolution over time, and have seen widespread adoption and customization by the community. The use of programs such as antiSMASH for BGC identification has become an indispensable first step in most workflows. The challenge of computational BGC dereplication is a significant one and is being tackled by curated community databases. Machine learning is playing an increasingly important role in BGC identification and annotation, as exemplified by the use of AlphaFold3 to predict functions of proteins. At the backend, dereplication tools such as GNPS have become indispensable for many as the first step in analysis of complex metabolite extracts.

Synthetic biology tools play a prominent role in translation of bioinformatic predictions into natural products, analogs, as well as biosynthetic enzymes. Once a certain predicted BGC has been selected for characterization, multiple approaches to interrogate the encoded natural products can be used, in both the native host and in heterologous hosts. Modern synthetic biology tools such as CRISPR-based gene activation, promoter engineering and ribosomal engineering have all found applications in natural product discovery. However, the one approach in which synthetic biology has played the most important role is the development of heterologous expression hosts for biosynthetic pathway characterization and engineering. From the examples presented in Section 3, it is clear that while the core concept of heterologous expression has been practiced for decades, the influx of new tools to construct pathways, engineer host and source biosynthetic genes has elevated the approach to a new level of complexity with respect to the length of pathways and number of genetic modifications possible. Essentially all classes of natural products have been heterologously produced in different model bacterial, fungal and plant hosts, highlighted by the impressive accomplishments of vinblastine and QS-21 biosynthesis in yeast. BGCs prioritized for bioactivities and novel structures can be readily mined using heterologous hosts, resulting in discovery of natural products with desired biological targets and new chemical scaffolds, respectively.

Natural products have served as leads for semisynthetic generation or chemical synthesis of more effective analogs. For example, the polyketide avermectin was isolated by Omura, while the more potent 22,23-dihydro derivative of avermectin B_1 known as ivermectin was obtained via chemical modification by Campbell. Similarly, the discovery of natural statins such as lovastatin and mevastatin has inspired semisynthesis of simvastatin and total synthesis of atorvastatin, all of which were blockbuster cholesterol lowering drugs. Section 4 provides examples of how biosynthetic pathways can be rationally engineered for the generation of complex natural product analogs with high structural precision, which are difficult to access by semisynthesis or total synthesis. Since each step of a biosynthetic pathway can be manipulated at the

gene level, many possible methods to engineer the structure of the final product can be identified and executed. Further involvement of tools such as directed evolution and metabolic engineering will set the stage for the integration of abiotic chemistry into microbial hosts, leading to much broader expansion of the chemical spaces associated with known natural products.

In sum, the convergences of several paradigm-shifting technologies, including cost-effective DNA synthesis, Next-generation DNA sequencing, CRISPR-based gene targeting, directed evolution, machine-learning based protein function prediction, etc in the last decade, have reenergized natural product discovery and engineering. We anticipate in the coming decade, integration of other synthetic biology and analytical technologies, such as photobiocatalysis, cell-free biosynthesis, AI-driven cellular engineering, and electron microscopy for structural elucidation, will further accelerate the discovery of new compounds from nature, as well as engineering of new compounds to nature.

Author Information

Corresponding Author

Yi Tang - Department of Chemical and Biomolecular Engineering and Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States; orcid.org/0000-0003-1597-0141; Email: yitang@ucla.edu

Authors

Kaushik Seshadri - Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, California 90095, United States; orcid.org/0000-0003-3436-4585

Abner N. D. Abad - Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, California 90095, United States; orcid.org/0009-0003-0516-256X

Kyle K. Nagasawa - Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States; orcid.org/0009-0005-1160-6053

Karl M. Yost - Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States; orcid.org/0009-0001-2532-703X

Colin W. Johnson - Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States; orcid.org/0000-0003-1092-1172

Moriel J. Dror - Department of Bioengineering, University of California, Los Angeles, California 90095, United States; orcid.org/0009-0008-4348-8882

Notes

The authors declare no competing financial interest.

Biographies

Kaushik Seshadri, born in 2001, received his undergraduate degree in Chemical Engineering from the Department of Chemical and Biomolecular Engineering at the University of California, Berkeley, in 2023. He is currently a PhD student in Prof. Yi Tang's lab at the University of California, Los Angeles. His current research interests focus on natural product biosynthesis and enzymology.

Abner N. D. Abad, born in 1998, received his undergraduate degree in Chemical Engineering from the Department of Chemical and Biomolecular Engineering at the University of California, Berkeley, in 2021. He is currently a PhD student in Prof. Yi Tang's lab at the University of California, Los Angeles. His current research interests focus on natural product discovery and biosynthesis.

Kyle K. Nagasawa, born in 2003, is an undergraduate student working toward an undergraduate degree in Biochemistry from the University of California, Los Angeles. He currently conducts research in Prof. Yi Tang's lab at the University of California, Los Angeles, and studies nontraditional core-forming enzymes towards the discovery of novel bioactive molecules.

Karl M. Yost, born in 2001, completed his undergraduate degree in Biochemistry at the University of California, Los Angeles in 2024. During this time, he conducted research in Professor Yi Tang's lab, focusing on structural characterization and biosynthetic pathway elucidation. He is now pursuing a Ph.D. in Chemistry and Chemical Biology at Harvard University.

Colin W. Johnson, born in 1993, received his B.A. in Biochemistry from Claremont McKenna College in 2016. He is currently a Ph.D. candidate in Professor Yi Tang's laboratory at the University of California, Los Angeles. His research interests focus on leveraging bioinformatics to uncover metabolites of therapeutic value assembled by noncanonical scaffolding enzymes in fungi.

Moriel J. Dror, born in 1998, received her B.S. in Chemical and Biomolecular Engineering from University of California, Berkeley in 2020. She is currently a Ph.D. candidate in Professor Yi Tang's laboratory at the University of California, Los Angeles. Her research interests focus on strain/metabolic engineering, specifically reconstituting plant pathways and optimizing biosynthetic titers in yeast.

Yi Tang, born in 1976, received his undergraduate degree in Chemical Engineering and Material Science from Penn State University. He received his Ph.D. in Chemical Engineering from California Institute of Technology under the guidance of Prof. David A. Tirrell. After NIH postdoctoral training in Chemical Biology from Prof. Chaitan Khosla at Stanford University, he started his independent career at University of California Los Angeles in 2004. He is currently the Parson's Family Foundation Professor in the Department of Chemical and Biomolecular Engineering at UCLA, and holds joint appointments in the Department of Chemistry and Biochemistry; and Department of Bioengineering.

Acknowledgements

This work was supported by the NIGMS 5R35GM118056 to YT. MJD is a fellow of the BioPacific Program supported by the NSF cooperative agreement (DMR-1933487). KKN was supported by the UCLA Undergraduate Research Center-Sciences Program. Figure 68 was created with BioRender.com.

Dedication

This review is dedicated to the life of Prof. Christopher T. Walsh.

Abbreviations

ACCase	acyl-CoA carboxylases
ACP	acyl carrier protein
ADC	antibody-drug conjugate
antiSMASH	antibiotics and Secondary Metabolite Analysis Shell
AT	acyltransferase
BAC	bacterial artificial chromosome
BAGEL	BActeriocin GEnome mining tool
BGC	biosynthetic gene cluster
BiG-FAM	Biosynthetic Gene Cluster Filies Database
Big-SCAPE	Biosynthetic Gene Similarity Clustering and Prospective Engine
BiG-SLICE	Biosynthetic Genes Super-Linear Clustering Engine
BiGCARP	Biosynthetic Gene Convolutional Autoencoding Representations of Proteins
BiLSTM	Bidirectional Long Short-Term Memory
BLAST	Basic Local Alignment Search Tool
CAPTURE	Cas12a-Assisted Precise Targeted cloning Using in vivo Cre-lox REcombination
CATCH	Cas9-Assisted Targeting of CHromosome segments
CD	conserved domain
CDPS	cyclodipeptide synthase
CLEAN	Contrastive Learning-enabled Enzyme Annotation

CMGnet	Conditional Molecular Generation net
CNN	convolutional neural networks
CONKAT-Seq	Co-Occurrence NetworK Analysis of Targeted SEQuences
CORASON	Core Analysis of Syntenic Orthologues to prioritize Natural product gene clusters
COSMIC	Confidence of Small Molecule Identifications
CRAGE	Chassis-independent Recombinase-Assisted Genome Engineering
CRF	conditional random fields
CRISPRa	CRISPR activation
CRISPRi	CRISPR interference
CCR	crotonyl-CoA reductase/carboxylase
CSA	Composite Spectra Analysis
CVD	cluster variation database
DAFdiscovery	Data-Fusion-based Discovery
DBTL	design, build, test, and learn
ddTAG	Dynamic Degradation of triacylglycerol (TAG)
DMAPP	dimethylallyl pyrophosphate
DMAT	dimethylallyl tryptophan synthase
DUF	domains of unknown function
ER	endoplasmic reticulum
FAC	fungal artificial chromosomes
FACS	fluorescence-activated cell sorting
FAS	fatty acid synthase
FBMN	Feature-Based Molecular Networking
FPP	farnesyl pyrophosphate
FPPS	farnesyl pyrophosphate synthase
FRET	fluorescence resonance energy transfer
FRIGG	Fungal Resistance Gene-directed Genome Mining
FSD	flanking subdomain
FunARTS	Fungal bioActive compound Resistant Target Seeker
GCF	gene cluster families
GECCO	GEne Cluster prediction with COnditional random fields
GFP	green fluorescent protein
GIPMA	Global Intensity-guided Peak Matching and Alignment
GNN	graph neural network
GNPS	Global Natural Products Social
GPP	geranylgeranyl pyrophosphate
GRAS	Generally Regarded as Safe
HAT	histone acetyltransferase
HDAC	histone deacetylase
HiTES	high-throughput elicitor screening
HMM	hidden Markov model
HP	hypothetical protein
HSQC	heteronuclear single quantum coherence
IDSL-MINT	Mass-INTerpreter
IMS	imaging mass spectrometry
IPK	isopentenyl phosphate kinase
IPP	isoprenyl pyrophosphate

IUP	Isopentenol Utilization Pathway
KGMN	Knowledge-Guided Multilayer Network
KS	ketosynthase
LAESI-MS	laser-ablation-coupled electrospray ionization MS
MAA	mycosporine-like amino acid
MassQL	Mass Spec Query Language
MASST	Mass Spectrometry Tool
MD	molecular dynamics
MEP	methylerythritol phosphate
MEV	mevalonate
MFS	major facilitator superfamily
MIA	monoterpene indole alkaloid
MIBiG	Minimum Information about a Biosynthetic Gene cluster
ML	machine learning
MN	molecular network
MNIO	multinuclear iron oxygenase
MS	mass spectrometry
NHEJ	non-homologous end joining
NLP	natural language processing
NMR	nuclear magnetic resonance spectroscopy
NRP	nonribosomal peptide
NRPS	nonribosomal peptide synthetase
ORF	open reading frame
OSMAC	one strain many compounds
PAM	phenylalanine aminomutase
PDB	precursor directed biosynthesis
PDB	precursor-directed biosynthesis
PEARL	PEptide Aminoacyl-tRNA Ligase
PKS	polyketide synthase
PLP	pyridoxial-5’-phosphate
PoTeM	polycyclic tetramate macrolactam
PRISM	Prediction Informatics for Secondary Metabolomes
PROTAC	proteolysis-targeting chimera
PTM	post-translation modification
RF	random field
Rg3m	Resistance Gene-Guided Genome Miner
RiPP	ribosomally synthesized and post-translationally modified peptide
RODEO	Rapid ORF Description and Evaluation Online
RRE	RiPP precursor recognition element
SAM	S-adenosylmethionine
SAR	structure-activity relationships
SCORE-metabolite-ID	Semi-automatic correlation analysis for reliable metabolite annotation
SGMNS	Structure-Guided Molecular Networking Strategy
SHY	Statistical HetereoSpectroscopy
SMART	Small Molecule Accurate Recognition Technology
SMURF	Secondary Metabolite Unknown Regions Finder
SNAC	synthetic diketide S-N-acylcysteamine

SNAP-MS	Structural similarity Network Annotation Platform for Mass Spectrometry
SRE	self-resistance enzyme
SSN	sequence similarity network
STOCSY	statistical total correlation spectroscopy
SVM	support vector machine
TAR	transformation-assisted recombination
TC	terpene cyclase
TE	thioesterase
THIQ	tetrahydroisoquinolines
TRY	high titer, rate, and yield
TS	terpene synthase
UHPLC-HRMS	ultra-high-performance liquid chromatography-high resolution mass spectrometry
XU	eXchange Unit
XUC	eXchange Unit between Condensation domains
XUT	eXchange Unit between Thiolation domains

References

(1) Walsh, C. T.; Tang, Y. Natural Product Biosynthesis; The Royal Society of Chemistry, **2022**. https://doi.org/10.1039/9781837671069.

(2) Demain, A. L.; Fang, A. The Natural Functions of Secondary Metabolites; **2000**; pp 1–39. https://doi.org/10.1007/3-540-44964-7_1.

(3) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J Nat Prod* **2020**, *83* (3), 770–803. https://doi.org/10.1021/acs.jnatprod.9b01285.

(4) Stratton, C. F.; Newman, D. J.; Tan, D. S. Cheminformatic Comparison of Approved Drugs from Natural Product versus Synthetic Origins. *Bioorg Med Chem Lett* **2015**, *25* (21), 4802–4807. https://doi.org/10.1016/j.bmcl.2015.07.014.

(5) Atanasov, A. G.; Zotchev, S. B.; Dirsch, V. M.; Supuran, C. T. Natural Products in Drug Discovery: Advances and Opportunities. *Nat Rev Drug* Discov **2021**, *20* (3), 200–216. https://doi.org/10.1038/s41573-020-00114-z.

(6) Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J. The Re-Emergence of Natural Products for Drug Discovery in the Genomics Era. *Nature Reviews Drug Discovery* **2015**. *14* (2), 111-129. https://doi.org/10.1038/nrd4510.

(7) Fu, Z.; Li, S.; Han, S.; Shi, C.; Zhang, Y. Antibody Drug Conjugate: The "Biological Missile" for Targeted Cancer Therapy. *Signal Transduct Target Ther* **2022**, 7 (1), 93. https://doi.org/10.1038/s41392-022-00947-7.

(8) Tsuchikama, K.; Anami, Y.; Ha, S. Y. Y.; Yamazaki, C. M. Exploring the next Generation of Antibody–Drug Conjugates. *Nat Rev Clin Oncol* **2024**, *21* (3), 203–223. https://doi.org/10.1038/s41571-023-00850-2.

(9) Schreiber, S. L. The Rise of Molecular Glues. *Cell* **2021**, *184* (1), 3–9. https://doi.org/10.1016/j.cell.2020.12.020.

(10) Isobe, Y.; Okumura, M.; McGregor, L. M.; Brittain, S. M.; Jones, M. D.; Liang, X.; White, R.; Forrester, W.; McKenna, J. M.; Tallarico, J. A.; Schirle, M.; Maimone, T. J.; Nomura, D. K. Manumycin Polyketides Act as Molecular Glues between UBR7 and P53. *Nat Chem Biol* **2020**, *16* (11), 1189–1198. https://doi.org/10.1038/s41589-020-0557-2.

(11) Li, Y.; Jia, Y.; Wang, X.; Shang, H.; Tian, Y. Protein-Targeted Degradation Agents Based on Natural Products. *Pharmaceuticals* **2022**, *16* (1), 46. https://doi.org/10.3390/ph16010046.

(12) Sakamoto, K. M.; Kim, K. B.; Kumagai, A.; Mercurio, F.; Crews, C. M.; Deshaies, R. J. Protacs: Chimeric Molecules That Target Proteins to the Skp1–Cullin–F Box Complex for Ubiquitination and Degradation. *Proceedings of the National Academy of Sciences* **2001**, *98* (15), 8554–8559. https://doi.org/10.1073/pnas.141230798.

(13) Spradlin, J. N.; Hu, X.; Ward, C. C.; Brittain, S. M.; Jones, M. D.; Ou, L.; To, M.; Proudfoot, A.; Ornelas, E.; Woldegiorgis, M.; Olzmann, J. A.; Bussiere, D. E.; Thomas, J. R.; Tallarico, J. A.; McKenna, J. M.; Schirle, M.; Maimone, T. J.; Nomura, D. K. Harnessing the Anti-Cancer Natural Product Nimbolide for Targeted Protein Degradation. *Nat Chem Biol* **2019**, *15* (7), 747–755. https://doi.org/10.1038/s41589-019-0304-8.

(14) Bachmann, B. O.; Van Lanen, S. G.; Baltz, R. H. Microbial Genome Mining for Accelerated Natural Products Discovery: Is a Renaissance in the Making? *J Ind Microbiol Biotechnol* **2014**, *41* (2), 175–184. https://doi.org/10.1007/s10295-013-1389-9.

(15) Challis, G. L. Mining Microbial Genomes for New Natural Products and Biosynthetic Pathways. *Microbiology* (N Y) **2008**, *154* (6), 1555–1569. https://doi.org/10.1099/mic.0.2008/018523-0.

(16) Medema, M. H.; Fischbach, M. A. Computational Approaches to Natural Product Discovery. *Nat Chem Biol* **2015**, *11* (9), 639–648. https://doi.org/10.1038/nchembio.1884.

(17) Staunton, J.; Weissman, K. J. Polyketide Biosynthesis: A Millennium Review. *Nat Prod Rep* **2001**, *18* (4), 380–416. https://doi.org/10.1039/a909079g.

(18) Shen, B. Polyketide Biosynthesis beyond the Type I, II and III Polyketide Synthase Paradigms. *Curr Opin Chem Biol* **2003**, *7* (2), 285–295. https://doi.org/10.1016/S1367-5931(03)00020-6.

(19) Chen, A.; Re, R. N.; Burkart, M. D. Type II Fatty Acid and Polyketide Synthases: Deciphering Protein–Protein and Protein–Substrate Interactions. *Nat Prod Rep* **2018**, *35* (10), 1029–1045. https://doi.org/10.1039/C8NP00040A.

(20) Winn, M.; Fyans, J. K.; Zhuo, Y.; Micklefield, J. Recent Advances in Engineering Nonribosomal Peptide Assembly Lines. *Natural Product Reports* **2016**. 33 (2), 317-347. https://doi.org/10.1039/c5np00099h.

(21) Ogasawara, Y.; Dairi, T. Biosynthesis of Oligopeptides Using ATP-Grasp Enzymes. *Chemistry – A European Journal* **2017**, *23* (45), 10714–10724. https://doi.org/10.1002/chem.201700674.

(22) Goswami, A.; Van Lanen, S. G. Enzymatic Strategies and Biocatalysts for Amide Bond Formation: Tricks of the Trade Outside of the Ribosome. *Mol Biosyst* **2015**, *11* (2), 338–353. https://doi.org/10.1039/C4MB00627E.

(23) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K.-D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H.-G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G.-L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A. Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature. *Nat. Prod. Rep.* **2013**, *30* (1), 108–160. https://doi.org/10.1039/C2NP20085F.

(24) Russell, A. H.; Truman, A. W. Genome Mining Strategies for Ribosomally Synthesised and Post-Translationally Modified Peptides. *Comput Struct Biotechnol J* **2020**, *18*, 1838–1851. https://doi.org/10.1016/j.csbj.2020.06.032.

(25) Gershenzon, J.; Dudareva, N. The Function of Terpene Natural Products in the Natural World. *Nat Chem Biol* **2007**, *3* (7), 408–414. https://doi.org/10.1038/nchembio.2007.5.

(26) Paddon, C. J.; Westfall, P. J.; Pitera, D. J.; Benjamin, K.; Fisher, K.; McPhee, D.; Leavell, M. D.; Tai, A.; Main, A.; Eng, D.; Polichuk, D. R.; Teoh, K. H.; Reed, D. W.; Treynor, T.; Lenihan, J.; Jiang, H.; Fleck, M.; Bajad, S.; Dang, G.; Dengrove, D.; Diola, D.; Dorin, G.; Ellens, K. W.; Fickes, S.; Galazzo, J.; Gaucher, S. P.; Geistlinger, T.; Henry, R.; Hepp, M.; Horning, T.; Iqbal, T.; Kizer, L.; Lieu, B.; Melis, D.; Moss, N.; Regentin, R.; Secrest, S.; Tsuruta, H.; Vazquez, R.; Westblade, L. F.; Xu, L.; Yu, M.; Zhang, Y.; Zhao, L.; Lievense, J.; Covello, P. S.; Keasling, J. D.; Reiling, K. K.; Renninger, N. S.; Newman, J. D. High-Level Semi-Synthetic Production of the Potent Antimalarial Artemisinin. *Nature* **2013**, *496* (7446), 528–532. https://doi.org/10.1038/nature12051.

(27) Jamieson, C. S.; Misa, J.; Tang, Y.; Billingsley, J. M. Biosynthesis and Synthetic Biology of Psychoactive Natural Products. *Chem Soc Rev* **2021**, *50* (12), 6950–7008. https://doi.org/10.1039/D1CS00065A.

(28) Lichman, B. R. The Scaffold-Forming Steps of Plant Alkaloid Biosynthesis. *Nat Prod Rep* **2021**, *38* (1), 103–129. https://doi.org/10.1039/D0NP00031K.

(29) Shiraishi, T.; Kuzuyama, T. *Recent Advances in the Biosynthesis of Nucleoside Antibiotics. J Antibiot (Tokyo)* **2019**, *72* (12), 913–923. https://doi.org/10.1038/s41429-019-0236-2.

(30) Kudo, F.; Eguchi, T. Biosynthetic Genes for Aminoglycoside *Antibiotics. J Antibiot (Tokyo)* **2009**, *62* (9), 471–481. https://doi.org/10.1038/ja.2009.76.

(31) Tang, L.; Shah, S.; Chung, L.; Carney, J.; Katz, L.; Khosla, C.; Julien, B. Cloning and Heterologous Expression of the Epothilone Gene Cluster. *Science (1979)* **2000**, *287* (5453), 640–642. https://doi.org/10.1126/science.287.5453.640.

(32) Yu, D.; Xu, F.; Zeng, J.; Zhan, J. Type III Polyketide Synthases in Natural Product Biosynthesis. *IUBMB Life* **2012**, *64* (4), 285–295. https://doi.org/10.1002/iub.1005.

(33) Adam, P.; Arigoni, D.; Bacher, A.; Eisenreich, W. Biosynthesis of Hyperforin in Hypericum p Erforatum. *J Med Chem* **2002**, *45* (21), 4786–4793. https://doi.org/10.1021/jm0209782.

(34) Barra, L.; Awakawa, T.; Shirai, K.; Hu, Z.; Bashiri, G.; Abe, I. β-NAD as a Building Block in Natural Product Biosynthesis. *Nature* **2021**, *600* (7890), 754–758. https://doi.org/10.1038/s41586-021-04214-7.

(35) Patteson, J. B.; Putz, A. T.; Tao, L.; Simke, W. C.; Bryant, L. H.; Britt, R. D.; Li, B. Biosynthesis of Fluopsin C, a Copper-Containing Antibiotic from Pseudomonas Aeruginosa. *Science (1979)* **2021**, *374* (6570), 1005–1009. https://doi.org/10.1126/science.abj6749.

(36) Yee, D. A.; Niwa, K.; Perlatti, B.; Chen, M.; Li, Y.; Tang, Y. Genome Mining for Unknown–Unknown Natural Products. *Nat Chem Biol* **2023**, *19* (5), 633–640. https://doi.org/10.1038/s41589-022-01246-6.

(37) Malpartida, F.; Hopwood, D. A. Molecular Cloning of the Whole Biosynthetic Pathway of a Streptomyces Antibiotic and Its Expression in a Heterologous Host. *Nature* **1984**, *309* (5967), 462-464. https://doi.org/10.1038/309462a0.

(38) Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. An Unusually Large Multifunctional Polypeptide in the Erythromycin-Producing Polyketide Synthase of Saccharopolyspora Erythraea. *Nature* **1990**, *348* (6297), 176–178. https://doi.org/10.1038/348176a0.

(39) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. Modular Organization of Genes Required for Complex Polyketide Biosynthesis. *Science (1979)* **1991**, *252* (5006), 675–679. https://doi.org/10.1126/science.2024119.

(40) Nützmann, H.; Huang, A.; Osbourn, A. Plant Metabolic Clusters – from Genetics to Genomics. *New Phytologist* **2016**, *211* (3), 771–789. https://doi.org/10.1111/nph.13981.

(41) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

(42) Bozhüyük, K. A. J.; Präve, L.; Kegler, C.; Schenk, L.; Kaiser, S.; Schelhas, C.; Shi, Y. N.; Kuttenlochner, W.; Schreiber, M.; Kandler, J.; Alanjary, M.; Mohiuddin, T. M.; Groll, M.; Hochberg, G. K. A.; Bode, H. B. Evolution-Inspired Engineering of Nonribosomal Peptide Synthetases. *Science* **2024**, *383* (6689), eadg4320. https://doi.org/10.1126/science.adg4320.

(43) Kelley, D. R.; Liu, B.; Delcher, A. L.; Pop, M.; Salzberg, S. L. Gene Prediction with Glimmer for Metagenomic Sequences Augmented by Classification and Clustering. *Nucleic Acids Res* **2012**, *40* (1), e9–e9. https://doi.org/10.1093/nar/gkr1067.

(44) Besemer, J.; Borodovsky, M. GeneMark: Web Software for Gene Finding in Prokaryotes, Eukaryotes and Viruses. *Nucleic Acids Res* **2005**, *33* (Web Server), W451–W454. https://doi.org/10.1093/nar/gki487.

(45) Stanke, M.; Morgenstern, B. AUGUSTUS: A Web Server for Gene Prediction in Eukaryotes That Allows User-Defined Constraints. *Nucleic Acids Res* **2005**, *33* (Web Server), W465–W467. https://doi.org/10.1093/nar/gki458.

(46) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J Mol Biol* **1990**, *215* (3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

(47) Khaldi, N.; Seifuddin, F. T.; Turner, G.; Haft, D.; Nierman, W. C.; Wolfe, K. H.; Fedorova, N. D. SMURF: Genomic Mapping of Fungal Secondary Metabolite Clusters. *Fungal Genetics and Biology* **2010**, *47* (9), 736–741. https://doi.org/10.1016/j.fgb.2010.06.003.

(48) Blin, K.; Shaw, S.; Augustijn, H. E.; Reitz, Z. L.; Biermann, F.; Alanjary, M.; Fetter, A.; Terlouw, B. R.; Metcalf, W. W.; Helfrich, E. J. N.; van Wezel, G. P.; Medema, M. H.; Weber, T. AntiSMASH 7.0: New and Improved Predictions for Detection, Regulation, Chemical Structures and Visualisation. *Nucleic Acids Res* **2023**, *51* (W1), W46–W50. https://doi.org/10.1093/nar/gkad344.

(49) Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R. AntiSMASH: Rapid Identification, Annotation and Analysis of Secondary Metabolite Biosynthesis Gene Clusters in Bacterial and Fungal Genome Sequences. *Nucleic Acids Res* **2011**, *39* (suppl_2), W339–W346. https://doi.org/10.1093/nar/gkr466.

(50) Blin, K.; Kim, H. U.; Medema, M. H.; Weber, T. Recent Development of AntiSMASH and Other Computational Approaches to Mine Secondary Metabolite Biosynthetic Gene Clusters. *Brief Bioinform* **2019**, *20* (4), 1103–1113. https://doi.org/10.1093/bib/bbx146.

(51) Kautsar, S. A.; Suarez Duran, H. G.; Blin, K.; Osbourn, A.; Medema, M. H. PlantiSMASH: Automated Identification, Annotation and Expression Analysis of Plant Biosynthetic Gene Clusters. *Nucleic Acids Res* **2017**, *45* (W1), W55–W63. https://doi.org/10.1093/nar/gkx305.

(52) Belknap, K. C.; Park, C. J.; Barth, B. M.; Andam, C. P. Genome Mining of Biosynthetic and Chemotherapeutic Gene Clusters in Streptomyces Bacteria. *Sci Rep* **2020**, *10* (1), 2003. https://doi.org/10.1038/s41598-020-58904-9.

(53) Skinnider, M. A.; Dejong, C. A.; Rees, P. N.; Johnston, C. W.; Li, H.; Webster, A. L. H.; Wyatt, M. A.; Magarvey, N. A. Genomes to Natural Products PRediction Informatics for Secondary Metabolomes (PRISM). *Nucleic Acids Res* **2015**, *43* (20), 9645-9662. https://doi.org/10.1093/nar/gkv1012.

(54) Cimermancic, P.; Medema, M. H.; Claesen, J.; Kurita, K.; Wieland Brown, L. C.; Mavrommatis, K.; Pati, A.; Godfrey, P. A.; Koehrsen, M.; Clardy, J.; Birren, B. W.; Takano, E.; Sali, A.; Linington, R. G.; Fischbach, M. A. Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. *Cell* **2014**, *158* (2), 412–421. https://doi.org/10.1016/j.cell.2014.06.034.

(55) King, A. M.; Zhang, Z.; Glassey, E.; Siuti, P.; Clardy, J.; Voigt, C. A. Systematic Mining of the Human Microbiome Identifies Antimicrobial Peptides with Diverse Activity Spectra. *Nat Microbiol* **2023**, *8* (12), 2420–2434. https://doi.org/10.1038/s41564-023-01524-6.

(56) Repka, L. M.; Chekan, J. R.; Nair, S. K.; van der Donk, W. A. Mechanistic Understanding of Lanthipeptide Biosynthetic Enzymes*. Chem Rev* **2017**, *117* (8), 5457–5520. https://doi.org/10.1021/acs.chemrev.6b00591.

(57) Kaletta, C.; Entian, K. D. Nisin, a Peptide Antibiotic: Cloning and Sequencing of the NisA Gene and Posttranslational Processing of Its Peptide Product. *J Bacteriol* **1989**, *171* (3), 1597–1601. https://doi.org/10.1128/jb.171.3.1597-1601.1989.

(58) LaMarche, M. J.; Leeds, J. A.; Brewer, J.; Dean, K.; Ding, J.; Dzink-Fox, J.; Gamber, G.; Jain, A.; Kerrigan, R.; Krastel, P.; Lee, K.; Lombardo, F.; McKenney, D.; Neckermann, G.; Osborne, C.; Palestrant, D.; Patane, M. A.; Rann, E. M.; Robinson, Z.; Schmitt, E.; Stams, T.; Tiamfook, S.; Yu, D.; Whitehead, L. Antibacterial and Solubility Optimization of Thiomuracin A. *J Med Chem* **2016**, *59* (14), 6920–6928. https://doi.org/10.1021/acs.jmedchem.6b00726.

(59) Rosengren, K. J.; Clark, R. J.; Daly, N. L.; Göransson, U.; Jones, A.; Craik, D. J. Microcin J25 Has a Threaded Sidechain-to-Backbone Ring Structure and Not a Head-to-Tail Cyclized Backbone. *J Am Chem Soc* **2003**, *125* (41), 12464–12474. https://doi.org/10.1021/ja0367703.

(60) Zhao, G.; Kosek, D.; Liu, H.-B.; Ohlemacher, S. I.; Blackburne, B.; Nikolskaya, A.; Makarova, K. S.; Sun, J.; Barry III, C. E.; Koonin, E. V.; Dyda, F.; Bewley, C. A. Structural Basis for a Dual Function ATP Grasp Ligase That Installs Single and Bicyclic ω-Ester Macrocycles in a New Multicore RiPP Natural Product. *J Am Chem Soc* **2021**, *143* (21), 8056–8068. https://doi.org/10.1021/jacs.1c02316.

(61) Sit, C. S.; van Belkum, M. J.; McKay, R. T.; Worobo, R. W.; Vederas, J. C. The 3D Solution Structure of Thurincin H, a Bacteriocin with Four Sulfur to A-Carbon Crosslinks. *Angewandte Chemie International Edition* **2011**, *50* (37), 8718–8721. https://doi.org/10.1002/anie.201102527.

(62) Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Grande Burgos, M. J.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; van Wezel, G. P.; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van

der Donk, W. A. New Developments in RiPP Discovery, Enzymology and Engineering. *Nat Prod Rep* **2021**, *38* (1), 130–239. https://doi.org/10.1039/D0NP00027B.

(63) Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. A Prevalent Peptide-Binding Domain Guides Ribosomal Natural Product Biosynthesis. *Nat Chem Biol* **2015**, *11* (8), 564–570. https://doi.org/10.1038/nchembio.1856.

(64) de Jong, A.; van Hijum, S. A. F. T.; Bijlsma, J. J. E.; Kok, J.; Kuipers, O. P. BAGEL: A Web-Based Bacteriocin Genome Mining Tool. *Nucleic Acids Res* **2006**, *34* (Web Server), W273–W279. https://doi.org/10.1093/nar/gkl237.

(65) van Heel, A. J.; de Jong, A.; Song, C.; Viel, J. H.; Kok, J.; Kuipers, O. P. BAGEL4: A User-Friendly Web Server to Thoroughly Mine RiPPs and Bacteriocins. *Nucleic Acids Res* **2018**, *46* (W1), W278–W281. https://doi.org/10.1093/nar/gky383.

(66) Agrawal, P.; Khater, S.; Gupta, M.; Sain, N.; Mohanty, D. RiPPMiner: A Bioinformatics Resource for Deciphering Chemical Structures of RiPPs Based on Prediction of Cleavage and Cross-Links. *Nucleic Acids Res* **2017**, *45* (W1), W80–W88. https://doi.org/10.1093/nar/gkx408.

(67) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.; Zakai, U. I.; Mitchell, D. A. A New Genome-Mining Tool Redefines the Lasso Peptide Biosynthetic Landscape. *Nat Chem Biol* **2017**, *13* (5), 470–478. https://doi.org/10.1038/nchembio.2319.

(68) Hudson, G. A.; Burkhart, B. J.; DiCaprio, A. J.; Schwalen, C. J.; Kille, B.; Pogorelov, T. V.; Mitchell, D. A. Bioinformatic Mapping of Radical S -Adenosylmethionine-Dependent Ribosomally Synthesized and Post-Translationally Modified Peptides Identifies New Cα, Cβ, and Cγ-Linked Thioether-Containing Peptides. *J Am Chem Soc* **2019**, *141* (20), 8228–8238. https://doi.org/10.1021/jacs.9b01519.

(69) Schwalen, C. J.; Hudson, G. A.; Kille, B.; Mitchell, D. A. Bioinformatic Expansion and Discovery of Thiopeptide Antibiotics. *J Am Chem Soc* **2018**, *140* (30), 9494–9501. https://doi.org/10.1021/jacs.8b03896.

(70) Merwin, N. J.; Mousa, W. K.; Dejong, C. A.; Skinnider, M. A.; Cannon, M. J.; Li, H.; Dial, K.; Gunabalasingam, M.; Johnston, C.; Magarvey, N. A. DeepRiPP Integrates Multiomics Data to Automate Discovery of Novel Ribosomally Synthesized Natural Products. *Proceedings of the National Academy of Sciences* **2020**, *117* (1), 371–380. https://doi.org/10.1073/pnas.1901493116.

(71) Kennedy, J.; Auclair, K.; Kendrew, S. G.; Park, C.; Vederas, J. C.; Richard Hutchinson, C. Modulation of Polyketide Synthase Activity by Accessory Proteins During Lovastatin Biosynthesis. *Science (1979)* **1999**, *284* (5418), 1368–1372. https://doi.org/10.1126/science.284.5418.1368.

(72) Lin, H.-C.; Chooi, Y.-H.; Dhingra, S.; Xu, W.; Calvo, A. M.; Tang, Y. The Fumagillin Biosynthetic Gene Cluster in Aspergillus Fumigatus Encodes a Cryptic Terpene Cyclase Involved in the Formation of β- Trans -Bergamotene. *J Am Chem Soc* **2013**, *135* (12), 4616–4619. https://doi.org/10.1021/ja312503y.

(73) Peterson, R. M.; Huang, T.; Rudolf, J. D.; Smanski, M. J.; Shen, B. Mechanisms of Self-Resistance in the Platensimycin- and Platencin-Producing Streptomyces Platensis MA7327 and MA7339 Strains. *Chem Biol* **2014**, *21* (3), 389–397. https://doi.org/10.1016/j.chembiol.2014.01.005.

(74) Kling, A.; Lukat, P.; Almeida, D. V.; Bauer, A.; Fontaine, E.; Sordello, S.; Zaburannyi, N.; Herrmann, J.; Wenzel, S. C.; König, C.; Ammerman, N. C.; Barrio, M. B.; Borchers, K.; Bordon-Pallier, F.; Brönstrup, M.; Courtemanche, G.; Gerlitz, M.; Geslin, M.; Hammann, P.; Heinz, D. W.; Hoffmann, H.; Klieber, S.; Kohlmann, M.; Kurz, M.; Lair, C.; Matter, H.; Nuermberger, E.; Tyagi, S.; Fraisse, L.; Grosset, J. H.; Lagrange, S.; Müller, R. Targeting DnaN for Tuberculosis Therapy Using Novel Griselimycins. *Science (1979)* **2015**, *348* (6239), 1106–1112. https://doi.org/10.1126/science.aaa4690.

(75) Xie, L.; Zang, X.; Cheng, W.; Zhang, Z.; Zhou, J.; Chen, M.; Tang, Y. Harzianic Acid from Trichoderma Afroharzianum Is a Natural Product Inhibitor of Acetohydroxyacid Synthase. *J Am Chem Soc* **2021**, *143* (25), 9575–9584. https://doi.org/10.1021/jacs.1c03988.

(76) Sato, M.; Sakano, S.; Nakahara, M.; Tamura, Y.; Hara, K.; Hashimoto, H.; Tang, Y.; Watanabe, K. Uncommon Arrangement of Self-Resistance Allows Biosynthesis of de Novo Purine Biosynthesis Inhibitor That Acts as an Immunosuppressor. *J Am Chem Soc* **2023**, *145* (49), 26883–26889. https://doi.org/10.1021/jacs.3c09600.

(77) Hu, Z.; Awakawa, T.; Ma, Z.; Abe, I. Aminoacyl Sulfonamide Assembly in SB-203208 Biosynthesis. *Nat Commun* **2019**, *10* (1), 184. https://doi.org/10.1038/s41467-018-08093-x.

(78) Zhang, W.; Zhang, X.; Feng, D.; Liang, Y.; Wu, Z.; Du, S.; Zhou, Y.; Geng, C.; Men, P.; Fu, C.; Huang, X.; Lu, X. Discovery of a Unique Flavonoid Biosynthesis Mechanism in Fungi by Genome Mining. *Angewandte Chemie International Edition* **2023**, *62* (12), e202215529. https://doi.org/10.1002/anie.202215529.

(79) Liu, N.; Abramyan, E. D.; Cheng, W.; Perlatti, B.; Harvey, C. J. B.; Bills, G. F.; Tang, Y. Targeted Genome Mining Reveals the Biosynthetic Gene Clusters of Natural Product CYP51 Inhibitors. *J Am Chem Soc* **2021**, *143* (16), 6043–6047. https://doi.org/10.1021/jacs.1c01516.

(80) Tang, X.; Li, J.; Millán-Aguiñaga, N.; Zhang, J. J.; O'Neill, E. C.; Ugalde, J. A.; Jensen, P. R.; Mantovani, S. M.; Moore, B. S. Identification of Thiotetronic Acid Antibiotic Biosynthetic Pathways by Target-Directed Genome Mining. *ACS Chem Biol* **2015**, *10* (12), 2841–2849. https://doi.org/10.1021/acschembio.5b00658.

(81) Yan, Y.; Liu, Q.; Zang, X.; Yuan, S.; Bat-Erdene, U.; Nguyen, C.; Gan, J.; Zhou, J.; Jacobsen, S. E.; Tang, Y. Resistance-Gene-Directed Discovery of a Natural-Product Herbicide with a New Mode of Action. *Nature* **2018**, *559* (7714), 415–418. https://doi.org/10.1038/s41586-018-0319-4.

(82) Dunbar, K. L.; Perlatti, B.; Liu, N.; Cornelius, A.; Mummau, D.; Chiang, Y.-M.; Hon, L.; Nimavat, M.; Pallas, J.; Kordes, S.; Ng, H. L.; Harvey, C. J. B. Resistance Gene–Guided Genome Mining Reveals the Roseopurpurins as Inhibitors of Cyclin-Dependent Kinases. *Proceedings of the National Academy of Sciences* **2023**, *120* (48), e2310522120. https://doi.org/10.1073/pnas.2310522120.

(83) Yeh, H.-H.; Ahuja, M.; Chiang, Y.-M.; Oakley, C. E.; Moore, S.; Yoon, O.; Hajovsky, H.; Bok, J.-W.; Keller, N. P.; Wang, C. C. C.; Oakley, B. R. Resistance Gene-Guided Genome Mining: Serial Promoter Exchanges in Aspergillus Nidulans Reveal the Biosynthetic Pathway for Fellutamide B, a Proteasome Inhibitor. *ACS Chem Biol* **2016**, *11* (8), 2275–2284. https://doi.org/10.1021/acschembio.6b00213.

(84) Culp, E. J.; Sychantha, D.; Hobson, C.; Pawlowski, A. C.; Prehna, G.; Wright, G. D. ClpP Inhibitors Are Produced by a Widespread Family of Bacterial Gene Clusters. *Nat Microbiol* **2022**, *7* (3), 451–462. https://doi.org/10.1038/s41564-022-01073-4.

(85) Panter, F.; Krug, D.; Baumann, S.; Müller, R. Self-Resistance Guided Genome Mining Uncovers New Topoisomerase Inhibitors from Myxobacteria. *Chem Sci* **2018**, *9* (21), 4898–4908. https://doi.org/10.1039/C8SC01325J.

(86) Yarlagadda, V.; Medina, R.; Johnson, T. A.; Koteva, K. P.; Cox, G.; Thaker, M. N.; Wright, G. D. Resistance-Guided Discovery of Elfamycin Antibiotic Producers with Antigonococcal Activity. *ACS Infect Dis* **2020**, *6* (12), 3163–3173. https://doi.org/10.1021/acsinfecdis.0c00467.

(87) Kjærbølling, I.; Vesth, T.; Andersen, M. R. Resistance Gene-Directed Genome Mining of 50 Aspergillus Species. *mSystems* **2019**, *4* (4), 10-1128. https://doi.org/10.1128/mSystems.00085-19.

(88) Jenkinson, C. B.; Podgorny, A. R.; Zhong, C.; Oakley, B. R. Computer-Aided, Resistance Gene-Guided Genome Mining for Proteasome and HMG-CoA Reductase Inhibitors. *J Ind Microbiol Biotechnol* **2023**, *50* (1), kuad045. https://doi.org/10.1093/jimb/kuad045.

(89) Yılmaz, T. M.; Mungan, M. D.; Berasategui, A.; Ziemert, N. FunARTS, the Fungal BioActive Compound Resistant Target Seeker, an Exploration Engine for Target-Directed Genome Mining in Fungi. *Nucleic Acids Res* **2023**, *51* (W1), W191–W197. https://doi.org/10.1093/nar/gkad386.

(90) Kautsar, S. A.; van der Hooft, J. J. J.; de Ridder, D.; Medema, M. H. BiG-SLiCE: A Highly Scalable Tool Maps the Diversity of 1.2 Million Biosynthetic Gene Clusters. *Gigascience* **2021**, *10* (1), giaa154. https://doi.org/10.1093/gigascience/giaa154.

(91) Epstein, S. C.; Charkoudian, L. K.; Medema, M. H. A Standardized Workflow for Submitting Data to the Minimum Information about a Biosynthetic Gene Cluster (MIBiG) Repository: Prospects for Research-Based Educational Experiences. *Stand Genomic Sci* **2018**, *13* (1), 16. https://doi.org/10.1186/s40793-018-0318-y.

(92) Terlouw, B. R.; Blin, K.; Navarro-Muñoz, J. C.; Avalon, N. E.; Chevrette, M. G.; Egbert, S.; Lee, S.; Meijer, D.; Recchia, M. J. J.; Reitz, Z. L.; van Santen, J. A.; Selem-Mojica, N.; Tørring, T.; Zaroubi, L.; Alanjary, M.; Aleti, G.; Aguilar, C.; Al-Salihi, S. A. A.; Augustijn, H. E.; Avelar-Rivas, J. A.; Avitia-Domínguez, L. A.; Barona-Gómez, F.; Bernaldo-Agüero, J.; Bielinski, V. A.; Biermann, F.; Booth, T. J.; Carrion Bravo, V. J.; Castelo-Branco, R.; Chagas, F. O.; Cruz-Morales, P.; Du, C.; Duncan, K. R.; Gavriilidou, A.; Gayrard, D.; Gutiérrez-García, K.; Haslinger, K.; Helfrich, E. J. N.; van der Hooft, J. J. J.; Jati, A. P.; Kalkreuter, E.; Kalyvas, N.; Kang, K. B.; Kautsar, S.; Kim, W.; Kunjapur, A. M.; Li, Y.-X.; Lin, G.-M.; Loureiro, C.; Louwen, J. J. R.; Louwen, N. L. L.; Lund, G.; Parra, J.; Philmus, B.; Pourmohsenin, B.; Pronk, L. J. U.; Rego, A.; Rex, D. A. B.; Robinson, S.; Rosas-Becerra, L. R.; Roxborough, E. T.; Schorn, M. A.; Scobie, D. J.; Singh, K. S.; Sokolova, N.; Tang, X.; Udwary, D.; Vigneshwari, A.; Vind, K.; Vromans, S. P. J. M.; Waschulin, V.; Williams, S. E.; Winter, J. M.; Witte, T. E.; Xie, H.; Yang, D.; Yu, J.; Zdouc, M.; Zhong, Z.; Collemare, J.; Linington, R. G.; Weber, T.; Medema, M. H. MIBiG 3.0: A Community-Driven Effort to Annotate Experimentally Validated Biosynthetic Gene Clusters. *Nucleic Acids Res* **2023**, *51* (D1), D603–D610. https://doi.org/10.1093/nar/gkac1049.

(93) Zdouc, M. M.; Blin, K.; Louwen, N. L. L.; Navarro, J.; Loureiro, C.; Bader, C. D.; Bailey, C. B.; Barra, L.; Booth, T. J.; Bozhüyük, K. A. J.; Cediel-Becerra, J. D. D.; Charlop-Powers, Z.; Chevrette, M. G.; Chooi, Y. H.; D'Agostino, P. M.; de Rond, T.; Del Pup, E.; Duncan, K. R.; Gu, W.; Hanif, N.; Helfrich, E. J. N.; Jenner, M.; Katsuyama, Y.; Korenskaia, A.; Krug, D.; Libis, V.; Lund, G. A.; Mantri, S.; Morgan, K. D.; Owen, C.; Phan, C.-S.; Philmus, B.; Reitz, Z. L.; Robinson, S. L.; Singh, K. S.; Teufel, R.; Tong, Y.; Tugizimana, F.; Ulanova, D.; Winter, J. M.; Aguilar, C.; Akiyama, D. Y.; Al-Salihi, S. A. A.; Alanjary, M.; Alberti, F.; Aleti, G.; Alharthi, S. A.; Rojo, M. Y. A.; Arishi, A. A.; Augustijn, H. E.; Avalon, N. E.; Avelar-Rivas, J. A.; Axt, K. K.; Barbieri, H. B.; Barbosa, J. C. J.; Barboza Segato, L. G.; Barrett, S. E.; Baunach, M.; Beemelmanns, C.; Beqaj, D.; Berger, T.; Bernaldo-Agüero, J.; Bettenbühl, S. M.; Bielinski, V. A.; Biermann, F.; Borges, R. M.; Borriss, R.; Breitenbach, M.; Bretscher, K. M.; Brigham, M. W.; Buedenbender, L.; Bulcock, B. W.; Cano-Prieto, C.; Capela, J.; Carrion, V. J.; Carter, R. S.; Castelo-Branco, R.; Castro-Falcón, G.; Chagas, F. O.; Charria-Girón, E.; Chaudhri, A. A.; Chaudhry, V.; Choi, H.; Choi, Y.; Choupannejad, R.; Chromy, J.; Donahey, M. S. C.; Collemare, J.; Connolly, J. A.; Creamer, K. E.; Crüsemann, M.; Cruz, A. A.; Cumsille, A.; Dallery, J.-F.; Damas-Ramos, L. C.; Damiani, T.; de Kruijff, M.; Martín, B. D.; Sala, G. Della; Dillen, J.; Doering, D. T.; Dommaraju, S. R.; Durusu, S.; Egbert, S.; Ellerhorst, M.; Faussurier, B.; Fetter, A.; Feuermann, M.; Fewer, D. P.; Foldi, J.; Frediansyah, A.; Garza, E. A.; Gavriilidou, A.; Gentile, A.; Gerke, J.; Gerstmans, H.; Gomez-Escribano, J. P.; González-Salazar, L. A.; Grayson, N. E.; Greco, C.; Gomez, J. E. G.; Guerra, S.; Flores, S. G.; Gurevich, A.; Gutiérrez-García, K.; Hart, L.; Haslinger, K.; He, B.; Hebra, T.; Hemmann, J. L.; Hindra, H.; Höing, L.; Holland, D. C.; Holme, J. E.; Horch, T.; Hrab, P.; Hu, J.; Huynh, T.-H.; Hwang, J.-Y.; Iacovelli, R.; Iftime, D.; Iorio, M.; Jayachandran, S.; Jeong, E.; Jing, J.; Jung, J. J.; Kakumu, Y.; Kalkreuter, E.; Kang, K. Bin; Kang, S.; Kim, W.; Kim, G. J.; Kim, H.; Kim, H. U.; Klapper, M.; Koetsier, R. A.; Kollten, C.; Kovács, Á. T.; Kriukova, Y.; Kubach, N.; Kunjapur, A. M.; Kushnareva, A. K.; Kust, A.; Lamber, J.; Larralde, M.; Larsen, N. J.; Launay, A. P.; Le, N.-T.-H.; Lebeer, S.; Lee, B. T.; Lee, K.; Lev, K. L.; Li, S.-M.; Li, Y.-X.; Licona-Cassani, C.; Lien, A.; Liu, J.; Lopez, J. A. V.; Machushynets, N. V.; Macias, M. I.; Mahmud, T.; Maleckis, M.; Martinez-Martinez, A. M.; Mast, Y.; Maximo, M. F.; McBride, C. M.; McLellan, R. M.; Bhatt, K. M.; Melkonian, C.; Merrild, A.; Metsä-Ketelä, M.; Mitchell, D. A.; Müller, A. V.; Nguyen, G.-S.; Nguyen, H. T.; Niedermeyer, T. H. J.; O'Hare, J. H.; Ossowicki, A.; Ostash, B. O.; Otani, H.; Padva, L.; Paliyal, S.; Pan, X.; Panghal, M.; Parade, D. S.; Park, J.; Parra, J.; Rubio, M. P.; Pham, H. T.; Pidot, S. J.; Piel, J.; Pourmohsenin, B.; Rakhmanov, M.; Ramesh, S.; Rasmussen, M. H.; Rego, A.; Reher, R.; Rice, A. J.; Rigolet, A.; Romero-Otero, A.; Rosas-Becerra, L. R.; Rosiles, P. Y.; Rutz, A.; Ryu, B.; Sahadeo, L.-A.; Saldanha, M.; Salvi, L.; Sánchez-Carvajal, E.; Santos-Medellin, C.; Sbaraini, N.; Schoellhorn, S. M.; Schumm, C.; Sehnal, L.; Selem, N.; Shah, A. D.; Shishido, T. K.; Sieber, S.; Silviani, V.; Singh, G.; Singh, H.; Sokolova, N.; Sonnenschein, E. C.; Sosio, M.; Sowa, S. T.; Steffen, K.; Stegmann, E.; Streiff, A. B.; Strüder, A.; Surup, F.; Svenningsen, T.; Sweeney, D.; Szenei, J.; Tagirdzhanov, A.; Tan, B.; Tarnowski, M. J.; Terlouw, B. R.; Rey, T.; Thome, N. U.; Torres Ortega, L. R.; Tørring, T.; Trindade, M.; Truman, A. W.; Tvilum, M.; Udwary, D. W.; Ulbricht, C.; Vader, L.; van Wezel, G. P.; Walmsley, M.; Warnasinghe, R.; Weddeling, H. G.; Weir, A. N. M.; Williams, K.; Williams, S. E.; Witte, T. E.; Rocca, S. M. W.; Yamada, K.; Yang, D.; Yang, D.; Yu, J.; Zhou, Z.; Ziemert, N.; Zimmer, L.; Zimmermann, A.; Zimmermann, C.; van der Hooft, J. J. J.; Linington, R. G.; Weber, T.; Medema, M. H. MIBiG 4.0: Advancing Biosynthetic Gene Cluster Curation through Global Collaboration. *Nucleic Acids Res* **2024**, gkae1115. https://doi.org/10.1093/nar/gkae1115.

(94) Fullam, A.; Letunic, I.; Schmidt, T. S. B.; Ducarmon, Q. R.; Karcher, N.; Khedkar, S.; Kuhn, M.; Larralde, M.; Maistrenko, O. M.; Malfertheiner, L.; Milanese, A.; Rodrigues, J. F. M.; Sanchis-López, C.; Schudoma, C.; Szklarczyk, D.; Sunagawa, S.; Zeller, G.; Huerta-Cepas, J.; von Mering, C.; Bork, P.; Mende, D. R. ProGenomes3: Approaching One Million Accurately and Consistently Annotated High-Quality Prokaryotic Genomes. *Nucleic Acids Res* **2023**, *51* (D1), D760–D766. https://doi.org/10.1093/nar/gkac1078.

(95) Navarro-Muñoz, J. C.; Selem-Mojica, N.; Mullowney, M. W.; Kautsar, S. A.; Tryon, J. H.; Parkinson, E. I.; De Los Santos, E. L. C.; Yeong, M.; Cruz-Morales, P.; Abubucker, S.; Roeters, A.; Lokhorst, W.; Fernandez-Guerra, A.; Cappelini, L. T. D.; Goering, A. W.; Thomson, R. J.; Metcalf, W. W.; Kelleher, N. L.; Barona-Gomez, F.; Medema, M. H. A Computational Framework to Explore Large-Scale Biosynthetic Diversity. *Nat Chem Biol* **2020**, *16* (1), 60–68. https://doi.org/10.1038/s41589-019-0400-9.

(96) Kautsar, S. A.; Blin, K.; Shaw, S.; Weber, T.; Medema, M. H. BiG-FAM: The Biosynthetic Gene Cluster Families Database. *Nucleic Acids Res* **2021**, *49* (D1), D490–D497. https://doi.org/10.1093/nar/gkaa812.

(97) Yuan, Y.; Shi, C.; Zhao, H. Machine Learning-Enabled Genome Mining and Bioactivity Prediction of Natural Products. *ACS Synth Biol* **2023**, *12* (9), 2650–2662. https://doi.org/10.1021/acssynbio.3c00234.

(98) Carroll, L. M.; Larralde, M.; Fleck, J. S.; Ponnudurai, R.; Milanese, A.; Barazzone, E. C.; Zeller, G. Accurate de Novo Identification of Biosynthetic Gene Clusters with GECCO. *bioRxiv* **2021**. Accessed 2024-07-15.

(99) Hannigan, G. D.; Prihoda, D.; Palicka, A.; Soukup, J.; Klempir, O.; Rampula, L.; Durcak, J.; Wurst, M.; Kotowski, J.; Chang, D.; Wang, R.; Piizzi, G.; Temesi, G.; Hazuda, D. J.; Woelk, C. H.; Bitton, D. A. A Deep Learning Genome-Mining Strategy for Biosynthetic Gene Cluster Prediction. *Nucleic Acids Res* **2019**, *47* (18), e110–e110. https://doi.org/10.1093/nar/gkz654.

(100) Rios-Martinez, C.; Bhattacharya, N.; Amini, A. P.; Crawford, L.; Yang, K. K. Deep Self-Supervised Learning for Biosynthetic Gene Cluster Detection and Product Classification. *PLoS Comput Biol* **2023**, *19* (5), e1011162. https://doi.org/10.1371/journal.pcbi.1011162.

(101) Romero, P. A.; Arnold, F. H. Exploring Protein Fitness Landscapes by Directed Evolution. *Nat Rev Mol Cell Biol* **2009**, *10* (12), 866–876. https://doi.org/10.1038/nrm2805.

(102) Schor, R.; Schotte, C.; Wibberg, D.; Kalinowski, J.; Cox, R. J. Three Previously Unrecognised Classes of Biosynthetic Enzymes Revealed during the Production of Xenovulene A. *Nat Commun* **2018**, *9* (1), 1963. https://doi.org/10.1038/s41467-018-04364-9.

(103) Zhang, L.; Zhang, B.; Zhu, A.; Liu, S. H.; Wu, R.; Zhang, X.; Xu, Z.; Tan, R. X.; Ge, H. M. Biosynthesis of Phomactin Platelet Activating Factor Antagonist Requires a Two-Enzyme Cascade. *Angewandte Chemie International Edition* **2023**, *62* (47), e202312996. https://doi.org/10.1002/anie.202312996.

(104) Du, Y.-L.; He, H.-Y.; Higgins, M. A.; Ryan, K. S. A Heme-Dependent Enzyme Forms the Nitrogen–Nitrogen Bond in Piperazate. *Nat Chem Biol* **2017**, *13* (8), 836–838. https://doi.org/10.1038/nchembio.2411.

(105) Shi, J.; Zang, X.; Zhao, Z.; Shen, Z.; Li, W.; Zhao, G.; Zhou, J.; Du, Y.-L. Conserved Enzymatic Cascade for Bacterial Azoxy Biosynthesis. *J Am Chem Soc* **2023**, *145* (49), 27131–27139. https://doi.org/10.1021/jacs.3c12018.

(106) Lichman, B. R.; Godden, G. T.; Hamilton, J. P.; Palmer, L.; Kamileen, M. O.; Zhao, D.; Vaillancourt, B.; Wood, J. C.; Sun, M.; Kinser, T. J.; Henry, L. K.; Rodriguez-Lopez, C.; Dudareva, N.; Soltis, D. E.; Soltis, P. S.; Buell, C. R.; O'Connor, S. E. The Evolutionary Origins of the Cat Attractant Nepetalactone in Catnip. *Sci Adv* **2020**, *6* (20), eaba0721. https://doi.org/10.1126/sciadv.aba0721.

(107) Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; Bodenstein, S. W.; Evans, D. A.; Hung, C.-C.; O'Neill, M.; Reiman, D.; Tunyasuvunakool, K.; Wu, Z.; Žemgulytė, A.; Arvaniti, E.; Beattie, C.; Bertolli, O.; Bridgland, A.; Cherepanov, A.; Congreve, M.; Cowen-Rivers, A. I.; Cowie, A.; Figurnov, M.; Fuchs, F. B.; Gladman, H.; Jain, R.; Khan, Y. A.; Low, C. M. R.; Perlin, K.; Potapenko, A.; Savy, P.; Singh, S.; Stecula, A.; Thillaisundaram, A.; Tong, C.; Yakneen, S.; Zhong, E. D.; Zielinski, M.; Žídek, A.; Bapst, V.;

Kohli, P.; Jaderberg, M.; Hassabis, D.; Jumper, J. M. Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3. *Nature* **2024**, *630* (8016), 493–500. https://doi.org/10.1038/s41586-024-07487-w.

(108) Senior, A. W.; Evans, R.; Jumper, J.; Kirkpatrick, J.; Sifre, L.; Green, T.; Qin, C.; Žídek, A.; Nelson, A. W. R.; Bridgland, A.; Penedones, H.; Petersen, S.; Simonyan, K.; Crossan, S.; Kohli, P.; Jones, D. T.; Silver, D.; Kavukcuoglu, K.; Hassabis, D. Improved Protein Structure Prediction Using Potentials from Deep Learning. *Nature* **2020**, *577* (7792), 706–710. https://doi.org/10.1038/s41586-019-1923-7.

(109) Holm, L. Dali Server: Structural Unification of Protein Families. Nucleic Acids Res **2022**, 50 (W1), W210–W215. https://doi.org/10.1093/nar/gkac387.

(110) Gligorijević, V.; Renfrew, P. D.; Kosciolek, T.; Leman, J. K.; Berenberg, D.; Vatanen, T.; Chandler, C.; Taylor, B. C.; Fisk, I. M.; Vlamakis, H.; Xavier, R. J.; Knight, R.; Cho, K.; Bonneau, R. Structure-Based Protein Function Prediction Using Graph Convolutional Networks. *Nat Commun* **2021**, *12* (1), 3168. https://doi.org/10.1038/s41467-021-23303-9.

(111) Sanderson, T.; Bileschi, M. L.; Belanger, D.; Colwell, L. J. ProteInfer, Deep Neural Networks for Protein Functional Inference. *Elife* **2023**, *12*, e80942. https://doi.org/10.7554/eLife.80942.

(112) Yu, T.; Cui, H.; Li, J. C.; Luo, Y.; Jiang, G.; Zhao, H. Enzyme Function Prediction Using Contrastive Learning. *Science (1979)* **2023**, *379* (6639), 1358–1363. https://doi.org/10.1126/science.adf2465.

(113) Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A Web Tool for Generating Protein Sequence Similarity Networks. Biochimica et Biophysica Acta (BBA) - *Proteins and Proteomics* **2015**, *1854* (8), 1019–1037. https://doi.org/10.1016/j.bbapap.2015.04.015.

(114) Zallot, R.; Oberg, N.; Gerlt, J. A. The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. *Biochemistry* **2019**, *58* (41), 4169–4182. https://doi.org/10.1021/acs.biochem.9b00735.

(115) Copp, J. N.; Akiva, E.; Babbitt, P. C.; Tokuriki, N. Revealing Unexplored Sequence-Function Space Using Sequence Similarity Networks. *Biochemistry* **2018**, *57* (31), 4651–4662. https://doi.org/10.1021/acs.biochem.8b00473.

(116) Grenade, N. L.; Chiriac, D. S.; Pasternak, A. R. O.; Babulic, J. L.; Rowland, B. E.; Howe, G. W.; Ross, A. C. Discovery of a Tambjamine Gene Cluster in Streptomyces Suggests Convergent Evolution in Bipyrrole Natural Product Biosynthesis. *ACS Chem Biol* **2023**, *18* (2), 223–229. https://doi.org/10.1021/acschembio.2c00685.

(117) Hu, D. X.; Withall, D. M.; Challis, G. L.; Thomson, R. J. Structure, Chemical Synthesis, and Biosynthesis of Prodiginine Natural Products. *Chem Rev* **2016**, *116* (14), 7818–7853. https://doi.org/10.1021/acs.chemrev.6b00024.

(118) Bode, H. B.; Bethe, B.; Höfs, R.; Zeeck, A. Big Effects from Small Changes: Possible Ways to Explore Nature's Chemical Diversity. *ChemBioChem* **2002**, *3* (7), 619. https://doi.org/10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9.

(119) Rutledge, P. J.; Challis, G. L. Discovery of Microbial Natural Products by Activation of Silent Biosynthetic Gene Clusters. *Nat Rev Microbiol* **2015**, *13* (8), 509–523. https://doi.org/10.1038/nrmicro3496.

(120) Lyu, H.-N.; Liu, H.-W.; Keller, N. P.; Yin, W.-B. Harnessing Diverse Transcriptional Regulators for Natural Product Discovery in Fungi. *Nat Prod Rep* **2020**, *37* (1), 6–16. https://doi.org/10.1039/C8NP00027A.

(121) Wilkinson, C. J.; Hughes-Thomas, Z. A.; Martin, C. J.; Böhm, I.; Mironenko, T.; Deacon, M.; Wheatcroft, M.; Wirtz, G.; Staunton, J.; Leadlay, P. F. Increasing the Efficiency of Heterologous Promoters in Actinomycetes. *J Mol Microbiol* Biotechnol **2002**, *4* (4), 417–426.

(122) Tay, D. W. P.; Tan, L. L.; Heng, E.; Zulkarnain, N.; Ching, K. C.; Wibowo, M.; Chin, E. J.; Tan, Z. Y. Q.; Leong, C. Y.; Ng, V. W. P.; Yang, L. K.; Seow, D. C. S.; Lim, Y. W.; Koh, W.; Koduru, L.; Kanagasundaram, Y.; Ng, S. B.; Lim, Y. H.; Wong, F. T. Exploring a General Multi-Pronged Activation Strategy for Natural Product Discovery in Actinomycetes. *Commun Biol* **2024**, *7* (1), 50. https://doi.org/10.1038/s42003-023-05648-7.

(123) Rued, B. E.; Covington, B. C.; Bushin, L. B.; Szewczyk, G.; Laczkovich, I.; Seyedsayamdost, M. R.; Federle, M. J. Quorum Sensing in Streptococcus Mutans Regulates Production of Tryglysin, a Novel RaS-RiPP Antimicrobial Compound. *mBio* **2021**, *12* (2), 10-1128. https://doi.org/10.1128/mBio.02688-20.

(124) Mukherji, R.; Zhang, S.; Chowdhury, S.; Stallforth, P. Chimeric LuxR Transcription Factors Rewire Natural Product Regulation. *Angewandte Chemie International Edition* **2020**, *59* (15), 6192–6195. https://doi.org/10.1002/anie.201914449.

(125) Seyedsayamdost, M. R. High-Throughput Platform for the Discovery of Elicitors of Silent Bacterial Gene Clusters. *Proceedings of the National Academy of Sciences* **2014**, *111* (20), 7266–7271. https://doi.org/10.1073/pnas.1400019111.

(126) Biggins, J. B.; Ternei, M. A.; Brady, S. F. Malleilactone, a Polyketide Synthase-Derived Virulence Factor Encoded by the Cryptic Secondary Metabolome of Burkholderia Pseudomallei Group Pathogens. *J Am Chem Soc* **2012**, *134* (32), 13192–13195. https://doi.org/10.1021/ja3052156.

(127) Xu, F.; Nazari, B.; Moon, K.; Bushin, L. B.; Seyedsayamdost, M. R. Discovery of a Cryptic Antifungal Compound from Streptomyces Albus J1074 Using High-Throughput Elicitor Screens. *J Am Chem Soc* **2017**, *139* (27), 9203–9212. https://doi.org/10.1021/jacs.7b02716.

(128) Lee, S. R.; Gallant, É.; Seyedsayamdost, M. R. Discovery of Cryptic Natural Products Using High-Throughput Elicitor Screening on Agar Media. *Biochemistry* **2024**. https://doi.org/10.1021/acs.biochem.4c00659.

(129) Xu, F.; Wu, Y.; Zhang, C.; Davis, K. M.; Moon, K.; Bushin, L. B.; Seyedsayamdost, M. R. A Genetics-Free Method for High-Throughput Discovery of Cryptic Microbial Metabolites. *Nat Chem Biol* **2019**, *15* (2), 161–168. https://doi.org/10.1038/s41589-018-0193-2.

(130) Chioti, V. T.; McWhorter, K. L.; Blue, T. C.; Li, Y.; Xu, F.; Jeffrey, P. D.; Davis, K. M.; Seyedsayamdost, M. R. Potent and Specific Antibiotic Combination Therapy against Clostridioides Difficile. *Nat Chem Biol* **2024**, *20* (7), 924–933. https://doi.org/10.1038/s41589-024-01651-z.

(131) Han, E. J.; Lee, S. R.; Townsend, C. A.; Seyedsayamdost, M. R. Targeted Discovery of Cryptic Enediyne Natural Products via FRET-Coupled High-Throughput Elicitor Screening. *ACS Chem Biol* **2023**, *18* (8), 1854–1862. https://doi.org/10.1021/acschembio.3c00281.

(132) Shima, J.; Hesketh, A.; Okamoto, S.; Kawamoto, S.; Ochi, K. Induction of Actinorhodin Production by RpsL (Encoding Ribosomal Protein S12) Mutations That Confer Streptomycin Resistance in Streptomyces Lividans and Streptomyces Coelicolor A3(2). *J Bacteriol* **1996**, *178* (24), 7276–7284. https://doi.org/10.1128/jb.178.24.7276-7284.1996.

(133) Dalebroux, Z. D.; Swanson, M. S. PpGpp: Magic beyond RNA Polymerase. *Nat Rev Microbiol* **2012**, *10* (3), 203–212. https://doi.org/10.1038/nrmicro2720.

(134) Zhang, Q.; Ren, J.-W.; Wang, W.; Zhai, J.; Yang, J.; Liu, N.; Huang, Y.; Chen, Y.; Pan, G.; Fan, K. A Versatile Transcription–Translation in One Approach for Activation of Cryptic Biosynthetic Gene Clusters. *ACS Chem Biol* **2020**, *15* (9), 2551–2557. https://doi.org/10.1021/acschembio.0c00581.

(135) Zhu, S.; Duan, Y.; Huang, Y. The Application of Ribosome Engineering to Natural Product Discovery and Yield Improvement in Streptomyces. *Antibiotics* **2019**, *8* (3), 133. https://doi.org/10.3390/antibiotics8030133.

(136) Guo, F.; Xiang, S.; Li, L.; Wang, B.; Rajasärkkä, J.; Gröndahl-Yli-Hannuksela, K.; Ai, G.; Metsä-Ketelä, M.; Yang, K. Targeted Activation of Silent Natural Product Biosynthesis Pathways by Reporter-Guided Mutant Selection. *Metab Eng* **2015**, *28*, 134–142. https://doi.org/10.1016/j.ymben.2014.12.006.

(137) Wang, X.; Zhou, H.; Chen, H.; Jing, X.; Zheng, W.; Li, R.; Sun, T.; Liu, J.; Fu, J.; Huo, L.; Li, Y.; Shen, Y.; Ding, X.; Müller, R.; Bian, X.; Zhang, Y. Discovery of Recombinases Enables Genome Mining of Cryptic Biosynthetic Gene Clusters in Burkholderiales Species. *Proceedings of the National Academy of Sciences* **2018**, *115* (18), E4255-E4263. https://doi.org/10.1073/pnas.1720941115.

(138) Jung, B.; Kim, S.; Lee, J. Microcyle Conidiation in Filamentous Fungi. *Mycobiology* **2014**, *42* (1), 1–5. https://doi.org/10.5941/MYCO.2014.42.1.1.

(139) Gaynes, R. The Discovery of Penicillin—New Insights After More Than 75 Years of Clinical Use. *Emerg Infect Dis* **2017**, *23* (5), 849–853. https://doi.org/10.3201/eid2305.161556.

(140) Brakhage, A. A. Regulation of Fungal Secondary Metabolism. Nat Rev *Microbiol* **2013**, *11* (1), 21–32. https://doi.org/10.1038/nrmicro2916.

(141) Bergmann, S.; Schümann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. Genomics-Driven Discovery of PKS-NRPS Hybrid Metabolites from Aspergillus Nidulans. *Nat Chem Biol* **2007**, *3* (4), 213–217. https://doi.org/10.1038/nchembio869.

(142) Sato, M.; Dander, J. E.; Sato, C.; Hung, Y.-S.; Gao, S.-S.; Tang, M.-C.; Hang, L.; Winter, J. M.; Garg, N. K.; Watanabe, K.; Tang, Y. Collaborative Biosynthesis of Maleimide- and Succinimide-Containing Natural Products by Fungal Polyketide Megasynthases. *J Am Chem Soc* **2017**, *139* (15), 5317–5320. https://doi.org/10.1021/jacs.7b02432.

(143) Zhang, T.; Wan, J.; Zhan, Z.; Bai, J.; Liu, B.; Hu, Y. Activation of an Unconventional Meroterpenoid Gene Cluster in Neosartorya Glabra Leads to the Production of New Berkeleyacetals. *Acta Pharm Sin B* **2018**, *8* (3), 478–487. https://doi.org/10.1016/j.apsb.2017.12.005.

(144) Bok, J. W.; Keller, N. P. LaeA, a Regulator of Secondary Metabolism in Aspergillus Spp. *Eukaryot Cell* **2004**, *3* (2), 527–535. https://doi.org/10.1128/EC.3.2.527-535.2004.

(145) Liu, Y.; Li, P.; Qi, C.; Zha, Z.; Meng, J.; Liu, C.; Han, J.; Zhou, Q.; Luo, Z.; Wang, J.; Zhu, H.; Ye, Y.; Chen, C.; Zhou, Y.; Zhang, Y. Cryptic Piperazine Derivatives Activated by Knocking out the Global Regulator LaeA in Aspergillus Flavipes. *Bioorg Med Chem* **2024**, *103*, 117685. https://doi.org/10.1016/j.bmc.2024.117685.

(146) Song, Z.; Zhou, S.; Zhang, H.; Keller, N. P.; Oakley, B. R.; Liu, X.; Yin, W.-B. Fungal Secondary Metabolism Is Governed by an RNA-Binding Protein CsdA/RsdA Complex. *Nat Commun* **2023**, *14* (1), 7351. https://doi.org/10.1038/s41467-023-43205-2.

(147) Keller, N. P. Fungal Secondary Metabolism: Regulation, Function and Drug Discovery. Nat Rev *Microbiol* **2019**, *17* (3), 167–180. https://doi.org/10.1038/s41579-018-0121-1.

(148) Bok, J. W.; Chiang, Y.-M.; Szewczyk, E.; Reyes-Dominguez, Y.; Davidson, A. D.; Sanchez, J. F.; Lo, H.-C.; Watanabe, K.; Strauss, J.; Oakley, B. R.; Wang, C. C. C.; Keller, N. P. Chromatin-Level Regulation of Biosynthetic Gene Clusters. *Nat Chem Biol* **2009**, *5* (7), 462–464. https://doi.org/10.1038/nchembio.177.

(149) Dallery, J.-F.; Le Goff, G.; Adelin, E.; Iorga, B. I.; Pigné, S.; O'Connell, R. J.; Ouazzani, J. Deleting a Chromatin Remodeling Gene Increases the Diversity of Secondary Metabolites Produced by Colletotrichum Higginsianum. *J Nat Prod* **2019**, *82* (4), 813–822. https://doi.org/10.1021/acs.jnatprod.8b00796.

(150) Mao, X.; Xu, W.; Li, D.; Yin, W.; Chooi, Y.; Li, Y.; Tang, Y.; Hu, Y. Epigenetic Genome Mining of an Endophytic Fungus Leads to the Pleiotropic Biosynthesis of Natural Products. *Angewandte Chemie International Edition* **2015**, *54* (26), 7592–7596. https://doi.org/10.1002/anie.201502452.

(151) Bai, J.; Mu, R.; Dou, M.; Yan, D.; Liu, B.; Wei, Q.; Wan, J.; Tang, Y.; Hu, Y. Epigenetic Modification in Histone Deacetylase Deletion Strain of Calcarisporium Arbuscula Leads to Diverse Diterpenoids. *Acta Pharm Sin B* **2018**, *8* (4), 687–697. https://doi.org/10.1016/j.apsb.2017.12.012.

(152) Lu, H.; Chen, H.-Y.; Luo, S.-R.; Guo, L.; Luo, D.; Ma, L.-F.; Zhan, Z.-J. A New Cleistanthane Diterpenoid from the Cultures of a Mutant of Calcarisporium Arbuscula NRRL 3705. *Phytochem Lett* **2024**, *60*, 239–242. https://doi.org/10.1016/j.phytol.2023.09.001.

(153) Fan, A.; Mi, W.; Liu, Z.; Zeng, G.; Zhang, P.; Hu, Y.; Fang, W.; Yin, W.-B. Deletion of a Histone Acetyltransferase Leads to the Pleiotropic Activation of Natural Products in Metarhizium Robertsii. *Org Lett* **2017**, *19* (7), 1686–1689. https://doi.org/10.1021/acs.orglett.7b00476.

(154) Guo, Z.; Zou, Z.-M. Discovery of New Secondary Metabolites by Epigenetic Regulation and NMR Comparison from the Plant Endophytic Fungus Monosporascus Eutypoides. *Molecules* **2020**, *25* (18), 4192. https://doi.org/10.3390/molecules25184192.

(155) Wu, G.; Zhou, H.; Zhang, P.; Wang, X.; Li, W.; Zhang, W.; Liu, X.; Liu, H.-W.; Keller, N. P.; An, Z.; Yin, W.-B. Polyketide Production of Pestaloficiols and Macrodiolide Ficiolides Revealed by Manipulations of Epigenetic Regulators in an Endophytic Fungus. *Org Lett* **2016**, *18* (8), 1832–1835. https://doi.org/10.1021/acs.orglett.6b00562.

(156) Ameruoso, A.; Villegas Kcam, M. C.; Cohen, K. P.; Chappell, J. Activating Natural Product Synthesis Using CRISPR Interference and Activation Systems in Streptomyces. *Nucleic Acids Res* **2022**, *50* (13), 7751–7760. https://doi.org/10.1093/nar/gkac556.

(157) Gilbert, L. A.; Horlbeck, M. A.; Adamson, B.; Villalta, J. E.; Chen, Y.; Whitehead, E. H.; Guimaraes, C.; Panning, B.; Ploegh, H. L.; Bassik, M. C.; Qi, L. S.; Kampmann, M.; Weissman, J. S. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **2014**, *159* (3), 647–661. https://doi.org/10.1016/j.cell.2014.09.029.

(158) Qi, L. S.; Larson, M. H.; Gilbert, L. A.; Doudna, J. A.; Weissman, J. S.; Arkin, A. P.; Lim, W. A. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **2013**, *152* (5), 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022.

(159) Adli, M. The CRISPR Tool Kit for Genome Editing and Beyond. *Nat Commun* **2018**, *9* (1), 1911. https://doi.org/10.1038/s41467-018-04252-2.

(160) Li, L.; Wei, K.; Zheng, G.; Liu, X.; Chen, S.; Jiang, W.; Lu, Y. CRISPR-Cpf1-Assisted Multiplex Genome Editing and Transcriptional Repression in Streptomyces. *Appl Environ Microbiol* **2018**, *84* (18), e00827-18. https://doi.org/10.1128/AEM.00827-18.

(161) Rill, A.; Zhao, L.; Bode, H. B. Genetic Toolbox for Photorhabdus and Xenorhabdus: PSEVA Based Heterologous Expression Systems and CRISPR/Cpf1 Based Genome Editing for Rapid Natural Product Profiling. *Microb Cell Fact* **2024**, *23* (1), 98. https://doi.org/10.1186/s12934-024-02363-8.

(162) Ke, J.; Robinson, D.; Wu, Z.-Y.; Kuftin, A.; Louie, K.; Kosina, S.; Northen, T.; Cheng, J.-F.; Yoshikuni, Y. CRAGE-CRISPR Facilitates Rapid Activation of Secondary Metabolite Biosynthetic Gene Clusters in Bacteria. *Cell Chem Biol* **2022**, *29* (4), 696-710.e4. https://doi.org/10.1016/j.chembiol.2021.08.009.

(163) Wang, G.; Zhao, Z.; Ke, J.; Engel, Y.; Shi, Y.-M.; Robinson, D.; Bingol, K.; Zhang, Z.; Bowen, B.; Louie, K.; Wang, B.; Evans, R.; Miyamoto, Y.; Cheng, K.; Kosina, S.; De Raad, M.; Silva, L.; Luhrs, A.; Lubbe, A.; Hoyt, D. W.; Francavilla, C.; Otani, H.; Deutsch, S.; Washton, N. M.; Rubin, E. M.; Mouncey, N. J.; Visel, A.; Northen, T.; Cheng, J.-F.; Bode, H. B.; Yoshikuni, Y. CRAGE Enables Rapid Activation of Biosynthetic Gene Clusters in Undomesticated Bacteria. *Nat Microbiol* **2019**, *4* (12), 2498–2510. https://doi.org/10.1038/s41564-019-0573-8.

(164) Zhang, M. M.; Wong, F. T.; Wang, Y.; Luo, S.; Lim, Y. H.; Heng, E.; Yeo, W. L.; Cobb, R. E.; Enghiad, B.; Ang, E. L.; Zhao, H. CRISPR–Cas9 Strategy for Activation of Silent Streptomyces Biosynthetic Gene Clusters. *Nat Chem Biol* **2017**, *13* (6), 607–609. https://doi.org/10.1038/nchembio.2341.

(165) Xie, F.; Zhao, H.; Liu, J.; Yang, X.; Neuber, M.; Agrawal, A. A.; Kaur, A.; Herrmann, J.; Kalinina, O. V.; Wei, X.; Müller, R.; Fu, C. Autologous DNA Mobilization and Multiplication Expedite Natural Products Discovery from Bacteria. *Science (1979)* **2024**, *386* (6727), eabq7333. https://doi.org/10.1126/science.abq7333.

(166) Shi, T.-Q.; Gao, J.; Wang, W.-J.; Wang, K.-F.; Xu, G.-Q.; Huang, H.; Ji, X.-J. CRISPR/Cas9-Based Genome Editing in the Filamentous Fungus Fusarium Fujikuroi and Its Application in Strain Engineering for Gibberellic Acid Production. *ACS Synth Biol* **2019**, *8* (2), 445–454. https://doi.org/10.1021/acssynbio.8b00478.

(167) Zheng, Y.-M.; Lin, F.-L.; Gao, H.; Zou, G.; Zhang, J.-W.; Wang, G.-Q.; Chen, G.-D.; Zhou, Z.-H.; Yao, X.-S.; Hu, D. Development of a Versatile and Conventional Technique for Gene Disruption in Filamentous Fungi Based on CRISPR-Cas9 Technology. *Sci Rep* **2017**, *7* (1), 9250. https://doi.org/10.1038/s41598-017-10052-3.

(168) Wang, Y.; Chen, H.; Ma, L.; Gong, M.; Wu, Y.; Bao, D.; Zou, G. Use of CRISPR-Cas Tools to Engineer Trichoderma Species. *Microb Biotechnol* **2022**, *15* (10), 2521–2532. https://doi.org/10.1111/1751-7915.14126.

(169) Liu, W.; An, C.; Shu, X.; Meng, X.; Yao, Y.; Zhang, J.; Chen, F.; Xiang, H.; Yang, S.; Gao, X.; Gao, S.-S. A Dual-Plasmid CRISPR/Cas System for Mycotoxin Elimination in Polykaryotic Industrial Fungi. *ACS Synth Biol* **2020**, *9* (8), 2087–2095. https://doi.org/10.1021/acssynbio.0c00178.

(170) Yuan, B.; Keller, N. P.; Oakley, B. R.; Stajich, J. E.; Wang, C. C. C. Manipulation of the Global Regulator McrA Upregulates Secondary Metabolite Production in Aspergillus Wentii Using CRISPR-Cas9 with In Vitro Assembled Ribonucleoproteins. *ACS Chem Biol* **2022**, *17* (10), 2828–2835. https://doi.org/10.1021/acschembio.2c00456.

(171) Roux, I.; Woodcraft, C.; Hu, J.; Wolters, R.; Gilchrist, C. L. M.; Chooi, Y.-H. CRISPR-Mediated Activation of Biosynthetic Gene Clusters for Bioactive Molecule Discovery in Filamentous Fungi. *ACS Synth Biol* **2020**, *9* (7), 1843–1854. https://doi.org/10.1021/acssynbio.0c00197.

(172) Mózsik, L.; Hoekzema, M.; de Kok, N. A. W.; Bovenberg, R. A. L.; Nygård, Y.; Driessen, A. J. M. CRISPR-Based Transcriptional Activation Tool for Silent Genes in Filamentous Fungi. *Sci Rep* **2021**, *11* (1), 1118. https://doi.org/10.1038/s41598-020-80864-3.

(173) Zhao, F.; Sun, C.; Liu, Z.; Cabrera, A.; Escobar, M.; Huang, S.; Yuan, Q.; Nie, Q.; Luo, K. L.; Lin, A.; Vanegas, J. A.; Zhu, T.; Hilton, I. B.; Gao, X. Multiplex Base-Editing Enables Combinatorial Epigenetic Regulation for Genome Mining of Fungal Natural Products. *J Am Chem Soc* **2023**, *145* (1), 413–421. https://doi.org/10.1021/jacs.2c10211.

(174) Wang, M.; Carver, J. J.; Phelan, V. V; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V; Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya P, C. A.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. Ø.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* **2016**, *34* (8), 828–837. https://doi.org/10.1038/nbt.3597.

(175) Petras, D.; Phelan, V. V.; Acharya, D.; Allen, A. E.; Aron, A. T.; Bandeira, N.; Bowen, B. P.; Belle-Oudry, D.; Boecker, S.; Cummings, D. A.; Deutsch, J. M.; Fahy, E.; Garg, N.; Gregor, R.; Handelsman, J.; Navarro-Hoyos, M.; Jarmusch, A. K.; Jarmusch, S. A.; Louie, K.; Maloney, K. N.; Marty, M. T.; Meijler, M. M.; Mizrahi, I.; Neve, R. L.; Northen, T. R.; Molina-Santiago, C.; Panitchpakdi, M.; Pullman, B.; Puri, A. W.; Schmid, R.; Subramaniam, S.; Thukral, M.; Vasquez-Castro, F.; Dorrestein, P. C.; Wang, M. GNPS Dashboard: Collaborative Exploration of Mass Spectrometry Data in the Web Browser. *Nat Methods* **2022**, *19* (2), 134–136. https://doi.org/10.1038/s41592-021-01339-5.

(176) Pang, Z.; Xu, L.; Viau, C.; Lu, Y.; Salavati, R.; Basu, N.; Xia, J. MetaboAnalystR 4.0: A Unified LC-MS Workflow for Global Metabolomics. *Nat Commun* **2024**, *15* (1), 3675. https://doi.org/10.1038/s41467-024-48009-6.

(177) Wang, X.-C.; Zhang, J.-N.; Zhao, J.-J.; Guo, X.-M.; Li, S.-F.; Zheng, Q.-X.; Liu, P.-P.; Lu, P.; Fu, H.-Y.; Yu, Y.-J.; She, Y. AntDAS-DDA: A New Platform for Data-Dependent Acquisition Mode-Based Untargeted Metabolomic Profiling Analysis with Advantage of Recognizing Insource Fragment Ions to Improve Compound Identification. *Anal Chem* **2023**, *95* (2), 638-649. https://doi.org/10.1021/acs.analchem.2c01795.

(178) Gaudêncio, S. P.; Bayram, E.; Lukić Bilela, L.; Cueto, M.; Díaz-Marrero, A. R.; Haznedaroglu, B. Z.; Jimenez, C.; Mandalakis, M.; Pereira, F.; Reyes, F.; Tasdemir, D. Advanced Methods for Natural Products Discovery: Bioactivity Screening, Dereplication, Metabolomics Profiling, Genomic Sequencing, Databases and Informatic Tools, and Structure Elucidation. *Mar Drugs* **2023**, *21* (5), 308. https://doi.org/10.3390/md21050308.

(179) Berger, T.; Alenfelder, J.; Steinmüller, S.; Heimann, D.; Gohain, N.; Petras, D.; Wang, M.; Berger, R.; Kostenis, E.; Reher, R. A MassQL-Integrated Molecular Networking Approach for the Discovery and Substructure Annotation of Bioactive Cyclic Peptides. *J Nat Prod* **2024**, *87* (4), 692–704. https://doi.org/10.1021/acs.jnatprod.3c00750.

(180) Selegato, D. M.; Zanatta, A. C.; Pilon, A. C.; Veloso, J. H.; Castro-Gamboa, I. Application of Feature-Based Molecular Networking and MassQL for the MS/MS Fragmentation Study of Depsipeptides. *Front Mol Biosci* **2023**, *10*, 1238475. https://doi.org/10.3389/fmolb.2023.1238475.

(181) Wang, M.; Jarmusch, A. K.; Vargas, F.; Aksenov, A. A.; Gauglitz, J. M.; Weldon, K.; Petras, D.; da Silva, R.; Quinn, R.; Melnik, A. V.; van der Hooft, J. J. J.; Caraballo-Rodríguez, A. M.; Nothias, L. F.; Aceves, C. M.; Panitchpakdi, M.; Brown, E.; Di Ottavio, F.; Sikora, N.; Elijah, E. O.; Labarta-Bajo, L.; Gentry, E. C.; Shalapour, S.; Kyle, K. E.; Puckett, S. P.; Watrous, J. D.; Carpenter, C. S.; Bouslimani, A.; Ernst, M.; Swafford, A. D.; Zúñiga, E. I.; Balunas, M. J.; Klassen, J. L.; Loomba, R.; Knight, R.; Bandeira, N.; Dorrestein, P. C. Mass Spectrometry Searches Using MASST. *Nat Biotechnol* **2020**, *38* (1), 23–26. https://doi.org/10.1038/s41587-019-0375-9.

(182) Mongia, M.; Yasaka, T. M.; Liu, Y.; Guler, M.; Lu, L.; Bhagwat, A.; Behsaz, B.; Wang, M.; Dorrestein, P. C.; Mohimani, H. Fast Mass Spectrometry Search and Clustering of Untargeted Metabolomics Data. *Nat Biotechnol* **2024**. https://doi.org/10.1038/s41587-023-01985-4.

(183) Zuffa, S.; Schmid, R.; Bauermeister, A.; P. Gomes, P. W.; Caraballo-Rodriguez, A. M.; El Abiead, Y.; Aron, A. T.; Gentry, E. C.; Zemlin, J.; Meehan, M. J.; Avalon, N. E.; Cichewicz, R. H.; Buzun, E.; Terrazas, M. C.; Hsu, C.-Y.; Oles, R.; Ayala, A. V.; Zhao, J.; Chu, H.; Kuijpers, M. C. M.; Jackrel, S. L.; Tugizimana, F.; Nephali, L. P.; Dubery, I. A.; Madala, N. E.; Moreira, E. A.; Costa-Lotufo, L. V.; Lopes, N. P.; Rezende-Teixeira, P.; Jimenez, P. C.; Rimal, B.; Patterson, A. D.; Traxler, M. F.; Pessotti, R. de C.; Alvarado-Villalobos, D.; Tamayo-Castillo, G.; Chaverri, P.; Escudero-Leyva, E.; Quiros-Guerrero, L.-M.; Bory, A. J.; Joubert, J.; Rutz, A.; Wolfender, J.-L.; Allard, P.-M.; Sichert, A.; Pontrelli, S.; Pullman, B. S.; Bandeira, N.; Gerwick, W. H.; Gindro, K.; Massana-Codina, J.; Wagner, B. C.; Forchhammer, K.; Petras, D.; Aiosa, N.; Garg, N.; Liebeke, M.; Bourceau, P.; Kang, K. Bin; Gadhavi, H.; de Carvalho, L. P. S.; Silva dos Santos, M.; Pérez-Lorente, A. I.; Molina-Santiago, C.; Romero, D.; Franke, R.; Brönstrup, M.; Vera Ponce de León, A.; Pope, P. B.; La Rosa, S. L.; La Barbera, G.; Roager, H. M.; Laursen, M. F.; Hammerle, F.; Siewert, B.; Peintner, U.; Licona-Cassani, C.; Rodriguez-Orduña, L.; Rampler, E.; Hildebrand, F.; Koellensperger, G.; Schoeny, H.; Hohenwallner, K.; Panzenboeck, L.; Gregor, R.; O'Neill, E. C.; Roxborough, E. T.; Odoi, J.; Bale, N. J.; Ding, S.; Sinninghe Damsté, J. S.; Guan, X. L.; Cui, J. J.; Ju, K.-S.; Silva, D. B.; Silva, F. M. R.; da Silva, G. F.; Koolen, H. H. F.; Grundmann, C.; Clement, J. A.; Mohimani, H.; Broders, K.; McPhail, K. L.; Ober-Singleton, S. E.; Rath, C. M.; McDonald, D.; Knight, R.; Wang, M.; Dorrestein, P. C. MicrobeMASST: A Taxonomically Informed Mass Spectrometry Search Tool for Microbial Metabolomics Data. *Nat Microbiol* **2024**, *9* (2), 336–345. https://doi.org/10.1038/s41564-023-01575-9.

(184) West, K. A.; Schmid, R.; Gauglitz, J. M.; Wang, M.; Dorrestein, P. C. FoodMASST a Mass Spectrometry Search Tool for Foods and Beverages. *NPJ Sci Food* **2022**, *6* (1), 22. https://doi.org/10.1038/s41538-022-00137-3.

(185) de Jonge, N. F.; Louwen, J. J. R.; Chekmeneva, E.; Camuzeaux, S.; Vermeir, F. J.; Jansen, R. S.; Huber, F.; van der Hooft, J. J. J. MS2Query: Reliable and Scalable MS2 Mass Spectra-Based Analogue Search. *Nat Commun* **2023**, *14* (1), 1752. https://doi.org/10.1038/s41467-023-37446-4.

(186) Huber, F.; Ridder, L.; Verhoeven, S.; Spaaks, J. H.; Diblen, F.; Rogers, S.; van der Hooft, J. J. J. Spec2Vec: Improved Mass Spectral Similarity Scoring through Learning of Structural Relationships. *PLoS Comput Biol* **2021**, *17* (2), e1008724. https://doi.org/10.1371/journal.pcbi.1008724.

(187) Huber, F.; van der Burg, S.; van der Hooft, J. J. J.; Ridder, L. MS2DeepScore: A Novel Deep Learning Similarity Measure to Compare Tandem Mass Spectra. *J Cheminform* **2021**, *13* (1), 84. https://doi.org/10.1186/s13321-021-00558-4.

(188) Godzien, J.; Gil de la Fuente, A.; Otero, A.; Barbas, C. Metabolite Annotation and Identification; **2018**; pp 415–445. https://doi.org/10.1016/bs.coac.2018.07.004.

(189) Shen, X.; Wu, S.; Liang, L.; Chen, S.; Contrepois, K.; Zhu, Z.-J.; Snyder, M. MetID: An R Package for Automatable Compound Annotation for LC-MS-Based Data. *Bioinformatics* **2022**, *38* (2), 568–569. https://doi.org/10.1093/bioinformatics/btab583.

(190) Del Carratore, F.; Eagles, W.; Borka, J.; Breitling, R. IpaPy2: Integrated Probabilistic Annotation (IPA) 2.0—an Improved Bayesian-Based Method for the Annotation of LC–MS/MS Untargeted Metabolomics Data. *Bioinformatics* **2023**, *39* (7), btad455. https://doi.org/10.1093/bioinformatics/btad455.

(191) Graça, G.; Cai, Y.; Lau, C.-H. E.; Vorkas, P. A.; Lewis, M. R.; Want, E. J.; Herrington, D.; Ebbels, T. M. D. Automated Annotation of Untargeted All-Ion Fragmentation LC–MS Metabolomics Data with MetaboAnnotatoR. *Anal Chem* **2022**, *94* (8), 3446–3455. https://doi.org/10.1021/acs.analchem.1c03032.

(192) Zhou, Z.; Luo, M.; Zhang, H.; Yin, Y.; Cai, Y.; Zhu, Z.-J. Metabolite Annotation from Knowns to Unknowns through Knowledge-Guided Multi-Layer Metabolic Networking. *Nat Commun* **2022**, *13* (1), 6656. https://doi.org/10.1038/s41467-022-34537-6.

(193) Zheng, F.; You, L.; Qin, W.; Ouyang, R.; Lv, W.; Guo, L.; Lu, X.; Li, E.; Zhao, X.; Xu, G. MetEx: A Targeted Extraction Strategy for Improving the Coverage and Accuracy of Metabolite Annotation in Liquid Chromatography–High-Resolution Mass Spectrometry Data. *Anal Chem* **2022**, *94* (24), 8561–8569. https://doi.org/10.1021/acs.analchem.1c04783.

(194) Strutz, J.; Shebek, K. M.; Broadbelt, L. J.; Tyo, K. E. J. MINE 2.0: Enhanced Biochemical Coverage for Peak Identification in Untargeted Metabolomics. *Bioinformatics* **2022**, *38* (13), 3484–3487. https://doi.org/10.1093/bioinformatics/btac331.

(195) Dührkop, K.; Shen, H.; Meusel, M.; Rousu, J.; Böcker, S. Searching Molecular Structure Databases with Tandem Mass Spectra Using CSI:FingerID. *Proceedings of the National Academy of Sciences* **2015**, *112* (41), 12580–12585. https://doi.org/10.1073/pnas.1509788112.

(196) Hoffmann, M. A.; Nothias, L.-F.; Ludwig, M.; Fleischauer, M.; Gentry, E. C.; Witting, M.; Dorrestein, P. C.; Dührkop, K.; Böcker, S. High-Confidence Structural Annotation of Metabolites Absent from Spectral Libraries. *Nat Biotechnol* **2022**, *40* (3), 411–421. https://doi.org/10.1038/s41587-021-01045-9.

(197) Baygi, S. F.; Kumar, Y.; Barupal, D. K. IDSL.CSA: Composite Spectra Analysis for Chemical Annotation of Untargeted Metabolomics Datasets. *Anal Chem* **2023**, *95* (25), 9480–9487. https://doi.org/10.1021/acs.analchem.3c00376.

(198) Aron, A. T.; Gentry, E. C.; McPhail, K. L.; Nothias, L.-F.; Nothias-Esposito, M.; Bouslimani, A.; Petras, D.; Gauglitz, J. M.; Sikora, N.; Vargas, F.; van der Hooft, J. J. J.; Ernst, M.; Kang, K. Bin; Aceves, C. M.; Caraballo-Rodríguez, A. M.; Koester, I.; Weldon, K. C.; Bertrand, S.; Roullier, C.; Sun, K.; Tehan, R. M.; Boya P., C. A.; Christian, M. H.; Gutiérrez, M.; Ulloa, A. M.; Tejeda Mora, J. A.; Mojica-Flores, R.; Lakey-Beitia, J.; Vásquez-Chaves, V.; Zhang, Y.; Calderón, A. I.; Tayler, N.; Keyzers, R. A.; Tugizimana, F.; Ndlovu, N.; Aksenov, A. A.; Jarmusch, A. K.; Schmid, R.; Truman, A. W.; Bandeira, N.; Wang, M.; Dorrestein, P. C. Reproducible Molecular Networking of Untargeted Mass Spectrometry Data Using GNPS. *Nat Protoc* **2020**, *15* (6), 1954–1991. https://doi.org/10.1038/s41596-020-0317-5.

(199) Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; Aicheler, F.; Aksenov, A. A.; Alka, O.; Allard, P.-M.; Barsch, A.; Cachet, X.; Caraballo-Rodriguez, A. M.; Da Silva, R. R.; Dang, T.; Garg, N.; Gauglitz, J. M.; Gurevich, A.; Isaac, G.; Jarmusch, A. K.; Kameník, Z.; Kang, K. Bin; Kessler, N.; Koester, I.; Korf, A.; Le Gouellec, A.; Ludwig, M.; Martin H., C.; McCall, L.-I.; McSayles, J.; Meyer, S. W.; Mohimani, H.; Morsy, M.; Moyne, O.; Neumann, S.; Neuweger, H.; Nguyen, N. H.; Nothias-Esposito, M.; Paolini, J.; Phelan, V. V.; Pluskal, T.; Quinn, R. A.; Rogers, S.; Shrestha, B.; Tripathi, A.; van der Hooft, J. J. J.; Vargas, F.; Weldon, K. C.; Witting, M.; Yang, H.; Zhang, Z.; Zubeil, F.; Kohlbacher, O.; Böcker, S.; Alexandrov, T.; Bandeira, N.; Wang, M.; Dorrestein, P. C. Feature-Based Molecular Networking in the GNPS Analysis Environment. Nat Methods **2020**, 17 (9), 905–908. https://doi.org/10.1038/s41592-020-0933-6.

(200) Morehouse, N. J.; Clark, T. N.; McMann, E. J.; van Santen, J. A.; Haeckl, F. P. J.; Gray, C. A.; Linington, R. G. Annotation of Natural Product Compound Families Using Molecular Networking Topology and Structural Similarity Fingerprinting. *Nat Commun* **2023**, *14* (1), 308. https://doi.org/10.1038/s41467-022-35734-z.

(201) Wang, X.; Li, C.; Li, Z.; Qi, Y.; Zhang, X.; Zhao, X.; Zhao, C.; Lin, X.; Lu, X.; Xu, G. A Structure-Guided Molecular Network Strategy for Global Untargeted Metabolomics Data Annotation. *Anal Chem* **2023**, *95* (31), 11603–11612. https://doi.org/10.1021/acs.analchem.3c00849.

(202) Baygi, S. F.; Barupal, D. K. IDSL_MINT: A Deep Learning Framework to Predict Molecular Fingerprints from Mass Spectra. *J Cheminform* **2024**, *16* (1), 8. https://doi.org/10.1186/s13321-024-00804-5.

(203) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. *Nat Methods* **2019**, *16* (4), 299–302. https://doi.org/10.1038/s41592-019-0344-8.

(204) Stravs, M. A.; Dührkop, K.; Böcker, S.; Zamboni, N. MSNovelist: De Novo Structure Generation from Mass Spectra. *Nat Methods* **2022**, *19* (7), 865–870. https://doi.org/10.1038/s41592-022-01486-3.

(205) Hong, Y.; Li, S.; Welch, C. J.; Tichy, S.; Ye, Y.; Tang, H. 3DMolMS: Prediction of Tandem Mass Spectra from 3D Molecular Conformations. *Bioinformatics* **2023**, *39* (6), btad354. https://doi.org/10.1093/bioinformatics/btad354.

(206) Lee, Y.-Y.; Guler, M.; Chigumba, D. N.; Wang, S.; Mittal, N.; Miller, C.; Krummenacher, B.; Liu, H.; Cao, L.; Kannan, A.; Narayan, K.; Slocum, S. T.; Roth, B. L.; Gurevich, A.; Behsaz, B.; Kersten, R. D.; Mohimani, H. HypoRiPPAtlas as an Atlas of Hypothetical Natural Products for Mass Spectrometry Database Search. *Nat Commun* **2023**, *14* (1), 4219. https://doi.org/10.1038/s41467-023-39905-4.

(207) Cardoso, S.; Cabral, D.; Maraschin, M.; Rocha, M. NMRFinder: A Novel Method for 1D 1H-NMR Metabolite Annotation. *Metabolomics* **2021**, *17* (2), 21. https://doi.org/10.1007/s11306-021-01772-9.

(208) Bruguière, A.; Derbré, S.; Dietsch, J.; Leguy, J.; Rahier, V.; Pottier, Q.; Bréard, D.; Suor-Cherer, S.; Viault, G.; Le Ray, A.-M.; Saubion, F.; Richomme, P. MixONat, a Software for the Dereplication of Mixtures Based on 13 C NMR Spectroscopy. *Anal Chem* **2020**, *92* (13), 8793–8801. https://doi.org/10.1021/acs.analchem.0c00193.

(209) Martínez-Treviño, S. H.; Uc-Cetina, V.; Fernández-Herrera, M. A.; Merino, G. Prediction of Natural Product Classes Using Machine Learning and 13 C NMR Spectroscopic Data. *J Chem Inf Model* **2020**, *60* (7), 3376–3386. https://doi.org/10.1021/acs.jcim.0c00293.

(210) Yao, L.; Yang, M.; Song, J.; Yang, Z.; Sun, H.; Shi, H.; Liu, X.; Ji, X.; Deng, Y.; Wang, X. Conditional Molecular Generation Net Enables Automated Structure Elucidation Based on 13 C NMR Spectra and Prior Knowledge. *Anal Chem* **2023**, *95* (12), 5393–5401. https://doi.org/10.1021/acs.analchem.2c05817.

(211) Zhang, C.; Idelbayev, Y.; Roberts, N.; Tao, Y.; Nannapaneni, Y.; Duggan, B. M.; Min, J.; Lin, E. C.; Gerwick, E. C.; Cottrell, G. W.; Gerwick, W. H. Small Molecule Accurate Recognition Technology (SMART) to Enhance Natural Products Research. *Sci Rep* **2017**, *7* (1), 14243. https://doi.org/10.1038/s41598-017-13923-x.

(212) Reher, R.; Kim, H. W.; Zhang, C.; Mao, H. H.; Wang, M.; Nothias, L.-F.; Caraballo-Rodriguez, A. M.; Glukhov, E.; Teke, B.; Leao, T.; Alexander, K. L.; Duggan, B. M.; Van Everbroeck, E. L.; Dorrestein, P. C.; Cottrell, G. W.; Gerwick, W. H. A Convolutional Neural Network-Based Approach for the Rapid Annotation of Molecularly Diverse Natural Products. *J Am Chem Soc* **2020**, *142* (9), 4114–4120. https://doi.org/10.1021/jacs.9b13786.

(213) Kim, H. W.; Zhang, C.; Reher, R.; Wang, M.; Alexander, K. L.; Nothias, L.-F.; Han, Y. K.; Shin, H.; Lee, K. Y.; Lee, K. H.; Kim, M. J.; Dorrestein, P. C.; Gerwick, W. H.; Cottrell, G. W. DeepSAT: Learning Molecular Structures from Nuclear Magnetic Resonance Data. *J Cheminform* **2023**, *15* (1), 71. https://doi.org/10.1186/s13321-023-00738-4.

(214) Kim, H. W.; Zhang, C.; Cottrell, G. W.; Gerwick, W. H. SMART-Miner: A Convolutional Neural Network-based Metabolite Identification from 1 H- 13 C HSQC Spectra. *Magnetic Resonance in Chemistry* **2022**, *60* (11), 1070–1075. https://doi.org/10.1002/mrc.5240.

(215) Wei, W.; Liao, Y.; Wang, Y.; Wang, S.; Du, W.; Lu, H.; Kong, B.; Yang, H.; Zhang, Z. Deep Learning-Based Method for Compound Identification in NMR Spectra of Mixtures. *Molecules* **2022**, *27* (12), 3653. https://doi.org/10.3390/molecules27123653.

(216) Takis, P. G.; Aggelidou, V. A.; Sands, C. J.; Louka, A. Mapping of 1 H NMR Chemical Shifts Relationship with Chemical Similarities for the Acceleration of Metabolic Profiling: Application on Blood Products. *Magnetic Resonance in Chemistry* **2023**, *61* (12), 759–769. https://doi.org/10.1002/mrc.5392.

(217) Du, H.; Gu, X.; Chen, J.; Bai, C.; Duan, X.; Hu, K. GIPMA: Global Intensity-Guided Peak Matching and Alignment for 2D 1H–13C HSQC-Based Metabolomics. *Anal Chem* **2023**, *95* (6), 3195-3203. https://doi.org/10.1021/acs.analchem.2c03323.

(218) Guan, Y.; Shree Sowndarya, S. V.; Gallegos, L. C.; St. John, P. C.; Paton, R. S. Real-Time Prediction of 1 H and 13 C Chemical Shifts with DFT Accuracy Using a 3D Graph Neural Network. *Chem Sci* **2021**, *12* (36), 12012–12026. https://doi.org/10.1039/D1SC03343C.

(219) Cloarec, O.; Dumas, M.-E.; Craig, A.; Barton, R. H.; Trygg, J.; Hudson, J.; Blancher, C.; Gauguier, D.; Lindon, J. C.; Holmes, E.; Nicholson, J. Statistical Total Correlation Spectroscopy: An Exploratory Approach for Latent Biomarker Identification from Metabolic 1 H NMR Data Sets. *Anal Chem* **2005**, *77* (5), 1282–1289. https://doi.org/10.1021/ac048630x.

(220) Crockford, D. J.; Holmes, E.; Lindon, J. C.; Plumb, R. S.; Zirah, S.; Bruce, S. J.; Rainville, P.; Stumpf, C. L.; Nicholson, J. K. Statistical Heterospectroscopy, an Approach to the Integrated Analysis of NMR and UPLC-MS Data Sets: Application in Metabonomic Toxicology Studies. *Anal Chem* **2006**, *78* (2), 363–371. https://doi.org/10.1021/ac051444m.

(221) Watermann, S.; Bode, M.-C.; Hackl, T. Identification of Metabolites from Complex Mixtures by 3D Correlation of 1H NMR, MS and LC Data Using the SCORE-Metabolite-ID Approach. *Sci Rep* **2023**, *13* (1), 15834. https://doi.org/10.1038/s41598-023-43056-3.

(222) Watermann, S.; Bode, M.-C.; Hackl, T. Identification of Metabolites from Complex Mixtures by 3D Correlation of 1H NMR, MS and LC Data Using the SCORE-Metabolite-ID Approach. *Sci Rep* **2023**, *13* (1), 15834. https://doi.org/10.1038/s41598-023-43056-3.

(223) Burg, R. W.; Miller, B. M.; Baker, E. E.; Birnbaum, J.; Currie, S. A.; Hartman, R.; Kong, Y.-L.; Monaghan, R. L.; Olson, G.; Putter, I.; Tunac, J. B.; Wallick, H.; Stapley, E. O.; Oiwa, R.; Ōmura, S. Avermectins, New Family of Potent Anthelmintic Agents: Producing Organism and Fermentation. *Antimicrob Agents Chemother* **1979**, *15* (3), 361–367. https://doi.org/10.1128/AAC.15.3.361.

(224) Shen, B. A New Golden Age of Natural Products Drug Discovery. *Cell* **2015**, *163* (6), 1297–1300. https://doi.org/10.1016/j.cell.2015.11.031.

(225) Hewage, R. T.; Aree, T.; Mahidol, C.; Ruchirawat, S.; Kittakoop, P. One Strain-Many Compounds (OSMAC) Method for Production of Polyketides, Azaphilones, and an Isochromanone Using the Endophytic Fungus Dothideomycete Sp. *Phytochemistry* **2014**, *108*, 87–94. https://doi.org/10.1016/j.phytochem.2014.09.013.

(226) Romano, S.; Jackson, S. A.; Patry, S.; Dobson, A. D. W. Extending the "One Strain Many Compounds" (OSMAC) Principle to Marine Microorganisms. *Mar Drugs* **2018**, *16* (7), 244. https://doi.org/10.3390/md16070244.

(227) Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C. Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of *E. Coli*. *Science (1979)* **2001**, *291* (5509), 1790–1792. https://doi.org/10.1126/science.1058092.

(228) Ajikumar, P. K.; Xiao, W.-H.; Tyo, K. E. J.; Wang, Y.; Simeon, F.; Leonard, E.; Mucha, O.; Phon, T. H.; Pfeifer, B.; Stephanopoulos, G. Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in *Escherichia Coli*. *Science (1979)* **2010**, *330* (6000), 70–74. https://doi.org/10.1126/science.1191652.

(229) Nett, R. S.; Lau, W.; Sattely, E. S. Discovery and Engineering of Colchicine Alkaloid Biosynthesis. *Nature* **2020**, *584* (7819), 148–153. https://doi.org/10.1038/s41586-020-2546-8.

(230) Lau, W.; Sattely, E. S. Six Enzymes from Mayapple That Complete the Biosynthetic Pathway to the Etoposide Aglycone. *Science (1979)* **2015**, *349* (6253), 1224–1228. https://doi.org/10.1126/science.aac7202.

(231) Li, Y.; Li, S.; Thodey, K.; Trenchard, I.; Cravens, A.; Smolke, C. D. Complete Biosynthesis of Noscapine and Halogenated Alkaloids in Yeast. *Proceedings of the National Academy of Sciences* **2018**, *115* (17), E3922–E3931. https://doi.org/10.1073/pnas.1721469115.

(232) Ro, D.-K.; Paradise, E. M.; Ouellet, M.; Fisher, K. J.; Newman, K. L.; Ndungu, J. M.; Ho, K. A.; Eachus, R. A.; Ham, T. S.; Kirby, J.; Chang, M. C. Y.; Withers, S. T.; Shiba, Y.; Sarpong, R.; Keasling, J. D. Production of the Antimalarial Drug Precursor Artemisinic Acid in Engineered Yeast. *Nature* **2006**, *440* (7086), 940–943. https://doi.org/10.1038/nature04640.

(233) Luo, X.; Reiter, M. A.; d'Espaux, L.; Wong, J.; Denby, C. M.; Lechner, A.; Zhang, Y.; Grzybowski, A. T.; Harth, S.; Lin, W.; Lee, H.; Yu, C.; Shin, J.; Deng, K.; Benites, V. T.; Wang, G.; Baidoo, E. E. K.; Chen, Y.; Dev, I.; Petzold, C. J.; Keasling, J. D. Complete Biosynthesis of Cannabinoids and Their Unnatural Analogues in Yeast. *Nature* **2019**, *567* (7746), 123–126. https://doi.org/10.1038/s41586-019-0978-9.

(234) Bond, C. M.; Tang, Y. Engineering Saccharomyces Cerevisiae for Production of Simvastatin. *Metab Eng* **2019**, *51*, 1–8. https://doi.org/10.1016/j.ymben.2018.09.005.

(235) Ma, S. M.; Li, J. W.-H.; Choi, J. W.; Zhou, H.; Lee, K. K. M.; Moorthie, V. A.; Xie, X.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C.; Tang, Y. Complete Reconstitution of a Highly Reducing Iterative Polyketide Synthase. *Science (1979)* **2009**, *326* (5952), 589–592. https://doi.org/10.1126/science.1175602.

(236) Galanie, S.; Thodey, K.; Trenchard, I. J.; Filsinger Interrante, M.; Smolke, C. D. Complete Biosynthesis of Opioids in Yeast. *Science (1979)* **2015**, *349* (6252), 1095–1100. https://doi.org/10.1126/science.aac9373.

(237) Bailey, A. M.; Alberti, F.; Kilaru, S.; Collins, C. M.; de Mattos-Shipley, K.; Hartley, A. J.; Hayes, P.; Griffin, A.; Lazarus, C. M.; Cox, R. J.; Willis, C. L.; O'Dwyer, K.; Spence, D. W.; Foster, G. D. Identification and Manipulation of the Pleuromutilin Gene Cluster from Clitopilus Passeckerianus for Increased Rapid Antibiotic Production. *Sci Rep* **2016**, *6* (1), 25202. https://doi.org/10.1038/srep25202.

(238) Smith, D. J.; Burnham, M. K. R.; Edwards, J.; Earl, A. J.; Turner, G. Cloning and Heterologous Expression of the Penicillin Biosynthetic Gene Cluster from Penicillium Chrysogenum. *Nat Biotechnol* **1990**, *8* (1), 39–41. https://doi.org/10.1038/nbt0190-39.

(239) Hoefgen, S.; Lin, J.; Fricke, J.; Stroe, M. C.; Mattern, D. J.; Kufs, J. E.; Hortschansky, P.; Brakhage, A. A.; Hoffmeister, D.; Valiante, V. Facile Assembly and Fluorescence-Based Screening Method for Heterologous Expression of Biosynthetic Pathways in Fungi. *Metab Eng* **2018**, *48*, 44–51. https://doi.org/10.1016/j.ymben.2018.05.014.

(240) Volk, M. J.; Tran, V. G.; Tan, S.-I.; Mishra, S.; Fatma, Z.; Boob, A.; Li, H.; Xue, P.; Martin, T. A.; Zhao, H. Metabolic Engineering: Methodologies and Applications. *Chem Rev* **2023**, *123* (9), 5521–5570. https://doi.org/10.1021/acs.chemrev.2c00403.

(241) Gao, J.; Zuo, Y.; Xiao, F.; Wang, Y.; Li, D.; Xu, J.; Ye, C.; Feng, L.; Jiang, L.; Liu, T.; Gao, D.; Ma, B.; Huang, L.; Xu, Z.; Lian, J. Biosynthesis of Catharanthine in Engineered Pichia Pastoris. *Nature Synthesis* **2023**, *2* (3), 231–242. https://doi.org/10.1038/s44160-022-00205-2.

(242) Lambalot, R. H.; Gehring, A. M.; Flugel, R. S.; Zuber, P.; LaCelle, M.; Marahiel, M. A.; Reid, R.; Khosla, C.; Walsh, C. T. A New Enzyme Superfamily — the Phosphopantetheinyl Transferases. *Chem Biol* **1996**, *3* (11), 923–936. https://doi.org/10.1016/S1074-5521(96)90181-7.

(243) Zhang, H.; Fang, L.; Osburne, M. S.; Pfeifer, B. A. The Continuing Development of E. Coli as a Heterologous Host for Complex Natural Product Biosynthesis. *Methods Mol Biol* **2016**, *1401*, 121. https://doi.org/10.1007/978-1-4939-3375-4_8.

(244) Meyer, A. J.; Segall-Shapiro, T. H.; Glassey, E.; Zhang, J.; Voigt, C. A. Escherichia Coli “Marionette” Strains with 12 Highly Optimized Small-Molecule Sensors. *Nat Chem Biol* **2019**, *15* (2), 196–204. https://doi.org/10.1038/s41589-018-0168-3.

(245) Chang, M. C. Y.; Eachus, R. A.; Trieu, W.; Ro, D.-K.; Keasling, J. D. Engineering Escherichia Coli for Production of Functionalized Terpenoids Using Plant P450s. *Nat Chem Biol* **2007**, *3* (5), 274–277. https://doi.org/10.1038/nchembio875.

(246) Ichinose, H.; Wariishi, H. Heterologous Expression and Mechanistic Investigation of a Fungal Cytochrome P450 (CYP5150A2): Involvement of Alternative Redox Partners. *Arch Biochem Biophys* **2012**, *518* (1), 8–15. https://doi.org/10.1016/j.abb.2011.12.010.

(247) Zhang, H.; Wang, Y.; Wu, J.; Skalina, K.; Pfeifer, B. A. Complete Biosynthesis of Erythromycin A and Designed Analogs Using E. Coli as a Heterologous Host. *Chem Biol* **2010**, *17* (11), 1232–1240. https://doi.org/10.1016/j.chembiol.2010.09.013.

(248) Balskus, E. P.; Walsh, C. T. The Genetic and Molecular Basis for Sunscreen Biosynthesis in Cyanobacteria. *Science (1979)* **2010**, *329* (5999), 1653–1656. https://doi.org/10.1126/science.1193637.

(249) Hwang, S.; Lee, Y.; Kim, J. H.; Kim, G.; Kim, H.; Kim, W.; Cho, S.; Palsson, B. O.; Cho, B.-K. Streptomyces as Microbial Chassis for Heterologous Protein Expression. *Front Bioeng Biotechnol* **2021**, *9*, 804295. https://doi.org/10.3389/fbioe.2021.804295.

(250) Liu, R.; Deng, Z.; Liu, T. Streptomyces Species: Ideal Chassis for Natural Product Discovery and Overproduction. *Metab Eng* **2018**, *50*, 74–84. https://doi.org/10.1016/j.ymben.2018.05.015.

(251) Lee, N.; Hwang, S.; Lee, Y.; Cho, S.; Palsson, B.; Cho, B.-K. Synthetic Biology Tools for Novel Secondary Metabolite Discovery in Streptomyces. *J Microbiol Biotechnol* **2019**, *29* (5), 667–686. https://doi.org/10.4014/jmb.1904.04015.

(252) Luo, Y.; Zhang, L.; Barton, K. W.; Zhao, H. Systematic Identification of a Panel of Strong Constitutive Promoters from *Streptomyces Albus*. *ACS Synth Biol* **2015**, *4* (9), 1001–1010. https://doi.org/10.1021/acssynbio.5b00016.

(253) Liu, X.; Liu, D.; Xu, M.; Tao, M.; Bai, L.; Deng, Z.; Pfeifer, B. A.; Jiang, M. Reconstitution of Kinamycin Biosynthesis within the Heterologous Host *Streptomyces Albus* J1074. *J Nat Prod* **2018**, *81* (1), 72–77. https://doi.org/10.1021/acs.jnatprod.7b00652.

(254) Peng, Q.; Gao, G.; Lü, J.; Long, Q.; Chen, X.; Zhang, F.; Xu, M.; Liu, K.; Wang, Y.; Deng, Z.; Li, Z.; Tao, M. Engineered Streptomyces Lividans Strains for Optimal Identification and Expression of Cryptic Biosynthetic Gene Clusters. *Front Microbiol* **2018**, *9* (DEC), 429498. https://doi.org/10.3389/fmicb.2018.03042.

(255) Thanapipatsiri, A.; Claesen, J.; Gomez-Escribano, J.-P.; Bibb, M.; Thamchaipenet, A. A Streptomyces Coelicolor Host for the Heterologous Expression of Type III Polyketide Synthase Genes. *Microb Cell Fact* **2015**, *14* (1), 145. https://doi.org/10.1186/s12934-015-0335-0.

(256) McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. Engineered Biosynthesis of Novel Polyketides. *Science (1979)* **1993**, *262* (5139), 1546–1550. https://doi.org/10.1126/science.8248802.

(257) Binnie, C.; Warren, M.; Butler, M. J. Cloning and Heterologous Expression in Streptomyces Lividans of Streptomyces Rimosus Genes Involved in Oxytetracycline Biosynthesis. *J Bacteriol* **1989**, *171* (2), 887–895. https://doi.org/10.1128/jb.171.2.887-895.1989.

(258) Smanski, M. J.; Casper, J.; Peterson, R. M.; Yu, Z.; Rajski, S. R.; Shen, B. Expression of the Platencin Biosynthetic Gene Cluster in Heterologous Hosts Yielding New Platencin Congeners. *J Nat Prod* **2012**, *75* (12), 2158–2167. https://doi.org/10.1021/np3005985.

(259) Shi, J.; Xu, X.; Liu, P. Y.; Hu, Y. L.; Zhang, B.; Jiao, R. H.; Bashiri, G.; Tan, R. X.; Ge, H. M. Discovery and Biosynthesis of Guanipiperazine from a NRPS-like Pathway. *Chem Sci* **2021**, *12* (8), 2925–2930. https://doi.org/10.1039/D0SC06135B.

(260) Takao, R.; Sakai, K.; Koshino, H.; Osada, H.; Takahashi, S. Identification of the Kinanthraquinone Biosynthetic Gene Cluster by Expression of an Atypical Response Regulator. *Biosci Biotechnol Biochem* **2021**, *85* (3), 714–721. https://doi.org/10.1093/bbb/zbaa082.

(261) Hashimoto, T.; Hashimoto, J.; Kagaya, N.; Nishimura, T.; Suenaga, H.; Nishiyama, M.; Kuzuyama, T.; Shin-ya, K. A Novel Oxazole-Containing Tetraene Compound, JBIR-159, Produced by Heterologous Expression of the Cryptic Trans-AT Type Polyketide Synthase Biosynthetic Gene Cluster. *J Antibiot (Tokyo)* **2021**, *74* (5), 354–358. https://doi.org/10.1038/s41429-021-00410-9.

(262) Xiao, W.; Tsunoda, T.; Maruyama, C.; Hamano, Y.; Ogasawara, Y.; Dairi, T. Peptide Epimerase-Dehydratase Complex Responsible for Biosynthesis of the Linaridin Class Ribosomal Peptides. *Biosci Biotechnol Biochem* **2023**, *87* (11), 1316–1322. https://doi.org/10.1093/bbb/zbad106.

(263) Kozakai, R.; Ono, T.; Hoshino, S.; Takahashi, H.; Katsuyama, Y.; Sugai, Y.; Ozaki, T.; Teramoto, K.; Teramoto, K.; Tanaka, K.; Abe, I.; Asamizu, S.; Onaka, H. Acyltransferase That Catalyses the Condensation of Polyketide and Peptide Moieties of Goadvionin Hybrid Lipopeptides. *Nat Chem* **2020**, *12* (9), 869–877. https://doi.org/10.1038/s41557-020-0508-2.

(264) Yang, R.; Shi, Q.; Huang, T.; Yan, Y.; Li, S.; Fang, Y.; Li, Y.; Liu, L.; Liu, L.; Wang, X.; Peng, Y.; Fan, J.; Zou, L.; Lin, S.; Chen, G. The Natural Pyrazolotriazine Pseudoiodinine from Pseudomonas Mosselii 923 Inhibits Plant Bacterial and Fungal Pathogens. *Nat Commun* **2023**, *14* (1), 734. https://doi.org/10.1038/s41467-023-36433-z.

(265) Gemperlein, K.; Zipf, G.; Bernauer, H. S.; Müller, R.; Wenzel, S. C. Metabolic Engineering of Pseudomonas Putida for Production of Docosahexaenoic Acid Based on a Myxobacterial PUFA Synthase. *Metab Eng* **2016**, *33*, 98–108. https://doi.org/10.1016/j.ymben.2015.11.001.

(266) Choi, K. R.; Cho, J. S.; Cho, I. J.; Park, D.; Lee, S. Y. Markerless Gene Knockout and Integration to Express Heterologous Biosynthetic Gene Clusters in Pseudomonas Putida. *Metab Eng* **2018**, *47*, 463–474. https://doi.org/10.1016/j.ymben.2018.05.003.

(267) Loeschcke, A.; Thies, S. Pseudomonas Putida—a Versatile Host for the Production of Natural Products. *Appl Microbiol Biotechnol* **2015**, *99* (15), 6197–6214. https://doi.org/10.1007/s00253-015-6745-4.

(268) Kampers, L. F. C.; Volkers, R. J. M.; Martins dos Santos, V. A. P. *Pseudomonas Putida* <scp>KT</Scp> 2440 Is <scp>HV</Scp> 1 Certified, Not <scp>GRAS</Scp>. *Microb Biotechnol* **2019**, *12* (5), 845–848. https://doi.org/10.1111/1751-7915.13443.

(269) Put, H.; Gerstmans, H.; Vande Capelle, H.; Fauvart, M.; Michiels, J.; Masschelein, J. *Bacillus Subtilis* as a Host for Natural Product Discovery and Engineering of Biosynthetic Gene Clusters. *Nat Prod Rep* **2024**, *41* (7), 1113–1151. https://doi.org/10.1039/D3NP00065F.

(270) Videau, P.; Wells, K. N.; Singh, A. J.; Gerwick, W. H.; Philmus, B. Assessment of Anabaena Sp. Strain PCC 7120 as a Heterologous Expression Host for Cyanobacterial Natural Products: Production of Lyngbyatoxin A. *ACS Synth Biol* **2016**, *5* (9), 978–988. https://doi.org/10.1021/acssynbio.6b00038.

(271) Taton, A.; Rohrer, S.; Diaz, B.; Reher, R.; Caraballo Rodriguez, A. M.; Pierce, M. L.; Dorrestein, P. C.; Gerwick, L.; Gerwick, W. H.; Golden, J. W. Heterologous Expression in *Anabaena* of the Columbamide Pathway from the Cyanobacterium *Moorena Bouillonii* and Production of New Analogs. *ACS Chem Biol* **2022**, *17* (7), 1910–1923. https://doi.org/10.1021/acschembio.2c00347.

(272) Taton, A.; Ecker, A.; Diaz, B.; Moss, N. A.; Anderson, B.; Reher, R.; Leão, T. F.; Simkovsky, R.; Dorrestein, P. C.; Gerwick, L.; Gerwick, W. H.; Golden, J. W. Heterologous Expression of Cryptomaldamide in a Cyanobacterial Host. *ACS Synth Biol* **2020**, *9* (12), 3364–3376. https://doi.org/10.1021/acssynbio.0c00431.

(273) Back, D.; O'Donnell, T. J.; Axt, K. K.; Gurr, J. R.; Vanegas, J. M.; Williams, P. G.; Philmus, B. Identification, Heterologous Expression, and Characterization of the Tolypodiol Biosynthetic Gene Cluster through an Integrated Approach. *ACS Chem Biol* **2023**, *18* (8), 1797–1807. https://doi.org/10.1021/acschembio.3c00225.

(274) Agarwal, V.; Blanton, J. M.; Podell, S.; Taton, A.; Schorn, M. A.; Busch, J.; Lin, Z.; Schmidt, E. W.; Jensen, P. R.; Paul, V. J.; Biggs, J. S.; Golden, J. W.; Allen, E. E.; Moore, B. S. Metagenomic Discovery of Polybrominated Diphenyl Ether Biosynthesis by Marine Sponges. *Nat Chem Biol* **2017**, *13* (5), 537–543. https://doi.org/10.1038/nchembio.2330.

(275) Zhan, C.; Lee, N.; Lan, G.; Dan, Q.; Cowan, A.; Wang, Z.; Baidoo, E. E. K.; Kakumanu, R.; Luckie, B.; Kuo, R. C.; McCauley, J.; Liu, Y.; Valencia, L.; Haushalter, R. W.; Keasling, J. D. Improved Polyketide Production in C. Glutamicum by Preventing Propionate-Induced Growth Inhibition. *Nat Metab* **2023**, *5* (7), 1127–1140. https://doi.org/10.1038/s42255-023-00830-x.

(276) Hirasawa, T.; Shimoyamada, Y.; Tachikawa, Y.; Satoh, Y.; Kawano, Y.; Dairi, T.; Ohtsu, I. Ergothioneine Production by Corynebacterium Glutamicum Harboring Heterologous Biosynthesis Pathways. *J Biosci Bioeng* **2023**, *135* (1), 25–33. https://doi.org/10.1016/j.jbiosc.2022.10.002.

(277) Liu, J.; Zhou, H.; Yang, Z.; Wang, X.; Chen, H.; Zhong, L.; Zheng, W.; Niu, W.; Wang, S.; Ren, X.; Zhong, G.; Wang, Y.; Ding, X.; Müller, R.; Zhang, Y.; Bian, X. Rational Construction of Genome-Reduced Burkholderiales Chassis Facilitates Efficient Heterologous Production of Natural Products from Proteobacteria. *Nat Commun* **2021**, *12* (1), 4347. https://doi.org/10.1038/s41467-021-24645-0.

(278) Hao, T.; Xie, Z.; Wang, M.; Liu, L.; Zhang, Y.; Wang, W.; Zhang, Z.; Zhao, X.; Li, P.; Guo, Z.; Gao, S.; Lou, C.; Zhang, G.; Merritt, J.; Horsman, G. P.; Chen, Y. An Anaerobic Bacterium Host System for Heterologous Expression of Natural Product Biosynthetic Gene Clusters. *Nat Commun* **2019**, *10* (1), 3665. https://doi.org/10.1038/s41467-019-11673-0.

(279) Li, J. S.; Barber, C. C.; Herman, N. A.; Cai, W.; Zafrir, E.; Du, Y.; Zhu, X.; Skyrud, W.; Zhang, W. Investigation of Secondary Metabolism in the Industrial Butanol Hyper-Producer *Clostridium Saccharoperbutylacetonicum* N1-4. *J Ind Microbiol Biotechnol* **2020**, *47* (3), 319–328. https://doi.org/10.1007/s10295-020-02266-8.

(280) Li, J. S.; Du, Y.; Gu, D.; Cai, W.; Green, A.; Ng, S.; Leung, A.; Del Rio Flores, A.; Zhang, W. Discovery and Biosynthesis of Clostyrylpyrones from the Obligate Anaerobe *Clostridium Roseum*. *Org Lett* **2020**, *22* (21), 8204–8209. https://doi.org/10.1021/acs.orglett.0c02656.

(281) Hwang, S.; Joung, C.; Kim, W.; Palsson, B.; Cho, B.-K. Recent Advances in Non-Model Bacterial Chassis Construction. *Curr Opin Syst Biol* **2023**, *36*, 100471. https://doi.org/10.1016/j.coisb.2023.100471.

(282) Brady, S. F. Construction of Soil Environmental DNA Cosmid Libraries and Screening for Clones That Produce Biologically Active Small Molecules. *Nat Protoc* **2007**, *2* (5), 1297–1305. https://doi.org/10.1038/nprot.2007.195.

(283) Li, Y.; Ma, Y.; Xia, Y.; Zhang, T.; Sun, S.; Gao, J.; Yao, H.; Wang, H. Discovery and Biosynthesis of Tricyclic Copper-Binding Ribosomal Peptides Containing Histidine-to-Butyrine Crosslinks. *Nat Commun* **2023**, *14* (1), 2944. https://doi.org/10.1038/s41467-023-38517-2.

(284) Yu, Z.; Huang, J.-P.; Yang, J.; Liu, C.; Yan, Y.; Wang, L.; Zhao, J.; Chen, Y.; Xiang, W.; Huang, S.-X. Discovery and Biosynthesis of Karnamicins as Angiotensin Converting Enzyme Inhibitors. *Nat Commun* **2023**, *14* (1), 209. https://doi.org/10.1038/s41467-023-35829-1.

(285) Xu, Z. F.; Zhou, Y. L.; Bo, S. T.; Zhang, S. Q.; Shi, J.; Xiang, L.; Xi, M. Y.; Zhang, B.; Xu, Z. R.; Yang, D.; Shen, B.; Tan, R. X.; Ge, H. M. Discovery and Biosynthetic Pathway Analysis of Cyclopentane–β-Lactone Globilactone A. *Nature Synthesis* **2023**, *3* (1), 99–110. https://doi.org/10.1038/s44160-023-00414-3.

(286) Shi, J.; Zeng, Y. J.; Zhang, B.; Shao, F. L.; Chen, Y. C.; Xu, X.; Sun, Y.; Xu, Q.; Tan, R. X.; Ge, H. M. Comparative Genome Mining and Heterologous Expression of an Orphan NRPS Gene Cluster Direct the Production of Ashimides. *Chem Sci* **2019**, *10* (10), 3042–3048. https://doi.org/10.1039/C8SC05670F.

(287) Popoff, A.; Hug, J. J.; Walesch, S.; Garcia, R.; Keller, L.; Müller, R. Structure and Biosynthesis of Myxofacyclines: Unique Myxobacterial Polyketides Featuring Varing and Rare Heterocycles □ **. *Chemistry – A European Journal* **2021**, *27* (67), 16654–16661. https://doi.org/10.1002/chem.202103095.

(288) Yan, F.; Auerbach, D.; Chai, Y.; Keller, L.; Tu, Q.; Hüttel, S.; Glemser, A.; Grab, H. A.; Bach, T.; Zhang, Y.; Müller, R. Biosynthesis and Heterologous Production of Vioprolides: Rational Biosynthetic Engineering and Unprecedented

4-Methylazetidinecarboxylic Acid Formation. *Angewandte Chemie International Edition* **2018**, *57* (28), 8754–8759. https://doi.org/10.1002/anie.201802479.

(289) Libis, V.; Antonovsky, N.; Zhang, M.; Shang, Z.; Montiel, D.; Maniko, J.; Ternei, M. A.; Calle, P. Y.; Lemetre, C.; Owen, J. G.; Brady, S. F. Uncovering the Biosynthetic Potential of Rare Metagenomic DNA Using Co-Occurrence Network Analysis of Targeted Sequences. *Nat Commun* **2019**, *10* (1), 3848. https://doi.org/10.1038/s41467-019-11658-z.

(290) Hashimoto, T.; Hashimoto, J.; Kozone, I.; Amagai, K.; Kawahara, T.; Takahashi, S.; Ikeda, H.; Shin-ya, K. Biosynthesis of Quinolidomicin, the Largest Known Macrolide of Terrestrial Origin: Identification and Heterologous Expression of a Biosynthetic Gene Cluster over 200 Kb. *Org Lett* **2018**, *20* (24), 7996–7999. https://doi.org/10.1021/acs.orglett.8b03570.

(291) Liu, S. H.; Wang, W.; Wang, K. B.; Zhang, B.; Li, W.; Shi, J.; Jiao, R. H.; Tan, R. X.; Ge, H. M. Heterologous Expression of a Cryptic Giant Type I PKS Gene Cluster Leads to the Production of Ansaseomycin. *Org Lett* **2019**, *21* (10), 3785–3788. https://doi.org/10.1021/acs.orglett.9b01237.

(292) Kudo, K.; Nishimura, T.; Izumikawa, M.; Kozone, I.; Hashimoto, J.; Fujie, M.; Suenaga, H.; Ikeda, H.; Satoh, N.; Shin-ya, K. Capability of a Large Bacterial Artificial Chromosome Clone Harboring Multiple Biosynthetic Gene Clusters for the Production of Diverse Compounds. *J Antibiot (Tokyo)* **2024**, *77* (5), 288–298. https://doi.org/10.1038/s41429-024-00711-9.

(293) Tan, B.; Chen, S.; Zhang, Q.; Chen, Y.; Zhu, Y.; Khan, I.; Zhang, W.; Zhang, C. Heterologous Expression Leads to Discovery of Diversified Lobophorin Analogues and a Flexible Glycosyltransferase. *Org Lett* **2020**, *22* (3), 1062–1066. https://doi.org/10.1021/acs.orglett.9b04597.

(294) Yang, Y. M.; Zhao, E. J.; Wei, W.; Xu, Z. F.; Shi, J.; Wu, X.; Zhang, B.; Igarashi, Y.; Jiao, R. H.; Liang, Y.; Tan, R. X.; Ge, H. M. Cytochrome P450 Catalyzes Benzene Ring Formation in the Biosynthesis of Trialkyl-Substituted Aromatic Polyketides. *Angewandte Chemie International Edition* **2023**, *62* (5), e202214026. https://doi.org/10.1002/anie.202214026.

(295) Gu, B.; Kim, D. G.; Kim, D.-K.; Kim, M.; Kim, H. U.; Oh, M.-K. Heterologous Overproduction of Oviedomycin by Refactoring Biosynthetic Gene Cluster and Metabolic Engineering of Host Strain Streptomyces Coelicolor. *Microb Cell Fact* **2023**, *22* (1), 212. https://doi.org/10.1186/s12934-023-02218-8.

(296) Nogawa, T.; Terai, A.; Amagai, K.; Hashimoto, J.; Futamura, Y.; Okano, A.; Fujie, M.; Satoh, N.; Ikeda, H.; Shin-ya, K.; Osada, H.; Takahashi, S. Heterologous Expression of the Biosynthetic Gene Cluster for Verticilactam and Identification of Analogues. *J Nat Prod* **2020**, *83* (12), 3598–3605. https://doi.org/10.1021/acs.jnatprod.0c00755.

(297) Zheng, Y.; Sakai, K.; Watanabe, K.; Takagi, H.; Sato-Shiozaki, Y.; Misumi, Y.; Miyanoiri, Y.; Kurisu, G.; Nogawa, T.; Takita, R.; Takahashi, S. Iron-Sulphur Protein Catalysed [4+2] Cycloadditions in Natural Product Biosynthesis. *Nat Commun* **2024**, *15* (1), 5779. https://doi.org/10.1038/s41467-024-50142-1.

(298) Jiang, G.; Zhang, Y.; Powell, M. M.; Zhang, P.; Zuo, R.; Zhang, Y.; Kallifidas, D.; Tieu, A. M.; Luesch, H.; Loria, R.; Ding, Y. High-Yield Production of Herbicidal Thaxtomins and Thaxtomin Analogs in a Nonpathogenic Streptomyces Strain. *Appl Environ Microbiol* **2018**, *84* (11), e00164-18. https://doi.org/10.1128/AEM.00164-18.

(299) Liu, W.; Zhai, S.; Zhang, L.; Chen, Y.; Liu, Z.; Ma, W.; Zhang, T.; Zhang, W.; Ma, L.; Zhang, C.; Zhang, W. Expanding the Chemical Diversity of Grisechelins via Heterologous Expression. *J Nat Prod* **2024**, *87* (2), 371–380. https://doi.org/10.1021/acs.jnatprod.3c01132.

(300) Jiang, W.; Zhao, X.; Gabrieli, T.; Lou, C.; Ebenstein, Y.; Zhu, T. F. Cas9-Assisted Targeting of CHromosome Segments CATCH Enables One-Step Targeted Cloning of Large Gene Clusters. *Nat Commun* **2015**, *6* (1), 8101. https://doi.org/10.1038/ncomms9101.

(301) Román-Hurtado, F.; Sánchez-Hidalgo, M.; Martín, J.; Ortiz-López, F.; Genilloud, O. Biosynthesis and Heterologous Expression of Cacaoidin, the First Member of the Lanthidin Family of RiPPs. *Antibiotics* **2021**, *10* (4), 403. https://doi.org/10.3390/antibiotics10040403.

(302) Román-Hurtado, F.; Sánchez-Hidalgo, M.; Martín, J.; Ortiz-López, F. J.; Carretero-Molina, D.; Reyes, F.; Genilloud, O. One Pathway, Two Cyclic Non-Ribosomal Pentapeptides: Heterologous Expression of BE-18257 Antibiotics and Pentaminomycins from Streptomyces Cacaoi CA-170360. *Microorganisms* **2021**, *9* (1), 135. https://doi.org/10.3390/microorganisms9010135.

(303) Li, H.; Pan, Y.; Liu, G. Multiplying the Heterologous Production of Spinosad through Tandem Amplification of Its Biosynthetic Gene Cluster in Streptomyces Coelicolor. *Microb Biotechnol* **2022**, *15* (5), 1550–1560. https://doi.org/10.1111/1751-7915.13965.

(304) Cao, M.; Zheng, C.; Yang, D.; Kalkreuter, E.; Adhikari, A.; Liu, Y.; Rateb, M. E.; Shen, B. Cryptic Sulfur Incorporation in Thioangucycline Biosynthesis. *Angewandte Chemie International Edition* **2021**, *60* (13), 7140–7147. https://doi.org/10.1002/anie.202015570.

(305) Zetsche, B.; Gootenberg, J. S.; Abudayyeh, O. O.; Slaymaker, I. M.; Makarova, K. S.; Essletzbichler, P.; Volz, S. E.; Joung, J.; van der Oost, J.; Regev, A.; Koonin, E. V.; Zhang, F. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **2015**, *163* (3), 759–771. https://doi.org/10.1016/j.cell.2015.09.038.

(306) Enghiad, B.; Huang, C.; Guo, F.; Jiang, G.; Wang, B.; Tabatabaei, S. K.; Martin, T. A.; Zhao, H. Cas12a-Assisted Precise Targeted Cloning Using in Vivo Cre-Lox Recombination. *Nat Commun* **2021**, *12* (1), 1171. https://doi.org/10.1038/s41467-021-21275-4.

(307) Ren, H.; Dommaraju, S. R.; Huang, C.; Cui, H.; Pan, Y.; Nesic, M.; Zhu, L.; Sarlah, D.; Mitchell, D. A.; Zhao, H. Genome Mining Unveils a Class of Ribosomal Peptides with Two Amino Termini. *Nat Commun* **2023**, *14* (1), 1624. https://doi.org/10.1038/s41467-023-37287-1.

(308) Ren, H.; Huang, C.; Pan, Y.; Dommaraju, S. R.; Cui, H.; Li, M.; Gadgil, M. G.; Mitchell, D. A.; Zhao, H. Non-Modular Fatty Acid Synthases Yield Distinct N-Terminal Acylation in Ribosomal Peptides. *Nat Chem* **2024**, *16* (8), 1320–1329. https://doi.org/10.1038/s41557-024-01491-3.

(309) Dhakal, D.; Kallifidas, D.; Chen, M.; Kokkaliari, S.; Chen, Q.-Y.; Paul, V. J.; Ding, Y.; Luesch, H. Heterologous Production of the C33–C45 Polyketide Fragment of Anticancer Apratoxins in a Cyanobacterial Host. *Org Lett* **2023**, *25* (13), 2238–2242. https://doi.org/10.1021/acs.orglett.3c00462.

(310) Liang, M.; Liu, L.; Xu, F.; Zeng, X.; Wang, R.; Yang, J.; Wang, W.; Karthik, L.; Liu, J.; Yang, Z.; Zhu, G.; Wang, S.; Bai, L.; Tong, Y.; Liu, X.; Wu, M.; Zhang, L.-X.; Tan, G.-Y. Activating Cryptic Biosynthetic Gene Cluster through a CRISPR–Cas12a-Mediated Direct Cloning Approach. *Nucleic Acids Res* **2022**, *50* (6), 3581–3592. https://doi.org/10.1093/nar/gkac181.

(311) Kouprina, N.; Larionov, V. Transformation-Associated Recombination (TAR) Cloning for Genomics Studies and Synthetic Biology. *Chromosoma* **2016**, *125* (4), 621–632. https://doi.org/10.1007/s00412-016-0588-3.

(312) Zhang, J. J.; Yamanaka, K.; Tang, X.; Moore, B. S. Direct Cloning and Heterologous Expression of Natural Product Biosynthetic Gene Clusters by Transformation-Associated Recombination. In *Methods in Enzymology*; Academic Press, 2019; Vol. 621, pp 87–110. https://doi.org/10.1016/bs.mie.2019.02.026.

(313) Groß, S.; Schnell, B.; Haack, P. A.; Auerbach, D.; Müller, R. In Vivo and in Vitro Reconstitution of Unique Key Steps in Cystobactamid Antibiotic Biosynthesis. *Nat Commun* **2021**, *12* (1), 1696. https://doi.org/10.1038/s41467-021-21848-3.

(314) Agarwal, V.; El Gamal, A. A.; Yamanaka, K.; Poth, D.; Kersten, R. D.; Schorn, M.; Allen, E. E.; Moore, B. S. Biosynthesis of Polybrominated Aromatic Organic Compounds by Marine Bacteria. *Nat Chem Biol* **2014**, *10* (8), 640–647. https://doi.org/10.1038/nchembio.1564.

(315) Tang, X.; Li, J.; Moore, B. S. Minimization of the Thiolactomycin Biosynthetic Pathway Reveals That the Cytochrome P450 Enzyme TlmF Is Required for Five-Membered Thiolactone Ring Formation. *ChemBioChem* **2017**, *18* (12), 1072–1076. https://doi.org/10.1002/cbic.201700090.

(316) Sigrist, R.; Luhavaya, H.; McKinnie, S. M. K.; Ferreira da Silva, A.; Jurberg, I. D.; Moore, B. S.; Gonzaga de Oliveira, L. Nonlinear Biosynthetic Assembly of Alpiniamide by a Hybrid *Cis* / *Trans* -AT PKS-NRPS. *ACS Chem Biol* **2020**, *15* (4), 1067–1077. https://doi.org/10.1021/acschembio.0c00081.

(317) Petrova, Y. D.; Zhao, J.; Webster, G.; Mullins, A. J.; Williams, K.; Alswat, A. S.; Challis, G. L.; Bailey, A. M.; Mahenthiralingam, E. Cloning and Expression of *Burkholderia* Polyyne Biosynthetic Gene Clusters in *Paraburkholderia* Hosts Provides a Strategy for Biopesticide Development. *Microb Biotechnol* **2022**, *15* (10), 2547–2561. https://doi.org/10.1111/1751-7915.14106.

(318) Wang, S.; Lin, S.; Fang, Q.; Gyampoh, R.; Lu, Z.; Gao, Y.; Clarke, D. J.; Wu, K.; Trembleau, L.; Yu, Y.; Kyeremeh, K.; Milne, B. F.; Tabudravu, J.; Deng, H. A Ribosomally Synthesised and Post-Translationally Modified Peptide Containing a β-Enamino Acid and a Macrocyclic Motif. *Nat Commun* **2022**, *13* (1), 5044. https://doi.org/10.1038/s41467-022-32774-3.

(319) Fu, J.; Bian, X.; Hu, S.; Wang, H.; Huang, F.; Seibert, P. M.; Plaza, A.; Xia, L.; Müller, R.; Stewart, A. F.; Zhang, Y. Full-Length RecE Enhances Linear-Linear Homologous Recombination and Facilitates Direct Cloning for Bioprospecting. *Nat Biotechnol* **2012**, *30* (5), 440–446. https://doi.org/10.1038/nbt.2183.

(320) Abbasi, M. N.; Fu, J.; Bian, X.; Wang, H.; Zhang, Y.; Li, A. Recombineering for Genetic Engineering of Natural Product Biosynthetic Pathways. *Trends Biotechnol* **2020**, *38* (7), 715–728. https://doi.org/10.1016/j.tibtech.2019.12.018.

(321) Maresca, M.; Erler, A.; Fu, J.; Friedrich, A.; Zhang, Y.; Stewart, A. F. Single-Stranded Heteroduplex Intermediates in λ Red Homologous Recombination. *BMC Mol Biol* **2010**, *11* (1), 54. https://doi.org/10.1186/1471-2199-11-54.

(322) Yang, R.; Feng, J.; Xiang, H.; Cheng, B.; Shao, L.-D.; Li, Y.-P.; Wang, H.; Hu, Q.-F.; Xiao, W.-L.; Matsuda, Y.; Wang, W.-G. Ketoreductase Domain-Catalyzed Polyketide Chain Release in Fungal Alkyl Salicylaldehyde Biosynthesis. *J Am Chem Soc* **2023**, *145* (20), 11293–11300. https://doi.org/10.1021/jacs.3c02011.

(323) Wang, H.; Li, Z.; Jia, R.; Yin, J.; Li, A.; Xia, L.; Yin, Y.; Müller, R.; Fu, J.; Stewart, A. F.; Zhang, Y. ExoCET: Exonuclease in Vitro Assembly Combined with RecET Recombination for Highly Efficient Direct DNA Cloning from Complex Genomes. *Nucleic Acids Res* **2018**, *46* (5), e28–e28. https://doi.org/10.1093/nar/gkx1249.

(324) Liu, Y.; Zhou, H.; Shen, Q.; Dai, G.; Yan, F.; Li, X.; Ren, X.; Sun, Q.; Tang, Y.-J.; Zhang, Y.; Bian, X. Discovery of Polycyclic Macrolide Shuangdaolides by Heterologous Expression of a Cryptic *Trans* -AT PKS Gene Cluster. *Org Lett* **2021**, *23* (17), 6967–6971. https://doi.org/10.1021/acs.orglett.1c02589.

(325) Song, C.; Luan, J.; Cui, Q.; Duan, Q.; Li, Z.; Gao, Y.; Li, R.; Li, A.; Shen, Y.; Li, Y.; Stewart, A. F.; Zhang, Y.; Fu, J.; Wang, H. Enhanced Heterologous Spinosad Production from a 79-Kb Synthetic Multioperon Assembly. *ACS Synth Biol* **2019**, *8* (1), 137–147. https://doi.org/10.1021/acssynbio.8b00402.

(326) Zhong, L.; Diao, X.; Zhang, N.; Li, F.; Zhou, H.; Chen, H.; Bai, X.; Ren, X.; Zhang, Y.; Wu, D.; Bian, X. Engineering and Elucidation of the Lipoinitiation Process in Nonribosomal Peptide Biosynthesis. *Nat Commun* **2021**, *12* (1). https://doi.org/10.1038/s41467-020-20548-8.

(327) Patel, J. R.; Oh, J.; Wang, S.; Crawford, J. M.; Isaacs, F. J. Cross-Kingdom Expression of Synthetic Genetic Elements Promotes Discovery of Metabolites in the Human Microbiome. *Cell* **2022**, *185* (9), 1487-1505.e14. https://doi.org/10.1016/j.cell.2022.03.008.

(328) Lin, D.; O'Callaghan, C. A. MetClo: Methylase-Assisted Hierarchical DNA Assembly Using a Single Type IIS Restriction Enzyme. *Nucleic Acids Res* **2018**, *46* (19), e113–e113. https://doi.org/10.1093/nar/gky596.

(329) Fang, H.; Zhao, J.; Zhao, X.; Dong, N.; Zhao, Y.; Zhang, D. Standardized Iterative Genome Editing Method for *Escherichia Coli* Based on CRISPR-Cas9. *ACS Synth Biol* **2024**, *13* (2), 613–623. https://doi.org/10.1021/acssynbio.3c00585.

(330) Kishore, S.; Del Rio Flores, A.; Lynch, S. R.; Yuet, K. P.; Khosla, C. Discovery and Characterization of the Fully Decorated Nocardiosis-Associated Polyketide Natural Product. *J Am Chem Soc* **2024**, *146* (6), 4212–4220. https://doi.org/10.1021/jacs.3c13670.

(331) Yuet, K. P.; Liu, C. W.; Lynch, S. R.; Kuo, J.; Michaels, W.; Lee, R. B.; McShane, A. E.; Zhong, B. L.; Fischer, C. R.; Khosla, C. Complete Reconstitution and Deorphanization of the 3 MDa Nocardiosis-Associated Polyketide Synthase. *J Am Chem Soc* **2020**, *142* (13), 5952–5957. https://doi.org/10.1021/jacs.0c00904.

(332) Del Rio Flores, A.; Khosla, C. Characterization of the Flavin-Dependent Monooxygenase Involved in the Biosynthesis of the Nocardiosis-Associated Polyketide. *Biochemistry* **2024**, *63* (21), 2868–2877. https://doi.org/10.1021/acs.biochem.4c00480.

(333) Hoshino, S.; Ijichi, S.; Asamizu, S.; Onaka, H. Insights into Arsenic Secondary Metabolism in Actinomycetes from the Structure and Biosynthesis of Bisenarsan. *J Am Chem Soc* **2023**, *145* (32), 17863–17871. https://doi.org/10.1021/jacs.3c04978.

(334) Gui, C.; Kalkreuter, E.; Lauterbach, L.; Yang, D.; Shen, B. Enediyne Natural Product Biosynthesis Unified by a Diiodotetrayne Intermediate. *Nat Chem Biol* **2024**, *20* (9), 1210–1219. https://doi.org/10.1038/s41589-024-01636-y.

(335) Cohen, D. R.; Townsend, C. A. A Dual Role for a Polyketide Synthase in Dynemicin Enediyne and Anthraquinone Biosynthesis. *Nature Chemistry 2017 10:2* **2017**, *10* (2), 231–236. https://doi.org/10.1038/nchem.2876.

(336) Bhardwaj, M.; Cui, Z.; Daniel Hankore, E.; Moonschi, F. H.; Saghaeiannejad Esfahani, H.; Kalkreuter, E.; Gui, C.; Yang, D.; Phillips, G. N.; Thorson, J. S.; Shen, B.; Van Lanen, S. G. A Discrete Intermediate for the Biosynthesis of Both the Enediyne Core and the Anthraquinone Moiety of Enediyne Natural Products. *Proceedings of the National Academy of Sciences* **2023**, *120* (9), e2220468120. https://doi.org/10.1073/pnas.2220468120.

(337) Gao, X.; Wang, P.; Tang, Y. Engineered Polyketide Biosynthesis and Biocatalysis in Escherichia Coli. *Appl Microbiol Biotechnol* **2010**, *88* (6), 1233–1242. https://doi.org/10.1007/s00253-010-2860-4.

(338) Wang, W.; Li, S.; Li, Z.; Zhang, J.; Fan, K.; Tan, G.; Ai, G.; Lam, S. M.; Shui, G.; Yang, Z.; Lu, H.; Jin, P.; Li, Y.; Chen, X.; Xia, X.; Liu, X.; Dannelly, H. K.; Yang, C.; Yang, Y.; Zhang, S.; Alterovitz, G.; Xiang, W.; Zhang, L. Harnessing the Intracellular Triacylglycerols for Titer Improvement of Polyketides in Streptomyces. *Nat Biotechnol* **2020**, *38* (1), 76–83. https://doi.org/10.1038/s41587-019-0335-4.

(339) Crooke, A. M.; Chand, A. K.; Cui, Z.; Balskus, E. P. Elucidation of Chalkophomycin Biosynthesis Reveals *N* -Hydroxypyrrole-Forming Enzymes. *J Am Chem Soc* **2024**, *146* (23), 16268–16280. https://doi.org/10.1021/jacs.4c04712.

(340) Gao, Y.; Walt, C.; Bader, C. D.; Müller, R. Genome-Guided Discovery of the Myxobacterial Thiolactone-Containing Sorangibactins. *ACS Chem Biol* **2023**, *18* (4), 924–932. https://doi.org/10.1021/acschembio.3c00063.

(341) Van Cura, D.; Ng, T. L.; Huang, J.; Hager, H.; Hartwig, J. F.; Keasling, J. D.; Balskus, E. P. Discovery of the Azaserine Biosynthetic Pathway Uncovers a Biological Route for A-Diazoester Production. *Angewandte Chemie International Edition* **2023**, *62* (28), e202304646. https://doi.org/10.1002/anie.202304646.

(342) Harris, N. C.; Sato, M.; Herman, N. A.; Twigg, F.; Cai, W.; Liu, J.; Zhu, X.; Downey, J.; Khalaf, R.; Martin, J.; Koshino, H.; Zhang, W. Biosynthesis of Isonitrile Lipopeptides by Conserved Nonribosomal Peptide Synthetase Gene Clusters in Actinobacteria. *Proceedings of the National Academy of Sciences* **2017**, *114* (27), 7025–7030. https://doi.org/10.1073/pnas.1705016114.

(343) Effert, J.; Westphalen, M.; Calderari, A.; Shi, Y.; Elamri, I.; Najah, S.; Grün, P.; Li, Y.; Gruez, A.; Weissman, K. J.; Bode, H. B. Pyrrolizwilline, a Unique Bacterial Alkaloid Assembled by a Nonribosomal Peptide Synthetase and Non-Enzymatic Dimerization. *Angewandte Chemie International Edition* **2024**, *63* (51), e202411258. https://doi.org/10.1002/anie.202411258.

(344) Cao, L.; Do, T.; Zhu, A.; Duan, J.; Alam, N.; Link, A. J. Genome Mining and Discovery of Imiditides, a Family of RiPPs with a Class-Defining Aspartimide Modification. *J Am Chem Soc* **2023**, *145* (34), 18834–18845. https://doi.org/10.1021/jacs.3c03991.

(345) Choi, B.; Elashal, H. E.; Cao, L.; Link, A. J. Mechanistic Analysis of the Biosynthesis of the Aspartimidylated Graspetide Amycolimiditide. *J Am Chem Soc* **2022**, *144* (47), 21628–21639. https://doi.org/10.1021/jacs.2c09004.

(346) Wang, X.; Chen, X.; Wang, Z.-J.; Zhuang, M.; Zhong, L.; Fu, C.; Garcia, R.; Müller, R.; Zhang, Y.; Yan, J.; Wu, D.; Huo, L. Discovery and Characterization of a Myxobacterial Lanthipeptide with Unique Biosynthetic Features and Anti-Inflammatory Activity. *J Am Chem Soc* **2023**, *145* (30), 16924–16937. https://doi.org/10.1021/jacs.3c06014.

(347) Ramesh, S.; Guo, X.; DiCaprio, A. J.; De Lio, A. M.; Harris, L. A.; Kille, B. L.; Pogorelov, T. V.; Mitchell, D. A. Bioinformatics-Guided Expansion and Discovery of Graspetides. *ACS Chem Biol* **2021**, *16* (12), 2787–2797. https://doi.org/10.1021/acschembio.1c00672.

(348) Zhang, Z. J.; Wu, C.; Moreira, R.; Dorantes, D.; Pappas, T.; Sundararajan, A.; Lin, H.; Pamer, E. G.; van der Donk, W. A. Activity of Gut-Derived Nisin-like Lantibiotics against Human Gut Pathogens and Commensals. *ACS Chem Biol* **2024**, *19* (2), 357–369. https://doi.org/10.1021/acschembio.3c00577.

(349) Pei, Z.-F.; Zhu, L.; Sarksian, R.; van der Donk, W. A.; Nair, S. K. Class V Lanthipeptide Cyclase Directs the Biosynthesis of a Stapled Peptide Natural Product. *J Am Chem Soc* **2022**, *144* (38), 17549–17557. https://doi.org/10.1021/jacs.2c06808.

(350) Sarksian, R.; Hegemann, J. D.; Simon, M. A.; Acedo, J. Z.; van der Donk, W. A. Unexpected Methyllanthionine Stereochemistry in the Morphogenetic Lanthipeptide SapT. *J Am Chem Soc* **2022**, *144* (14), 6373–6382. https://doi.org/10.1021/jacs.2c00517.

(351) Liu, C. L.; Wang, Z. J.; Shi, J.; Yan, Z. Y.; Zhang, G. D.; Jiao, R. H.; Tan, R. X.; Ge, H. M. P450-Modified Multicyclic Cyclophane-Containing Ribosomally Synthesized and Post-Translationally Modified Peptides. *Angewandte Chemie* **2024**, *136* (10), e202314046. https://doi.org/10.1002/ange.202314046.

(352) Clark, K. A.; Seyedsayamdost, M. R. Bioinformatic Atlas of Radical SAM Enzyme-Modified RiPP Natural Products Reveals an Isoleucine–Tryptophan Crosslink. *J Am Chem Soc* **2022**, *144* (39), 17876–17888. https://doi.org/10.1021/jacs.2c06497.

(353) Johnson, B. A.; Clark, K. A.; Bushin, L. B.; Spolar, C. N.; Seyedsayamdost, M. R. Expanding the Landscape of Noncanonical Amino Acids in RiPP Biosynthesis. *J Am Chem Soc* **2024**, *146* (6), 3805–3815. https://doi.org/10.1021/jacs.3c10824.

(354) Hu, Y. L.; Yin, F. Z.; Shi, J.; Ma, S. Y.; Wang, Z. R.; Tan, R. X.; Jiao, R. H.; Ge, H. M. P450-Modified Ribosomally Synthesized Peptides with Aromatic Cross-Links. *J Am Chem Soc* **2023**, *145* (50), 27325–27335. https://doi.org/10.1021/jacs.3c07416.

(355) Carson, D. V.; Juarez, R. J.; Do, T.; Yang, Z. J.; Link, A. J. Antimicrobial Lasso Peptide Cloacaenodin Utilizes a Unique TonB-Dependent Transporter to Access Susceptible Bacteria. *ACS Chem Biol* **2024**, *19* (4), 981–991. https://doi.org/10.1021/acschembio.4c00009.

(356) Elashal, H. E.; Koos, J. D.; Cheung-Lee, W. L.; Choi, B.; Cao, L.; Richardson, M. A.; White, H. L.; Link, A. J. Biosynthesis and Characterization of Fuscimiditide, an Aspartimidylated Graspetide. *Nat Chem* **2022**, *14* (11), 1325–1334. https://doi.org/10.1038/s41557-022-01022-y.

(357) Ayikpoe, R. S.; Shi, C.; Battiste, A. J.; Eslami, S. M.; Ramesh, S.; Simon, M. A.; Bothwell, I. R.; Lee, H.; Rice, A. J.; Ren, H.; Tian, Q.; Harris, L. A.; Sarksian, R.; Zhu, L.; Frerk, A. M.; Precord, T. W.; van der Donk, W. A.; Mitchell, D. A.; Zhao, H. A Scalable Platform to Discover Antimicrobials of Ribosomal Origin. *Nat Commun* **2022**, *13* (1), 6135. https://doi.org/10.1038/s41467-022-33890-w.

(358) Kenney, G. E.; Dassama, L. M. K.; Pandelia, M.-E.; Gizzi, A. S.; Martinie, R. J.; Gao, P.; DeHart, C. J.; Schachner, L. F.; Skinner, O. S.; Ro, S. Y.; Zhu, X.; Sadek, M.; Thomas, P. M.; Almo, S. C.; Bollinger, J. M.; Krebs, C.; Kelleher, N. L.;

Rosenzweig, A. C. The Biosynthesis of Methanobactin. *Science (1979)* **2018**, *359* (6382), 1411–1416. https://doi.org/10.1126/science.aap9437.

(359) Ayikpoe, R. S.; Zhu, L.; Chen, J. Y.; Ting, C. P.; van der Donk, W. A. Macrocyclization and Backbone Rearrangement During RiPP Biosynthesis by a SAM-Dependent Domain-of-Unknown-Function 692. *ACS Cent Sci* **2023**, *9* (5), 1008–1018. https://doi.org/10.1021/acscentsci.3c00160.

(360) Ting, C. P.; Funk, M. A.; Halaby, S. L.; Zhang, Z.; Gonen, T.; van der Donk, W. A. Use of a Scaffold Peptide in the Biosynthesis of Amino Acid–Derived Natural Products. *Science (1979)* **2019**, *365* (6450), 280–284. https://doi.org/10.1126/science.aau6232.

(361) Chioti, V. T.; Clark, K. A.; Ganley, J. G.; Han, E. J.; Seyedsayamdost, M. R. N–Cα Bond Cleavage Catalyzed by a Multinuclear Iron Oxygenase from a Divergent Methanobactin-like RiPP Gene Cluster. *J Am Chem Soc* **2024**, *146* (11), 7313–7323. https://doi.org/10.1021/jacs.3c11740.

(362) Nguyen, D. T.; Zhu, L.; Gray, D. L.; Woods, T. J.; Padhi, C.; Flatt, K. M.; Mitchell, D. A.; van der Donk, W. A. Biosynthesis of Macrocyclic Peptides with C-Terminal β-Amino-α-Keto Acid Groups by Three Different Metalloenzymes. *ACS Cent Sci* **2024**, *10* (5), 1022–1032. https://doi.org/10.1021/acscentsci.4c00088.

(363) Bushin, L. B.; Covington, B. C.; Rued, B. E.; Federle, M. J.; Seyedsayamdost, M. R. Discovery and Biosynthesis of Streptosactin, a Sactipeptide with an Alternative Topology Encoded by Commensal Bacteria in the Human Microbiome. *J Am Chem Soc* **2020**, *142* (38), 16265–16275. https://doi.org/10.1021/jacs.0c05546.

(364) Imai, Y.; Meyer, K. J.; Iinishi, A.; Favre-Godal, Q.; Green, R.; Manuse, S.; Caboni, M.; Mori, M.; Niles, S.; Ghiglieri, M.; Honrao, C.; Ma, X.; Guo, J. J.; Makriyannis, A.; Linares-Otoya, L.; Böhringer, N.; Wuisan, Z. G.; Kaur, H.; Wu, R.; Mateus, A.; Typas, A.; Savitski, M. M.; Espinoza, J. L.; O'Rourke, A.; Nelson, K. E.; Hiller, S.; Noinaj, N.; Schäberle, T. F.; D'Onofrio, A.; Lewis, K. A New Antibiotic Selectively Kills Gram-Negative Pathogens. *Nature* **2019**, *576* (7787), 459–464. https://doi.org/10.1038/s41586-019-1791-1.

(365) Guo, S.; Wang, S.; Ma, S.; Deng, Z.; Ding, W.; Zhang, Q. Radical SAM-Dependent Ether Crosslink in Daropeptide Biosynthesis. *Nat Commun* **2022**, *13* (1), 2361. https://doi.org/10.1038/s41467-022-30084-2.

(366) Nguyen, H.; Made Kresna, I. D.; Böhringer, N.; Ruel, J.; Mora, E. de la; Kramer, J.-C.; Lewis, K.; Nicolet, Y.; Schäberle, T. F.; Yokoyama, K. Characterization of a Radical SAM Oxygenase for the Ether Crosslinking in Darobactin Biosynthesis. *J Am Chem Soc* **2022**, *144* (41), 18876–18886. https://doi.org/10.1021/jacs.2c05565.

(367) Wuisan, Z. G.; Kresna, I. D. M.; Böhringer, N.; Lewis, K.; Schäberle, T. F. Optimization of Heterologous Darobactin A Expression and Identification of the Minimal Biosynthetic Gene Cluster. *Metab Eng* **2021**, *66*, 123–136. https://doi.org/10.1016/j.ymben.2021.04.007.

(368) Pei, Z.-F.; Vior, N. M.; Zhu, L.; Truman, A. W.; Nair, S. K. Biosynthesis of Peptide–Nucleobase Hybrids in Ribosomal Peptides. *Nat Chem Biol* **2024**. https://doi.org/10.1038/s41589-024-01736-9.

(369) Hubrich, F.; Kandy, S. K.; Chepkirui, C.; Padhi, C.; Mordhorst, S.; Moosmann, P.; Zhu, T.; Gugger, M.; Chekan, J. R.; Piel, J. Ribosomal Peptides with Polycyclic Isoprenoid Moieties. *Chem* **2024**, *10* (10), 3224–3242. https://doi.org/10.1016/j.chempr.2024.07.026.

(370) Daniels, P. N.; Lee, H.; Splain, R. A.; Ting, C. P.; Zhu, L.; Zhao, X.; Moore, B. S.; van der Donk, W. A. A Biosynthetic Pathway to Aromatic Amines That Uses Glycyl-TRNA as Nitrogen Donor. *Nat Chem* **2022**, *14* (1), 71–77. https://doi.org/10.1038/s41557-021-00802-2.

(371) Ramos Figueroa, J.; Zhu, L.; van der Donk, W. A. Unexpected Transformations during Pyrroloiminoquinone Biosynthesis. *J Am Chem Soc* **2024**, *146* (20), 14235–14245. https://doi.org/10.1021/jacs.4c03677.

(372) Yu, Y.; van der Donk, W. A. Biosynthesis of 3-Thia-α-Amino Acids on a Carrier Peptide. *Proceedings of the National Academy of Sciences* **2022**, *119* (29), e2205285119. https://doi.org/10.1073/pnas.2205285119.

(373) Ward, V. C. A.; Chatzivasileiou, A. O.; Stephanopoulos, G. Metabolic Engineering of Escherichia Coli for the Production of Isoprenoids. *FEMS Microbiol Lett* **2018**, *365* (10), 79. https://doi.org/10.1093/femsle/fny079.

(374) Martin, V. J. J.; Pitera, D. J.; Withers, S. T.; Newman, J. D.; Keasling, J. D. Engineering a Mevalonate Pathway in Escherichia Coli for Production of Terpenoids. *Nat Biotechnol* **2003**, *21* (7), 796–802. https://doi.org/10.1038/nbt833.

(375) Aguilar, F.; Ekramzadeh, K.; Scheper, T.; Beutel, S. Whole-Cell Production of Patchouli Oil Sesquiterpenes in *Escherichia Coli*: Metabolic Engineering and Fermentation Optimization in Solid–Liquid Phase Partitioning Cultivation. *ACS Omega* **2020**, *5* (50), 32436–32446. https://doi.org/10.1021/acsomega.0c04590.

(376) Sun, C.; Zhang, R.; Xie, C. Biosynthesis of (*R*)-(+)-perillyl Alcohol by *Escherichia Coli* Expressing Neryl Pyrophosphate Synthase. *Eng Life Sci* **2022**, *22* (5), 407–416. https://doi.org/10.1002/elsc.202100135.

(377) Zhang, J.; Wang, X.; Zhang, X.; Zhang, Y.; Wang, F.; Li, X. Sesquiterpene Synthase Engineering and Targeted Engineering of α-Santalene Overproduction in *Escherichia Coli*. *J Agric Food Chem* **2022**, *70* (17), 5377–5385. https://doi.org/10.1021/acs.jafc.2c00754.

(378) Lund, S.; Hall, R.; Williams, G. J. An Artificial Pathway for Isoprenoid Biosynthesis Decoupled from Native Hemiterpene Metabolism. *ACS Synth Biol* **2019**, *8* (2), 232–238. https://doi.org/10.1021/acssynbio.8b00383.

(379) Chatzivasileiou, A. O.; Ward, V.; Edgar, S. M.; Stephanopoulos, G. Two-Step Pathway for Isoprenoid Synthesis. *Proceedings of the National Academy of Sciences* **2019**, *116* (2), 506–511. https://doi.org/10.1073/pnas.1812935116.

(380) Ma, X.; Liang, H.; Pan, Q.; Prather, K. L. J.; Sinskey, A. J.; Stephanopoulos, G.; Zhou, K. Optimization of the Isopentenol Utilization Pathway for Isoprenoid Synthesis in *Escherichia Coli*. *J Agric Food Chem* **2022**, *70* (11), 3512–3520. https://doi.org/10.1021/acs.jafc.2c00014.

(381) Pan, Q.; Ma, X.; Liang, H.; Liu, Y.; Zhou, Y.; Stephanopoulos, G.; Zhou, K. Biosynthesis of Geranate via Isopentenol Utilization Pathway in *Escherichia Coli*. *Biotechnol Bioeng* **2023**, *120* (1), 230–238. https://doi.org/10.1002/bit.28255.

(382) Xiao, Z.; Yang, Q.; Lin, X.; Li, F.-R.; Zhang, X.; Xu, H.-M.; Wang, Z.; Wang, J.; Dong, L.-B. Cytochrome P450-Mediated Skeleton Rearrangement of Taxadiene in an Engineered *Escherichia Coli* System. *Org Lett* **2024**, *26* (8), 1640–1644. https://doi.org/10.1021/acs.orglett.4c00103.

(383) Kumar, V.; Johnson, B. P.; Dimas, D. A.; Singh, S. Novel Homologs of Isopentenyl Phosphate Kinase Reveal Class-Wide Substrate Flexibility. *ChemCatChem* **2021**, *13* (17), 3781–3788. https://doi.org/10.1002/cctc.202100595.

(384) Hausjell, J.; Halbwirth, H.; Spadiut, O. Recombinant Production of Eukaryotic Cytochrome P450s in Microbial Cell Factories. *Biosci Rep* **2018**, *38* (2). https://doi.org/10.1042/BSR20171290.

(385) Dueber, J. E.; Wu, G. C.; Malmirchegini, G. R.; Moon, T. S.; Petzold, C. J.; Ullal, A. V; Prather, K. L. J.; Keasling, J. D. Synthetic Protein Scaffolds Provide Modular Control over Metabolic Flux. *Nat Biotechnol* **2009**, *27* (8), 753–759. https://doi.org/10.1038/nbt.1557.

(386) Kokorin, A.; Parshin, P. D.; Bakkes, P. J.; Pometun, A. A.; Tishkov, V. I.; Urlacher, V. B. Genetic Fusion of P450 BM3 and Formate Dehydrogenase towards Self-Sufficient Biocatalysts with Enhanced Activity. *Sci Rep* **2021**, *11* (1), 21706. https://doi.org/10.1038/s41598-021-00957-5.

(387) Park, S. Y.; Eun, H.; Lee, M. H.; Lee, S. Y. Metabolic Engineering of Escherichia Coli with Electron Channelling for the Production of Natural Products. *Nat Catal* **2022**, *5* (8), 726–737. https://doi.org/10.1038/s41929-022-00820-4.

(388) Wang, Z.; Dai, Y.; Azi, F.; Wang, Z.; Xu, W.; Wang, D.; Dong, M.; Xia, X. Constructing Protein-Scaffolded Multienzyme Assembly Enhances the Coupling Efficiency of the P450 System for Efficient Daidzein Biosynthesis from (2 *S*)-Naringenin. *J Agric Food Chem* **2024**, *72* (11), 5849–5859. https://doi.org/10.1021/acs.jafc.3c09854.

(389) Lescallette, A. R.; Dunn, Z. D.; Manning, V. A.; Trippe, K. M.; Li, B. Biosynthetic Origin of Formylaminooxyvinylglycine and Characterization of the Formyltransferase GvgI. *Biochemistry* **2022**, *61* (19), 2159–2164. https://doi.org/10.1021/acs.biochem.2c00374.

(390) de Rond, T.; Asay, J. E.; Moore, B. S. Co-Occurrence of Enzyme Domains Guides the Discovery of an Oxazolone Synthetase. *Nat Chem Biol* **2021**, *17* (7), 794–799. https://doi.org/10.1038/s41589-021-00808-4.

(391) Yu, L.; Zhou, W.; She, Y.; Ma, H.; Cai, Y.-S.; Jiang, M.; Deng, Z.; Price, N. P. J.; Chen, W. Efficient Biosynthesis of Nucleoside Cytokinin Angustmycin A Containing an Unusual Sugar System. *Nat Commun* **2021**, *12* (1), 6633. https://doi.org/10.1038/s41467-021-26928-y.

(392) Shiraishi, T.; Xia, J.; Kato, T.; Kuzuyama, T. Biosynthesis of the Nucleoside Antibiotic Angustmycins: Identification and Characterization of the Biosynthetic Gene Cluster Reveal Unprecedented Dehydratase Required for Exo-Glycal Formation. *J Antibiot (Tokyo)* **2021**, *74* (11), 830–833. https://doi.org/10.1038/s41429-021-00466-7.

(393) D'Agostino, P. M.; Seel, C. J.; Ji, X.; Gulder, T.; Gulder, T. A. M. Biosynthesis of Cyanobacterin, a Paradigm for Furanolide Core Structure Assembly. *Nat Chem Biol* **2022**, *18* (6), 652–658. https://doi.org/10.1038/s41589-022-01013-7.

(394) Del Rio Flores, A.; Twigg, F. F.; Du, Y.; Cai, W.; Aguirre, D. Q.; Sato, M.; Dror, M. J.; Narayanamoorthy, M.; Geng, J.; Zill, N. A.; Zhai, R.; Zhang, W. Biosynthesis of Triacsin Featuring an N-Hydroxytriazene Pharmacophore. *Nat Chem Biol* **2021**, *17* (12), 1305–1313. https://doi.org/10.1038/s41589-021-00895-3.

(395) Ushimaru, R.; Lyu, J.; Ling, M.; Abe, I. Multiple C–C Bond Cleavage Reactions Catalyzed by Tolyporphin Tetrapyrrole Biosynthetic Enzymes. *J Am Chem Soc* **2023**, *145* (17), 9834–9839. https://doi.org/10.1021/jacs.3c01993.

(396) Hibi, G.; Shiraishi, T.; Umemura, T.; Nemoto, K.; Ogura, Y.; Nishiyama, M.; Kuzuyama, T. Discovery of Type II Polyketide Synthase-like Enzymes for the Biosynthesis of Cispentacin. *Nat Commun* **2023**, *14* (1), 8065. https://doi.org/10.1038/s41467-023-43731-z.

(397) Romo, A. J.; Shiraishi, T.; Ikeuchi, H.; Lin, G.-M.; Geng, Y.; Lee, Y.-H.; Liem, P. H.; Ma, T.; Ogasawara, Y.; Shin-ya, K.; Nishiyama, M.; Kuzuyama, T.; Liu, H. The Amipurimycin and Miharamycin Biosynthetic Gene Clusters: Unraveling the Origins of 2-Aminopurinyl Peptidyl Nucleoside Antibiotics. *J Am Chem Soc* **2019**, *141* (36), 14152–14159. https://doi.org/10.1021/jacs.9b03021.

(398) Goffeau, A.; Barrell, G.; Bussey, H.; Davis, R. W.; Dujon, B.; Feldmann, H.; Galibert, F.; Hoheisel, J. D.; Jacq, C.; Johnston, M.; Louis, E. J.; Mewes, H. W.; Murakami, Y.; Philippsen, P.; Tettelin, H.; Oliver, S. G. Life with 6000 Genes. *Science (1979)* **1996**, *274* (5287), 546–567. https://doi.org/10.1126/SCIENCE.274.5287.546.

(399) Wong, G.; Lim, L. R.; Tan, Y. Q.; Go, M. K.; Bell, D. J.; Freemont, P. S.; Yew, W. S. Reconstituting the Complete Biosynthesis of D-Lysergic Acid in Yeast. *Nat Commun* **2022**, *13* (1), 712. https://doi.org/10.1038/s41467-022-28386-6.

(400) Zhao, M.; Ma, J.; Zhang, L.; Qi, H. Engineering Strategies for Enhanced Heterologous Protein Production by Saccharomyces Cerevisiae. *Microb Cell Fact* **2024**, *23* (1), 32. https://doi.org/10.1186/s12934-024-02299-z.

(401) Werck-Reichhart, D.; Feyereisen, R. Cytochromes P450: A Success Story. *Genome Biol* **2000**, *1* (6), reviews3003.1. https://doi.org/10.1186/gb-2000-1-6-reviews3003.

(402) Corsello, M. A.; Garg, N. K. Synthetic Chemistry Fuels Interdisciplinary Approaches to the Production of Artemisinin. *Nat Prod Rep* **2015**, *32* (3), 359–366. https://doi.org/10.1039/C4NP00113C.

(403) Westfall, P. J.; Pitera, D. J.; Lenihan, J. R.; Eng, D.; Woolard, F. X.; Regentin, R.; Horning, T.; Tsuruta, H.; Melis, D. J.; Owens, A.; Fickes, S.; Diola, D.; Benjamin, K. R.; Keasling, J. D.; Leavell, M. D.; McPhee, D. J.; Renninger, N. S.; Newman, J. D.; Paddon, C. J. Production of Amorphadiene in Yeast, and Its Conversion to Dihydroartemisinic Acid, Precursor to the Antimalarial Agent Artemisinin. *Proceedings of the National Academy of Sciences* **2012**, *109* (3), E111–E118. https://doi.org/10.1073/pnas.1110740109.

(404) Zha, W.; An, T.; Li, T.; Zhu, J.; Gao, K.; Sun, Z.; Xu, W.; Lin, P.; Zi, J. Reconstruction of the Biosynthetic Pathway of Santalols under Control of the GAL Regulatory System in Yeast. *ACS Synth Biol* **2020**, *9* (2), 449–456. https://doi.org/10.1021/acssynbio.9b00479.

(405) Zha, W.; Zhang, F.; Shao, J.; Ma, X.; Zhu, J.; Sun, P.; Wu, R.; Zi, J. Rationally Engineering Santalene Synthase to Readjust the Component Ratio of Sandalwood Oil. *Nature Communications 2022 13:1* **2022**, *13* (1), 1–10. https://doi.org/10.1038/s41467-022-30294-8.

(406) Zhi, Y.; Dai, C.; Fang, X.; Xiao, X.; Lu, H.; Chen, F.; Chen, R.; Ma, W.; Deng, Z.; Lu, L.; Liu, T. Gene-Directed In Vitro Mining Uncovers the Insect-Repellent Constituent from Mugwort (*Artemisia Argyi*). *J Am Chem Soc* **2024**, *146* (45), 30883–30892. https://doi.org/10.1021/jacs.4c08857.

(407) Xu, Y.; Wang, X.; Zhang, C.; Zhou, X.; Xu, X.; Han, L.; Lv, X.; Liu, Y.; Liu, S.; Li, J.; Du, G.; Chen, J.; Ledesma-Amaro, R.; Liu, L. De Novo Biosynthesis of Rubusoside and Rebaudiosides in Engineered Yeasts. *Nature Communications 2022 13:1* **2022**, *13* (1), 1–12. https://doi.org/10.1038/s41467-022-30826-2.

(408) Park, J. H.; Bassalo, M. C.; Lin, G.-M.; Chen, Y.; Doosthosseini, H.; Schmitz, J.; Roubos, J. A.; Voigt, C. A. Design of Four Small-Molecule-Inducible Systems in the Yeast Chromosome, Applied to Optimize Terpene Biosynthesis. *ACS Synth Biol* **2023**, *12* (4), 1119–1132. https://doi.org/10.1021/acssynbio.2c00607.

(409) Thimmappa, R.; Geisler, K.; Louveau, T.; O'Maille, P.; Osbourn, A. Triterpene Biosynthesis in Plants. *Annu Rev Plant Biol* **2014**, *65* (1), 225–257. https://doi.org/10.1146/annurev-arplant-050312-120229.

(410) Liu, Y.; Zhao, X.; Gan, F.; Chen, X.; Deng, K.; Crowe, S. A.; Hudson, G. A.; Belcher, M. S.; Schmidt, M.; Astolfi, M. C. T.; Kosina, S. M.; Pang, B.; Shao, M.; Yin, J.; Sirirungruang, S.; Iavarone, A. T.; Reed, J.; Martin, L. B. B.; El-Demerdash, A.; Kikuchi, S.; Misra, R. C.; Liang, X.; Cronce, M. J.; Chen, X.; Zhan, C.; Kakumanu, R.; Baidoo, E. E. K.; Chen, Y.; Petzold, C. J.; Northen, T. R.; Osbourn, A.; Scheller, H.; Keasling, J. D. Complete Biosynthesis of QS-21 in Engineered Yeast. *Nature* **2024**, *629* (8013), 937–944. https://doi.org/10.1038/s41586-024-07345-9.

(411) Zhuang, Y.; Yang, G.-Y.; Chen, X.; Liu, Q.; Zhang, X.; Deng, Z.; Feng, Y. Biosynthesis of Plant-Derived Ginsenoside Rh2 in Yeast via Repurposing a Key Promiscuous Microbial Enzyme. *Metab Eng* **2017**, *42*, 25–32. https://doi.org/10.1016/j.ymben.2017.04.009.

(412) Yuan, W.; Jiang, C.; Wang, Q.; Fang, Y.; Wang, J.; Wang, M.; Xiao, H. Biosynthesis of Mushroom-Derived Type II Ganoderic Acids by Engineered Yeast. *Nat Commun* **2022**, *13* (1), 7740. https://doi.org/10.1038/s41467-022-35500-1.

(413) Xu, S.; Teng, X.; Li, Y. Optimization of Campesterol-Producing Yeast Strains as a Feasible Platform for the Functional Reconstitution of Plant Membrane-Bound Enzymes. *ACS Synth Biol* **2023**, *12* (4), 1109–1118. https://doi.org/10.1021/acssynbio.2c00599.

(414) Tao, H.; Lauterbach, L.; Bian, G.; Chen, R.; Hou, A.; Mori, T.; Cheng, S.; Hu, B.; Lu, L.; Mu, X.; Li, M.; Adachi, N.; Kawasaki, M.; Moriya, T.; Senda, T.; Wang, X.; Deng, Z.; Abe, I.; Dickschat, J. S.; Liu, T. Discovery of Non-Squalene Triterpenes. *Nature* **2022**, *606* (7913), 414–419. https://doi.org/10.1038/s41586-022-04773-3.

(415) Pyne, M. E.; Kevvai, K.; Grewal, P. S.; Narcross, L.; Choi, B.; Bourgeois, L.; Dueber, J. E.; Martin, V. J. J. A Yeast Platform for High-Level Synthesis of Tetrahydroisoquinoline Alkaloids. *Nat Commun* **2020**, *11* (1), 3337. https://doi.org/10.1038/s41467-020-17172-x.

(416) Srinivasan, P.; Smolke, C. D. Biosynthesis of Medicinal Tropane Alkaloids in Yeast. *Nature* **2020**, *585* (7826), 614–619. https://doi.org/10.1038/s41586-020-2650-9.

(417) Huang, J.; Huang, S. Bioproduction of Monoterpene Indole Alkaloids in a Single Cell Factory. *Engineering Microbiology* **2022**, *2* (4), 100050. https://doi.org/10.1016/j.engmic.2022.100050.

(418) Thodey, K.; Galanie, S.; Smolke, C. D. A Microbial Biomanufacturing Platform for Natural and Semisynthetic Opioids. *Nat Chem Biol* **2014**, *10* (10), 837–844. https://doi.org/10.1038/nchembio.1613.

(419) Grewal, P. S.; Samson, J. A.; Baker, J. J.; Choi, B.; Dueber, J. E. Peroxisome Compartmentalization of a Toxic Enzyme Improves Alkaloid Production. *Nature Chemical Biology 2020 17:1* **2020**, *17* (1), 96–103. https://doi.org/10.1038/s41589-020-00668-4.

(420) Srinivasan, P.; Smolke, C. D. Engineering a Microbial Biosynthesis Platform for de Novo Production of Tropane Alkaloids. *Nat Commun* **2019**, *10* (1), 3634. https://doi.org/10.1038/s41467-019-11588-w.

(421) Srinivasan, P.; Smolke, C. D. Engineering Cellular Metabolite Transport for Biosynthesis of Computationally Predicted Tropane Alkaloid Derivatives in Yeast. *Proceedings of the National Academy of Sciences* **2021**, *118* (25), e2104460118. https://doi.org/10.1073/pnas.2104460118.

(422) Brown, S.; Clastre, M.; Courdavault, V.; O'Connor, S. E. De Novo Production of the Plant-Derived Alkaloid Strictosidine in Yeast. *Proceedings of the National Academy of Sciences* **2015**, *112* (11), 3205–3210. https://doi.org/10.1073/pnas.1423555112.

(423) Zhang, J.; Hansen, L. G.; Gudich, O.; Viehrig, K.; Lassen, L. M. M.; Schrübbers, L.; Adhikari, K. B.; Rubaszka, P.; Carrasquer-Alvarez, E.; Chen, L.; D'Ambrosio, V.; Lehka, B.; Haidar, A. K.; Nallapareddy, S.; Giannakou, K.; Laloux, M.; Arsovska, D.; Jørgensen, M. A. K.; Chan, L. J. G.; Kristensen, M.; Christensen, H. B.; Sudarsan, S.; Stander, E. A.; Baidoo, E.; Petzold, C. J.; Wulff, T.; O'Connor, S. E.; Courdavault, V.; Jensen, M. K.; Keasling, J. D. A Microbial Supply Chain for Production of the Anti-Cancer Drug Vinblastine. *Nature 2022 609:7926* **2022**, *609* (7926), 341–347. https://doi.org/10.1038/s41586-022-05157-3.

(424) Liu, T.; Gou, Y.; Zhang, B.; Gao, R.; Dong, C.; Qi, M.; Jiang, L.; Ding, X.; Li, C.; Lian, J. Construction of Ajmalicine and Sanguinarine de Novo Biosynthetic Pathways Using Stable Integration Sites in Yeast. *Biotechnol Bioeng* **2022**, *119* (5), 1314–1326. https://doi.org/10.1002/bit.28040.

(425) Liu, T.; Huang, Y.; Jiang, L.; Dong, C.; Gou, Y.; Lian, J. Efficient Production of Vindoline from Tabersonine by Metabolically Engineered Saccharomyces Cerevisiae. *Commun Biol* **2021**, *4* (1), 1089. https://doi.org/10.1038/s42003-021-02617-w.

(426) Misa, J.; Billingsley, J. M.; Niwa, K.; Yu, R. K.; Tang, Y. Engineered Production of Strictosidine and Analogues in Yeast. *ACS Synth Biol* **2022**, *11* (4), 1639–1649. https://doi.org/10.1021/acssynbio.2c00037.

(427) Yee, D. A.; DeNicola, A. B.; Billingsley, J. M.; Creso, J. G.; Subrahmanyam, V.; Tang, Y. Engineered Mitochondrial Production of Monoterpenes in Saccharomyces Cerevisiae. *Metab Eng* **2019**, *55*, 76–84. https://doi.org/10.1016/j.ymben.2019.06.004.

(428) Guo, J.; Gao, D.; Lian, J.; Qu, Y. De Novo Biosynthesis of Antiarrhythmic Alkaloid Ajmaline. *Nat Commun* **2024**, *15* (1), 457. https://doi.org/10.1038/s41467-024-44797-z.

(429) Bradley, S. A.; Lehka, B. J.; Hansson, F. G.; Adhikari, K. B.; Rago, D.; Rubaszka, P.; Haidar, A. K.; Chen, L.; Hansen, L. G.; Gudich, O.; Giannakou, K.; Lengger, B.; Gill, R. T.; Nakamura, Y.; de Bernonville, T. D.; Koudounas, K.; Romero-Suarez, D.; Ding, L.; Qiao, Y.; Frimurer, T. M.; Petersen, A. A.; Besseau, S.; Kumar, S.; Gautron, N.; Melin, C.; Marc, J.; Jeanneau, R.; O'Connor, S. E.; Courdavault, V.; Keasling, J. D.; Zhang, J.; Jensen, M. K. Biosynthesis of Natural and Halogenated Plant Monoterpene Indole Alkaloids in Yeast. *Nat Chem Biol* **2023**, *19* (12), 1551–1560. https://doi.org/10.1038/s41589-023-01430-2.

(430) Schotte, C.; Jiang, Y.; Grzech, D.; Dang, T.-T. T.; Laforest, L. C.; León, F.; Mottinelli, M.; Nadakuduti, S. S.; McCurdy, C. R.; O'Connor, S. E. Directed Biosynthesis of Mitragynine Stereoisomers. *J Am Chem Soc* **2023**, *145* (9), 4957–4963. https://doi.org/10.1021/jacs.2c13644.

(431) Holtz, M.; Rago, D.; Nedermark, I.; Hansson, F. G.; Lehka, B. J.; Hansen, L. G.; Marcussen, N. E. J.; Veneman, W. J.; Ahonen, L.; Wungsintaweekul, J.; Acevedo-Rocha, C. G.; Dirks, R. P.; Zhang, J.; Keasling, J. D.; Jensen, M. K. Metabolic Engineering of Yeast for de Novo Production of Kratom Monoterpene Indole Alkaloids. *Metab Eng* **2024**, *86*, 135–146. https://doi.org/10.1016/j.ymben.2024.09.011.

(432) Florea, S.; Panaccione, D. G.; Schardl, C. L. Ergot Alkaloids of the Family Clavicipitaceae. *Phytopathology* **2017**, *107* (5), 504–518. https://doi.org/10.1094/PHYTO-12-16-0435-RVW.

(433) Wu, N.; Yao, M.; Xiao, W.; Dong, T.; Ma, H.; Du, X.; Wang, Y.; Yuan, Y. Modular Pathway Compartmentalization for Agroclavine Overproduction in *Saccharomyces Cerevisiae*. *ACS Synth Biol* **2023**, *12* (4), 1133–1145. https://doi.org/10.1021/acssynbio.2c00626.

(434) Lee, K. K. M.; Silva, N. A. Da; Kealey, J. T. Determination of the Extent of Phosphopantetheinylation of Polyketide Synthases Expressed in Escherichia Coli and Saccharomyces Cerevisiae. *Anal Biochem* **2009**, *394* (1), 75–80. https://doi.org/10.1016/j.ab.2009.07.010.

(435) Xie, X.; Meehan, M. J.; Xu, W.; Dorrestein, P. C.; Tang, Y. Acyltransferase Mediated Polyketide Release from a Fungal Megasynthase. *J Am Chem Soc* **2009**, *131* (24), 8388–8389. https://doi.org/10.1021/ja903203g.

(436) Harvey, C. J. B.; Tang, M.; Schlecht, U.; Horecka, J.; Fischer, C. R.; Lin, H.-C.; Li, J.; Naughton, B.; Cherry, J.; Miranda, M.; Li, Y. F.; Chu, A. M.; Hennessy, J. R.; Vandova, G. A.; Inglis, D.; Aiyar, R. S.; Steinmetz, L. M.; Davis, R. W.; Medema, M. H.; Sattely, E.; Khosla, C.; St. Onge, R. P.; Tang, Y.; Hillenmeyer, M. E. HEx: A Heterologous Expression Platform for the Discovery of Fungal Natural Products. *Sci Adv* **2018**, *4* (4), eaar5459. https://doi.org/10.1126/sciadv.aar5459.

(437) Li, F.; Lin, Z.; Krug, P. J.; Catrow, J. L.; Cox, J. E.; Schmidt, E. W. Animal FAS–like Polyketide Synthases Produce Diverse Polypropionates. *Proceedings of the National Academy of Sciences* **2023**, *120* (38), e2305575120. https://doi.org/10.1073/pnas.2305575120.

(438) Cooke, T. F.; Fischer, C. R.; Wu, P.; Jiang, T.-X.; Xie, K. T.; Kuo, J.; Doctorov, E.; Zehnder, A.; Khosla, C.; Chuong, C.-M.; Bustamante, C. D. Genetic Mapping and Biochemical Basis of Yellow Feather Pigmentation in Budgerigars. *Cell* **2017**, *171* (2), 427-439.e21. https://doi.org/10.1016/j.cell.2017.08.016.

(439) Awan, A. R.; Blount, B. A.; Bell, D. J.; Shaw, W. M.; Ho, J. C. H.; McKiernan, R. M.; Ellis, T. Biosynthesis of the Antibiotic Nonribosomal Peptide Penicillin in Baker's Yeast. *Nat Commun* **2017**, *8* (1), 15202. https://doi.org/10.1038/ncomms15202.

(440) Valliere, M. A.; Korman, T. P.; Woodall, N. B.; Khitrov, G. A.; Taylor, R. E.; Baker, D.; Bowie, J. U. A Cell-Free Platform for the Prenylation of Natural Products and Application to Cannabinoid Production. *Nat Commun* **2019**, *10* (1), 565. https://doi.org/10.1038/s41467-019-08448-y.

(441) Zhang, Y.; Guo, J.; Gao, P.; Yan, W.; Shen, J.; Luo, X.; Keasling, J. D. Development of an Efficient Yeast Platform for Cannabigerolic Acid Biosynthesis. *Metab Eng* **2023**, *80*, 232–240. https://doi.org/10.1016/j.ymben.2023.10.004.

(442) Pandey, R. P.; Parajuli, P.; Koffas, M. A. G.; Sohng, J. K. Microbial Production of Natural and Non-Natural Flavonoids: Pathway Engineering, Directed Evolution and Systems/Synthetic Biology. *Biotechnol Adv* **2016**, *34* (5), 634–662. https://doi.org/10.1016/j.biotechadv.2016.02.012.

(443) Isogai, S.; Tominaga, M.; Kondo, A.; Ishii, J. Plant Flavonoid Production in Bacteria and Yeasts. *Frontiers in Chemical Engineering* **2022**, *4*, 880694. https://doi.org/10.3389/fceng.2022.880694.

(444) Delmulle, T.; De Maeseneire, S. L.; De Mey, M. Challenges in the Microbial Production of Flavonoids. *Phytochemistry Reviews* **2018**, *17* (2), 229–247. https://doi.org/10.1007/s11101-017-9515-3.

(445) Yang, S.; Chen, R.; Cao, X.; Wang, G.; Zhou, Y. J. De Novo Biosynthesis of the Hops Bioactive Flavonoid Xanthohumol in Yeast. *Nature Communications 2024 15:1* **2024**, *15* (1), 1–10. https://doi.org/10.1038/s41467-023-44654-5.

(446) Xiao, Z.; Wang, Y.; Liu, J.; Zhang, S.; Tan, X.; Zhao, Y.; Mao, J.; Jiang, N.; Zhou, J.; Shan, Y. Systematic Engineering of *Saccharomyces Cerevisiae* Chassis for Efficient Flavonoid-7- *O* -Disaccharide Biosynthesis. *ACS Synth Biol* **2023**, *12* (9), 2740–2749. https://doi.org/10.1021/acssynbio.3c00348.

(447) Liu, Q.; Yu, T.; Li, X.; Chen, Y.; Campbell, K.; Nielsen, J.; Chen, Y. Rewiring Carbon Metabolism in Yeast for High Level Production of Aromatic Chemicals. *Nature Communications 2019 10:1* **2019**, *10* (1), 1–13. https://doi.org/10.1038/s41467-019-12961-5.

(448) Jia, Z.-C.; Liu, D.; Ma, H.-D.; Cui, Y.-H.; Li, H.-M.; Li, X.; Yuan, Y.-J. Yeast Metabolic Engineering for Biosynthesis of Caffeic Acid-Derived Phenethyl Ester and Phenethyl Amide. *ACS Synth Biol* **2023**, *12* (12), 3635–3645. https://doi.org/10.1021/acssynbio.3c00413.

(449) Chen, R.; Gao, J.; Yu, W.; Chen, X.; Zhai, X.; Chen, Y.; Zhang, L.; Zhou, Y. J. Engineering Cofactor Supply and Recycling to Drive Phenolic Acid Biosynthesis in Yeast. *Nature Chemical Biology 2022 18:5* **2022**, *18* (5), 520–529. https://doi.org/10.1038/s41589-022-01014-6.

(450) Babaei, M.; Thomsen, P. T.; Pastor, M. C.; Jensen, M. K.; Borodina, I. Coupling High-Throughput and Targeted Screening for Identification of Nonobvious Metabolic Engineering Targets. *ACS Synth Biol* **2024**, *13* (1), 168–182. https://doi.org/10.1021/acssynbio.3c00396.

(451) Wang, P.; Fan, Z.; Wei, W.; Yang, C.; Wang, Y.; Shen, X.; Yan, X.; Zhou, Z. Biosynthesis of the Plant Coumarin Osthole by Engineered *Saccharomyces Cerevisiae*. *ACS Synth Biol* **2023**, *12* (8), 2455–2462. https://doi.org/10.1021/acssynbio.3c00321.

(452) Wang, Z.; Zhou, Y.; Wang, Y.; Yan, X. Reconstitution and Optimization of the Marmesin Biosynthetic Pathway in Yeast. *ACS Synth Biol* **2023**, *12* (10), 2922–2933. https://doi.org/10.1021/acssynbio.3c00267.

(453) Peng, H.; Chen, R.; Shaw, W. M.; Hapeta, P.; Jiang, W.; Bell, D. J.; Ellis, T.; Ledesma-Amaro, R. Modular Metabolic Engineering and Synthetic Coculture Strategies for the Production of Aromatic Compounds in Yeast. *ACS Synth Biol* **2023**, *12* (6), 1739–1749. https://doi.org/10.1021/acssynbio.3c00047.

(454) Zhang, G.; Wang, H.; Zhang, Z.; Verstrepen, K. J.; Wang, Q.; Dai, Z. Metabolic Engineering of Yarrowia Lipolytica for Terpenoids Production: Advances and Perspectives. *Crit Rev Biotechnol* **2022**, *42* (4), 618–633. https://doi.org/10.1080/07388551.2021.1947183.

(455) Sáez-Sáez, J.; Wang, G.; Marella, E. R.; Sudarsan, S.; Cernuda Pastor, M.; Borodina, I. Engineering the Oleaginous Yeast Yarrowia Lipolytica for High-Level Resveratrol Production. *Metab Eng* **2020**, *62*, 51–61. https://doi.org/10.1016/j.ymben.2020.08.009.

(456) Ma, Y.; Liu, N.; Greisen, P.; Li, J.; Qiao, K.; Huang, S.; Stephanopoulos, G. Removal of Lycopene Substrate Inhibition Enables High Carotenoid Productivity in Yarrowia Lipolytica. *Nature Communications 2022 13:1* **2022**, *13* (1), 1–11. https://doi.org/10.1038/s41467-022-28277-w.

(457) Ganesan, V.; Monteiro, L.; Pedada, D.; Stohr, A.; Blenner, M. High-Efficiency Multiplexed Cytosine Base Editors for Natural Product Synthesis in *Yarrowia Lipolytica*. *ACS Synth Biol* **2023**, *12* (10), 3082–3091. https://doi.org/10.1021/acssynbio.3c00435.

(458) Song, Z.; Lin, W.; Duan, X.; Song, L.; Wang, C.; Yang, H.; Lu, X.; Ji, X.; Tian, Y.; Liu, H. Increased Cordycepin Production in *Yarrowia Lipolytica* Using Combinatorial Metabolic Engineering Strategies. *ACS Synth Biol* **2023**, *12* (3), 780–787. https://doi.org/10.1021/acssynbio.2c00570.

(459) Peña, D. A.; Gasser, B.; Zanghellini, J.; Steiger, M. G.; Mattanovich, D. Metabolic Engineering of Pichia Pastoris. **2018**. https://doi.org/10.1016/j.ymben.2018.04.017.

(460) Zhang, X.; Wang, D.; Duan, Y.; Zheng, X.; Lin, Y.; Liang, S. Production of Lycopene by Metabolically Engineered Pichia Pastoris. *Biosci Biotechnol Biochem* **2020**, *84* (3), 463–470. https://doi.org/10.1080/09168451.2019.1693250.

(461) Gao, J.; Xu, J.; Zuo, Y.; Ye, C.; Jiang, L.; Feng, L.; Huang, L.; Xu, Z.; Lian, J. Synthetic Biology Toolkit for Marker-Less Integration of Multigene Pathways into *Pichia Pastoris* via CRISPR/Cas9. *ACS Synth Biol* **2022**, *11* (2), 623–633. https://doi.org/10.1021/acssynbio.1c00307.

(462) Muzafar, S.; Sharma, R. D.; Chauhan, N.; Prasad, R. Intron Distribution and Emerging Role of Alternative Splicing in Fungi. *FEMS Microbiol Lett* **2021**, *368* (19). https://doi.org/10.1093/femsle/fnab135.

(463) Nielsen, M. R.; Wollenberg, R. D.; Westphal, K. R.; Sondergaard, T. E.; Wimmer, R.; Gardiner, D. M.; Sørensen, J. L. Heterologous Expression of Intact Biosynthetic Gene Clusters in Fusarium Graminearum. *Fungal Genetics and Biology* **2019**, *132*, 103248. https://doi.org/10.1016/j.fgb.2019.103248.

(464) Liu, N.; Hung, Y.-S.; Gao, S.-S.; Hang, L.; Zou, Y.; Chooi, Y.-H.; Tang, Y. Identification and Heterologous Production of a Benzoyl-Primed Tricarboxylic Acid Polyketide Intermediate from the Zaragozic Acid A Biosynthetic Pathway. *Org Lett* **2017**, *19* (13), 3560–3563. https://doi.org/10.1021/acs.orglett.7b01534.

(465) Yee, D. A.; Tang, Y. Investigating Fungal Biosynthetic Pathways Using Heterologous Gene Expression: Aspergillus Nidulans as a Heterologous Host. In *Methods in Molecular Biology*; Humana Press Inc., 2022; Vol. 2489, pp 41–52. https://doi.org/10.1007/978-1-0716-2273-5_3.

(466) Chiang, Y.; Ahuja, M.; Oakley, C. E.; Entwistle, R.; Asokan, A.; Zutz, C.; Wang, C. C. C.; Oakley, B. R. Development of Genetic Dereplication Strains in *Aspergillus Nidulans* Results in the Discovery of Aspercryptin. *Angewandte Chemie International Edition* **2016**, *55* (5), 1662–1665. https://doi.org/10.1002/anie.201507097.

(467) Unkles, S. E.; Valiante, V.; Mattern, D. J.; Brakhage, A. A. Synthetic Biology Tools for Bioprospecting of Natural Products in Eukaryotes. *Chem Biol* **2014**, *21* (4), 502–508. https://doi.org/10.1016/j.chembiol.2014.02.010.

(468) Milne, N.; Thomsen, P.; Knudsen, N.; Rubaszka, P.; Kristensen, M.; Borodina, I. Metabolic Engineering of Saccharomyces Cerevisiae for the de Novo Production of Psilocybin and Related Tryptamine Derivatives. *Metab Eng* **2020**, *60*, 25–36. https://doi.org/10.1016/j.ymben.2019.12.007.

(469) Adams, A. M.; Kaplan, N. A.; Wei, Z.; Brinton, J. D.; Monnier, C. S.; Enacopol, A. L.; Ramelot, T. A.; Jones, J. A. In Vivo Production of Psilocybin in E. Coli. *Metab Eng* **2019**, *56*, 111–119. https://doi.org/10.1016/j.ymben.2019.09.009.

(470) Mattern, D. J.; Valiante, V.; Horn, F.; Petzke, L.; Brakhage, A. A. Rewiring of the Austinoid Biosynthetic Pathway in Filamentous Fungi. *ACS Chem Biol* **2017**, *12* (12), 2927–2933. https://doi.org/10.1021/acschembio.7b00814.

(471) Jarczynska, Z. D.; Garcia Vanegas, K.; Deichmann, M.; Nørskov Jensen, C.; Scheeper, M. J.; Futyma, M. E.; Strucko, T.; Jares Contesini, F.; Sparholt Jørgensen, T.; Blæsbjerg Hoof, J.; Hasbro Mortensen, U. A Versatile in Vivo DNA Assembly Toolbox for Fungal Strain Engineering. *ACS Synth Biol* **2022**, *11* (10), 3251–3263. https://doi.org/10.1021/acssynbio.2c00159.

(472) Bok, J. W.; Ye, R.; Clevenger, K. D.; Mead, D.; Wagner, M.; Krerowicz, A.; Albright, J. C.; Goering, A. W.; Thomas, P. M.; Kelleher, N. L.; Keller, N. P.; Wu, C. C. Fungal Artificial Chromosomes for Mining of the Fungal Secondary Metabolome. *BMC Genomics* **2015**, *16* (1), 343. https://doi.org/10.1186/s12864-015-1561-x.

(473) Park, S. C.; Steffan, B. N.; Yun Lim, F.; Gupta, R.; Ayaloglu Butun, F.; Chen, H.; Ye, R.; Decker, T.; Wu, C. C.; Kelleher, N. L.; Woo Bok, J.; Keller, N. P. Terpenoid Balance in *Aspergillus Nidulans* Unveiled by Heterologous Squalene Synthase Expression. *Sci Adv* **2024**, *10* (8), 7416. https://doi.org/10.1126/sciadv.adk7416.

(474) Caesar, L. K.; Butun, F. A.; Robey, M. T.; Ayon, N. J.; Gupta, R.; Dainko, D.; Bok, J. W.; Nickles, G.; Stankey, R. J.; Johnson, D.; Mead, D.; Cank, K. B.; Earp, C. E.; Raja, H. A.; Oberlies, N. H.; Keller, N. P.; Kelleher, N. L. Correlative Metabologenomics of 110 Fungi Reveals Metabolite–Gene Cluster Pairs. *Nat Chem Biol* **2023**, *19* (7), 846–854. https://doi.org/10.1038/s41589-023-01276-8.

(475) Li, H.; Hu, J.; Wei, H.; Solomon, P. S.; Vuong, D.; Lacey, E.; Stubbs, K. A.; Piggott, A. M.; Chooi, Y.-H. Chemical Ecogenomics-Guided Discovery of Phytotoxic α-Pyrones from the Fungal Wheat Pathogen *Parastagonospora Nodorum*. *Org Lett* **2018**, *20* (19), 6148–6152. https://doi.org/10.1021/acs.orglett.8b02617.

(476) Li, H.; Gilchrist, C. L. M.; Phan, C.-S.; Lacey, H. J.; Vuong, D.; Moggach, S. A.; Lacey, E.; Piggott, A. M.; Chooi, Y.-H. Biosynthesis of a New Benzazepine Alkaloid Nanangelenin A from *Aspergillus Nanangensis* Involves an Unusual <scp>l</Scp> -Kynurenine-Incorporating NRPS Catalyzing Regioselective Lactamization. *J Am Chem Soc* **2020**, *142* (15), 7145–7152. https://doi.org/10.1021/jacs.0c01605.

(477) Hu, J.; Sarrami, F.; Li, H.; Zhang, G.; Stubbs, K. A.; Lacey, E.; Stewart, S. G.; Karton, A.; Piggott, A. M.; Chooi, Y.-H. Heterologous Biosynthesis of Elsinochrome A Sheds Light on the Formation of the Photosensitive Perylenequinone System. *Chem Sci* **2019**, *10* (5), 1457–1465. https://doi.org/10.1039/C8SC02870B.

(478) Li, H.; Wei, H.; Hu, J.; Lacey, E.; Sobolev, A. N.; Stubbs, K. A.; Solomon, P. S.; Chooi, Y.-H. Genomics-Driven Discovery of Phytotoxic Cytochalasans Involved in the Virulence of the Wheat Pathogen *Parastagonospora Nodorum*. *ACS Chem Biol* **2020**, *15* (1), 226–233. https://doi.org/10.1021/acschembio.9b00791.

(479) Clevenger, K. D.; Bok, J. W.; Ye, R.; Miley, G. P.; Verdan, M. H.; Velk, T.; Chen, C.; Yang, K.; Robey, M. T.; Gao, P.; Lamprecht, M.; Thomas, P. M.; Islam, M. N.; Palmer, J. M.; Wu, C. C.; Keller, N. P.; Kelleher, N. L. A Scalable Platform

to Identify Fungal Secondary Metabolites and Their Gene Clusters. *Nat Chem Biol* **2017**, *13* (8), 895–901. https://doi.org/10.1038/nchembio.2408.

(480) Clevenger, K. D.; Ye, R.; Bok, J. W.; Thomas, P. M.; Islam, M. N.; Miley, G. P.; Robey, M. T.; Chen, C.; Yang, K.; Swyers, M.; Wu, E.; Gao, P.; Wu, C. C.; Keller, N. P.; Kelleher, N. L. Interrogation of Benzomalvin Biosynthesis Using Fungal Artificial Chromosomes with Metabolomic Scoring (FAC-MS): Discovery of a Benzodiazepine Synthase Activity. *Biochemistry* **2018**, *57* (23), 3237–3243. https://doi.org/10.1021/acs.biochem.8b00076.

(481) Robey, M. T.; Ye, R.; Bok, J. W.; Clevenger, K. D.; Islam, M. N.; Chen, C.; Gupta, R.; Swyers, M.; Wu, E.; Gao, P.; Thomas, P. M.; Wu, C. C.; Keller, N. P.; Kelleher, N. L. Identification of the First Diketomorpholine Biosynthetic Pathway Using FAC-MS Technology. *ACS Chem Biol* **2018**, *13* (5), 1142–1147. https://doi.org/10.1021/acschembio.8b00024.

(482) Hansen, B. G.; Salomonsen, B.; Nielsen, M. T.; Nielsen, J. B.; Hansen, N. B.; Nielsen, K. F.; Regueira, T. B.; Nielsen, J.; Patil, K. R.; Mortensen, U. H. Versatile Enzyme Expression and Characterization System for Aspergillus Nidulans, with the Penicillium Brevicompactum Polyketide Synthase Gene from the Mycophenolic Acid Gene Cluster as a Test Case. *Appl Environ Microbiol* **2011**, *77* (9), 3044–3051. https://doi.org/10.1128/AEM.01768-10.

(483) Vanegas, K. G.; Jarczynska, Z. D.; Strucko, T.; Mortensen, U. H. Cpf1 Enables Fast and Efficient Genome Editing in Aspergilli. *Fungal Biol Biotechnol* **2019**, *6* (1), 6. https://doi.org/10.1186/s40694-019-0069-6.

(484) Yang, X.; Feng, P.; Yin, Y.; Bushley, K.; Spatafora, J. W.; Wang, C. Cyclosporine Biosynthesis in *Tolypocladium Inflatum* Benefits Fungal Adaptation to the Environment. *mBio* **2018**, *9* (5). https://doi.org/10.1128/mBio.01211-18.

(485) Lee, C.; Chen, L.; Chiang, C.; Lai, C.; Lin, H. The Biosynthesis of Norsesquiterpene Aculenes Requires Three Cytochrome P450 Enzymes to Catalyze a Stepwise Demethylation Process. *Angewandte Chemie* **2019**, *131* (51), 18585–18589. https://doi.org/10.1002/ange.201910200.

(486) Yee, D. A.; Kakule, T. B.; Cheng, W.; Chen, M.; Chong, C. T. Y.; Hai, Y.; Hang, L. F.; Hung, Y.-S.; Liu, N.; Ohashi, M.; Okorafor, I. C.; Song, Y.; Tang, M.; Zhang, Z.; Tang, Y. Genome Mining of Alkaloidal Terpenoids from a Hybrid Terpene and Nonribosomal Peptide Biosynthetic Pathway. *J Am Chem Soc* **2020**, *142* (2), 710–714. https://doi.org/10.1021/jacs.9b13046.

(487) Cheng, W.; Chen, M.; Ohashi, M.; Tang, Y. Biosynthesis of Terpenoid–Pyrrolobenzoxazine Hybrid Natural Product CJ-12662. *Angewandte Chemie International Edition* **2022**, *61* (12), e202116928. https://doi.org/10.1002/anie.202116928.

(488) Yan, D.; Wang, K.; Bai, S.; Liu, B.; Bai, J.; Qi, X.; Hu, Y. Flavin-Dependent Monooxygenase-Mediated 1,2-Oxazine Construction *via* Meisenheimer Rearrangement in the Biosynthesis of Paeciloxazine. *J Am Chem Soc* **2022**, *144* (9), 4269–4276. https://doi.org/10.1021/jacs.2c00881.

(489) Zhang, J.-M.; Wang, H.-H.; Liu, X.; Hu, C.-H.; Zou, Y. Heterologous and Engineered Biosynthesis of Nematocidal Polyketide–Nonribosomal Peptide Hybrid Macrolactone from Extreme Thermophilic Fungi. *J Am Chem Soc* **2020**, *142* (4), 1957–1965. https://doi.org/10.1021/jacs.9b11410.

(490) Zhang, Y.; Wu, L.; Kerr, T. A.; Garg, N. K.; Tang, Y. Fragment-Guided Genome Mining of Octacyclic Cyclophane Alkaloids from Fungi. *J Am Chem Soc* **2024**, *146* (34), 23933–23942. https://doi.org/10.1021/jacs.4c06734.

(491) Liu, M.; Ohashi, M.; Hung, Y.-S.; Scherlach, K.; Watanabe, K.; Hertweck, C.; Tang, Y. AoiQ Catalyzes Geminal Dichlorination of 1,3-Diketone Natural Products. *J Am Chem Soc* **2021**, *143* (19), 7267–7271. https://doi.org/10.1021/jacs.1c02868.

(492) Chankhamjon, P.; Tsunematsu, Y.; Ishida-Ito, M.; Sasa, Y.; Meyer, F.; Boettger-Schmidt, D.; Urbansky, B.; Menzel, K.; Scherlach, K.; Watanabe, K.; Hertweck, C. Regioselective Dichlorination of a Non-Activated Aliphatic Carbon Atom and Phenolic Bismethylation by a Multifunctional Fungal Flavoenzyme. *Angewandte Chemie International Edition* **2016**, *55* (39), 11955–11959. https://doi.org/10.1002/anie.201604516.

(493) Tohyama, S.; Tomura, A.; Ikeda, N.; Hatano, M.; Odanaka, J.; Kubota, Y.; Umekita, M.; Igarashi, M.; Sawa, R.; Morino, T. Discovery and Characterization of NK13650s, Naturally Occurring P300-Selective Histone Acetyltransferase Inhibitors. *J Org Chem* **2012**, *77* (20), 9044–9052. https://doi.org/10.1021/jo301534b.

(494) Zhang, Z.; Tamura, Y.; Tang, M.; Qiao, T.; Sato, M.; Otsu, Y.; Sasamura, S.; Taniguchi, M.; Watanabe, K.; Tang, Y. Biosynthesis of the Immunosuppressant (-)-FR901483. *J Am Chem Soc* **2021**, *143* (1), 132–136. https://doi.org/10.1021/jacs.0c12352.

(495) Chipman, D.; Barak, Z.; Schloss, J. V. Biosynthesis of 2-Aceto-2-Hydroxy Acids: Acetolactate Synthases and Acetohydroxyacid Synthases. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1998**, *1385* (2), 401–419. https://doi.org/10.1016/S0167-4838(98)00083-1.

(496) Shang, Z.; Arishi, A. A.; Wu, C.; Lao, F.; Gilchrist, C. L. M.; Moggach, S. A.; Lacey, E.; Piggott, A. M.; Chooi, Y. Self-Resistance Gene-Guided Discovery of the Molecular Basis for Biosynthesis of the Fatty Acid Synthase Inhibitor Cerulenin. *Angewandte Chemie International Edition* **2024**, e202414941. https://doi.org/10.1002/anie.202414941.

(497) Bat-Erdene, U.; Kanayama, D.; Tan, D.; Turner, W. C.; Houk, K. N.; Ohashi, M.; Tang, Y. Iterative Catalysis in the Biosynthesis of Mitochondrial Complex II Inhibitors Harzianopyridone and Atpenin B. *J Am Chem Soc* **2020**, *142* (19), 8550–8554. https://doi.org/10.1021/jacs.0c03438.

(498) Zhang, M.; Sun, C.; Liu, Y.; Feng, H.; Chang, H.; Cao, S.; Li, G.; Yang, S.; Hou, J.; Zhu-Salzman, K.; Zhang, H.; Qin, Q. Transcriptome Analysis and Functional Validation Reveal a Novel Gene, *BcCGF1* , That Enhances Fungal Virulence by Promoting Infection-related Development and Host Penetration. *Mol Plant Pathol* **2020**, *21* (6), 834–853. https://doi.org/10.1111/mpp.12934.

(499) Mapuranga, J.; Chang, J.; Zhang, L.; Zhang, N.; Yang, W. Fungal Secondary Metabolites and Small RNAs Enhance Pathogenicity during Plant-Fungal Pathogen Interactions. *Journal of Fungi* **2022**, *9* (1), 4. https://doi.org/10.3390/jof9010004.

(500) Zhu, L.; Tang, L.; Tian, X.; Bai, Y.; Huang, L. Two Polyketide Synthase Genes, *VpPKS10* and *VpPKS33* , Regulated by VpLaeA Are Essential to the Virulence of *Valsa Pyri*. *Phytopathology* **2024**, *114* (9), 2071–2083. https://doi.org/10.1094/PHYTO-12-23-0498-R.

(501) Ugai, T.; Minami, A.; Fujii, R.; Tanaka, M.; Oguri, H.; Gomi, K.; Oikawa, H. Heterologous Expression of Highly Reducing Polyketide Synthase Involved in Betaenone Biosynthesis. *Chemical Communications* **2015**, *51* (10), 1878–1881. https://doi.org/10.1039/C4CC09512J.

(502) Li, H.; Hu, J.; Wei, H.; Solomon, P. S.; Stubbs, K. A.; Chooi, Y. Biosynthesis of a Tricyclo[6.2.2.0 ^{2,7}]Dodecane System by a Berberine Bridge Enzyme-Like Aldolase. *Chemistry – A European Journal* **2019**, *25* (66), 15062–15066. https://doi.org/10.1002/chem.201904360.

(503) Arishi, A. A.; Shang, Z.; Lacey, E.; Crombie, A.; Vuong, D.; Li, H.; Bracegirdle, J.; Turner, P.; Lewis, W.; Flematti, G. R.; Piggott, A. M.; Chooi, Y.-H. Discovery and Heterologous Biosynthesis of Glycosylated Polyketide Luteodienoside A Reveals Unprecedented Glucinol-Mediated Product Offloading by a Fungal Carnitine *O* -Acyltransferase Domain. *Chem Sci* **2024**, *15* (9), 3349–3356. https://doi.org/10.1039/D3SC05008D.

(504) Martín, J. F. Transport Systems, Intracellular Traffic of Intermediates and Secretion of β-Lactam Antibiotics in Fungi. *Fungal Biology and Biotechnology 2020 7:1* **2020**, *7* (1), 1–15. https://doi.org/10.1186/S40694-020-00096-Y.

(505) Meijer, W. H.; Gidijala, L.; Fekken, S.; Kiel, J. A. K. W.; van den Berg, M. A.; Lascaris, R.; Bovenberg, R. A. L.; van der Klei, I. J. Peroxisomes Are Required for Efficient Penicillin Biosynthesis in *Penicillium Chrysogenum*. *Appl Environ Microbiol* **2010**, *76* (17), 5702–5709. https://doi.org/10.1128/AEM.02327-09.

(506) Kiel, J. A. K. W.; van den Berg, M. A.; Fusetti, F.; Poolman, B.; Bovenberg, R. A. L.; Veenhuis, M.; van der Klei, I. J. Matching the Proteome to the Genome: The Microbody of Penicillin-Producing Penicillium Chrysogenum Cells. *Funct Integr Genomics* **2009**, *9* (2), 167–184. https://doi.org/10.1007/s10142-009-0110-6.

(507) Herr, A.; Fischer, R. Improvement of Aspergillus Nidulans Penicillin Production by Targeting AcvA to Peroxisomes. *Metab Eng* **2014**, *25*, 131–139. https://doi.org/10.1016/j.ymben.2014.07.002.

(508) Jamieson, C. S.; Ohashi, M.; Liu, F.; Tang, Y.; Houk, K. N. The Expanding World of Biosynthetic Pericyclases: Cooperation of Experiment and Theory for Discovery. *Nat Prod Rep* **2019**, *36* (5), 698–713. https://doi.org/10.1039/C8NP00075A.

(509) Dan, Q.; Newmister, S. A.; Klas, K. R.; Fraley, A. E.; McAfoos, T. J.; Somoza, A. D.; Sunderhaus, J. D.; Ye, Y.; Shende, V. V.; Yu, F.; Sanders, J. N.; Brown, W. C.; Zhao, L.; Paton, R. S.; Houk, K. N.; Smith, J. L.; Sherman, D. H.; Williams, R. M. Fungal Indole Alkaloid Biogenesis through Evolution of a Bifunctional Reductase/Diels–Alderase. *Nat Chem* **2019**, *11* (11), 972–980. https://doi.org/10.1038/s41557-019-0326-6.

(510) Kim, L. J.; Ohashi, M.; Zhang, Z.; Tan, D.; Asay, M.; Cascio, D.; Rodriguez, J. A.; Tang, Y.; Nelson, H. M. Prospecting for Natural Products by Genome Mining and Microcrystal Electron Diffraction. *Nat Chem Biol* **2021**, *17* (8), 872–877. https://doi.org/10.1038/s41589-021-00834-2.

(511) Ohashi, M.; Kakule, T. B.; Tang, M.-C.; Jamieson, C. S.; Liu, M.; Zhao, Y.-L.; Houk, K. N.; Tang, Y. Biosynthesis of *Para*-Cyclophane-Containing Hirsutellone Family of Fungal Natural Products. *J Am Chem Soc* **2021**, *143* (15), 5605–5609. https://doi.org/10.1021/jacs.1c00098.

(512) Zhang, Z.; Jamieson, C. S.; Zhao, Y.-L.; Li, D.; Ohashi, M.; Houk, K. N.; Tang, Y. Enzyme-Catalyzed Inverse-Electron Demand Diels–Alder Reaction in the Biosynthesis of Antifungal Ilicicolin H. *J Am Chem Soc* **2019**, *141* (14), 5659–5663. https://doi.org/10.1021/jacs.9b02204.

(513) Ohashi, M.; Jamieson, C. S.; Cai, Y.; Tan, D.; Kanayama, D.; Tang, M.-C.; Anthony, S. M.; Chari, J. V.; Barber, J. S.; Picazo, E.; Kakule, T. B.; Cao, S.; Garg, N. K.; Zhou, J.; Houk, K. N.; Tang, Y. An Enzymatic Alder-Ene Reaction. *Nature* **2020**, *586* (7827), 64–69. https://doi.org/10.1038/s41586-020-2743-5.

(514) Ohashi, M.; Liu, F.; Hai, Y.; Chen, M.; Tang, M.; Yang, Z.; Sato, M.; Watanabe, K.; Houk, K. N.; Tang, Y. SAM-Dependent Enzyme-Catalysed Pericyclic Reactions in Natural Product Biosynthesis. *Nature* **2017**, *549* (7673), 502–506. https://doi.org/10.1038/nature23882.

(515) Li, L.; Yu, P.; Tang, M.-C.; Zou, Y.; Gao, S.-S.; Hung, Y.-S.; Zhao, M.; Watanabe, K.; Houk, K. N.; Tang, Y. Biochemical Characterization of a Eukaryotic Decalin-Forming Diels–Alderase. *J Am Chem Soc* **2016**, *138* (49), 15837–15840. https://doi.org/10.1021/jacs.6b10452.

(516) Tan, D.; Jamieson, C. S.; Ohashi, M.; Tang, M.-C.; Houk, K. N.; Tang, Y. Genome-Mined Diels–Alderase Catalyzes Formation of the *Cis* -Octahydrodecalins of Varicidin A and B. *J Am Chem Soc* **2019**, *141* (2), 769–773. https://doi.org/10.1021/jacs.8b12010.

(517) Chen, Q.; Yuan, G.; Yuan, T.; Zeng, H.; Zou, Z.-R.; Tu, Z.; Gao, J.; Zou, Y. Set of Cytochrome P450s Cooperatively Catalyzes the Synthesis of a Highly Oxidized and Rearranged Diterpene-Class Sordarinane Architecture. *J Am Chem Soc* **2022**, *144* (8), 3580–3589. https://doi.org/10.1021/jacs.1c12427.

(518) Sun, Z.; Jamieson, C. S.; Ohashi, M.; Houk, K. N.; Tang, Y. Discovery and Characterization of a Terpene Biosynthetic Pathway Featuring a Norbornene-Forming Diels-Alderase. *Nat Commun* **2022**, *13* (1), 2568. https://doi.org/10.1038/s41467-022-30288-6.

(519) Niwa, K.; Ohashi, M.; Xie, K.; Chiang, C.-Y.; Jamieson, C. S.; Sato, M.; Watanabe, K.; Liu, F.; Houk, K. N.; Tang, Y. Biosynthesis of Polycyclic Natural Products from Conjugated Polyenes via Tandem Isomerization and Pericyclic Reactions. *J Am Chem Soc* **2023**, *145* (25), 13520–13525. https://doi.org/10.1021/jacs.3c02380.

(520) Liu, J.; Lu, J.; Zhang, C.; Zhou, Q.; Jamieson, C. S.; Shang, C.; Houk, K. N.; Zhou, J.; Hu, Y. Tandem Intermolecular [4 + 2] Cycloadditions Are Catalysed by Glycosylated Enzymes for Natural Product Biosynthesis. *Nat Chem* **2023**, *15* (8), 1083–1090. https://doi.org/10.1038/s41557-023-01260-8.

(521) Guggenheim, S.; Flavin, M. Cystathionine γ-Synthase. *Journal of Biological Chemistry* **1969**, *244* (22), 6217–6227. https://doi.org/10.1016/S0021-9258(18)63527-6.

(522) Cui, Z.; Overbay, J.; Wang, X.; Liu, X.; Zhang, Y.; Bhardwaj, M.; Lemke, A.; Wiegmann, D.; Niro, G.; Thorson, J. S.; Ducho, C.; Van Lanen, S. G. Pyridoxal-5′-Phosphate-Dependent Alkyl Transfer in Nucleoside Antibiotic Biosynthesis. *Nat Chem Biol* **2020**, *16* (8), 904–911. https://doi.org/10.1038/s41589-020-0548-3.

(523) Chen, M.; Liu, C.-T.; Tang, Y. Discovery and Biocatalytic Application of a PLP-Dependent Amino Acid γ-Substitution Enzyme That Catalyzes C–C Bond Formation. *J Am Chem Soc* **2020**, *142* (23), 10506–10515. https://doi.org/10.1021/jacs.0c03535.

(524) Gao, J.; Liu, S.; Zhou, C.; Lara, D.; Zou, Y.; Hai, Y. A Pyridoxal 5′-Phosphate-Dependent Mannich Cyclase. *Nat Catal* **2023**, *6* (6), 476–486. https://doi.org/10.1038/s41929-023-00963-y.

(525) Hai, Y.; Chen, M.; Huang, A.; Tang, Y. Biosynthesis of Mycotoxin Fusaric Acid and Application of a PLP-Dependent Enzyme for Chemoenzymatic Synthesis of Substituted <scp>l</Scp> -Pipecolic Acids. *J Am Chem Soc* **2020**, *142* (46), 19668–19677. https://doi.org/10.1021/jacs.0c09352.

(526) Zhang, J.-M.; Liu, X.; Wei, Q.; Ma, C.; Li, D.; Zou, Y. Berberine Bridge Enzyme-like Oxidase-Catalysed Double Bond Isomerization Acts as the Pathway Switch in Cytochalasin Synthesis. *Nat Commun* **2022**, *13* (1), 225. https://doi.org/10.1038/s41467-021-27931-z.

(527) Liu, S.; Nie, Q.; Liu, Z.; Patil, S.; Gao, X. Fungal P450 Deconstructs the 2,5-Diazabicyclo[2.2.2]Octane Ring *En Route* to the Complete Biosynthesis of 21 *R* -Citrinadin A. *J Am Chem Soc* **2023**, *145* (26), 14251–14259. https://doi.org/10.1021/jacs.3c02109.

(528) Li, H.; Shu, S.; Kalaitzis, J. A.; Shang, Z.; Vuong, D.; Crombie, A.; Lacey, E.; Piggott, A. M.; Chooi, Y.-H. Genome Mining of *Aspergillus Hancockii* Unearths Cryptic Polyketide Hancockinone A Featuring a Prenylated 6/6/6/5 Carbocyclic Skeleton. *Org Lett* **2021**, *23* (22), 8789–8793. https://doi.org/10.1021/acs.orglett.1c03283.

(529) Chen, X.-W.; Rao, L.; Chen, J.-L.; Zou, Y. Unexpected Assembly Machinery for 4(3H)-Quinazolinone Scaffold Synthesis. *Nat Commun* **2022**, *13* (1), 6522. https://doi.org/10.1038/s41467-022-34340-3.

(530) Wei, G.; Shu, X.; Yao, Y.; Zhang, K.; Zhang, J.; Gao, S.-S. Heterologous Production of Unnatural Flavipucine Family Products Provides Insights into Flavipucines Biosynthesis. *Org Lett* **2021**, *23* (20), 7708–7712. https://doi.org/10.1021/acs.orglett.1c02566.

(531) Chen, X.-W.; Liu, Z.; Dai, S.; Zou, Y. Discovery, Characterization and Engineering of the Free <scp>l</Scp> -Histidine *C4* -Prenyltransferase. *J Am Chem Soc* **2024**, *146* (34), 23686–23691. https://doi.org/10.1021/jacs.4c08388.

(532) Xu, R.; Zou, Y. Biosynthesis of (-)-Vinigrol. *Angewandte Chemie International Edition* **2024**, e202416795. https://doi.org/10.1002/anie.202416795.

(533) Jin, F. J.; Maruyama, J.; Juvvadi, P. R.; Arioka, M.; Kitamoto, K. Development of a Novel Quadruple Auxotrophic Host Transformation System by *ArgB* Gene Disruption Using *AdeA* Gene and Exploiting Adenine Auxotrophy in *Aspergillus Oryzae*. *FEMS Microbiol Lett* **2004**, *239* (1), 79–85. https://doi.org/10.1016/j.femsle.2004.08.025.

(534) de Mattos-Shipley, K. M. J.; Lazarus, C. M.; Williams, K. Investigating Fungal Biosynthetic Pathways Using Heterologous Gene Expression: Aspergillus Oryzae as a Heterologous Host. In *Methods in Molecular Biology*; Humana Press Inc., 2022; Vol. 2489, pp 23–39. https://doi.org/10.1007/978-1-0716-2273-5_2.

(535) Dao, T. T.; de Mattos-Shipley, K. M. J.; Prosser, I. M.; Williams, K.; Zacharova, M. K.; Lazarus, C. M.; Willis, C. L.; Bailey, A. M. Cleaning the Cellular Factory–Deletion of McrA in Aspergillus Oryzae NSAR1 and the Generation of a Novel Kojic Acid Deficient Strain for Cleaner Heterologous Production of Secondary Metabolites. *Frontiers in Fungal Biology* **2021**, *2*, 632542. https://doi.org/10.3389/ffunb.2021.632542.

(536) OIKAWA, H. Heterologous Production of Fungal Natural Products: Reconstitution of Biosynthetic Gene Clusters in Model Host <I>Aspergillus Oryzae</I> *Proceedings of the Japan Academy, Series B* **2020**, *96* (9), 420–430. https://doi.org/10.2183/pjab.96.031.

(537) Sakai, K.; Kinoshita, H.; Shimizu, T.; Nihira, T. Construction of a Citrinin Gene Cluster Expression System in Heterologous Aspergillus Oryzae. *J Biosci Bioeng* **2008**, *106* (5), 466–472. https://doi.org/10.1263/jbb.106.466.

(538) Itoh, T.; Tokunaga, K.; Matsuda, Y.; Fujii, I.; Abe, I.; Ebizuka, Y.; Kushiro, T. Reconstitution of a Fungal Meroterpenoid Biosynthesis Reveals the Involvement of a Novel Family of Terpene Cyclases. *Nat Chem* **2010**, *2* (10), 858–864. https://doi.org/10.1038/nchem.764.

(539) Awakawa, T.; Liu, W.; Bai, T.; Taniguchi, T.; Abe, I. Orthoester Formation in Fungal Meroterpenoid Austalide F Biosynthesis. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2023**, *378* (1871), 20220037. https://doi.org/10.1098/rstb.2022.0037.

(540) Awakawa, T.; Abe, I. Reconstitution of Polyketide-Derived Meroterpenoid Biosynthetic Pathway in Aspergillus Oryzae. *Journal of Fungi* **2021**, *7* (6), 486. https://doi.org/10.3390/jof7060486.

(541) Tagami, K.; Liu, C.; Minami, A.; Noike, M.; Isaka, T.; Fueki, S.; Shichijo, Y.; Toshima, H.; Gomi, K.; Dairi, T.; Oikawa, H. Reconstitution of Biosynthetic Machinery for Indole-Diterpene Paxilline in *Aspergillus Oryzae*. *J Am Chem Soc* **2013**, *135* (4), 1260–1263. https://doi.org/10.1021/ja3116636.

(542) Jiang, Y.; Ozaki, T.; Harada, M.; Miyasaka, T.; Sato, H.; Miyamoto, K.; Kanazawa, J.; Liu, C.; Maruyama, J.; Adachi, M.; Nakazaki, A.; Nishikawa, T.; Uchiyama, M.; Minami, A.; Oikawa, H. Biosynthesis of Indole Diterpene Lolitrems: Radical-Induced Cyclization of an Epoxyalcohol Affording a Characteristic Lolitremane Skeleton. *Angewandte Chemie International Edition* **2020**, *59* (41), 17996–18002. https://doi.org/10.1002/anie.202007280.

(543) Ozaki, T.; Minami, A.; Oikawa, H. Biosynthesis of Indole Diterpenes: A Reconstitution Approach in a Heterologous Host. *Nat Prod Rep* **2023**, *40* (1), 202–213. https://doi.org/10.1039/D2NP00031H.

(544) Zhou, L.; Abe, I.; Awakawa, T. Biosynthesis of Dihydroxyardeemin by Heterologous Expression. *Tetrahedron* **2022**, *127*, 133095. https://doi.org/10.1016/j.tet.2022.133095.

(545) Meng, L.; Awakawa, T.; Li, X.; Quan, Z.; Yang, S.; Wang, B.; Abe, I. Discovery of (±)-Penindolenes Reveals an Unusual Indole Ring Cleavage Pathway Catalyzed by P450 Monooxygenase. *Angewandte Chemie* **2024**, *136* (26), e202403963. https://doi.org/10.1002/ange.202403963.

(546) Tao, H.; Mori, T.; Wei, X.; Matsuda, Y.; Abe, I. One Polyketide Synthase, Two Distinct Products: *Trans* -Acting Enzyme-Controlled Product Divergence in Calbistrin Biosynthesis. *Angewandte Chemie International Edition* **2021**, *60* (16), 8851–8858. https://doi.org/10.1002/anie.202016525.

(547) Wang, W.-G.; Wang, H.; Du, L.-Q.; Li, M.; Chen, L.; Yu, J.; Cheng, G.-G.; Zhan, M.-T.; Hu, Q.-F.; Zhang, L.; Yao, M.; Matsuda, Y. Molecular Basis for the Biosynthesis of an Unusual Chain-Fused Polyketide, Gregatin A. *J Am Chem Soc* **2020**, *142* (18), 8464–8472. https://doi.org/10.1021/jacs.0c02337.

(548) Williams, K.; Szwalbe, A. J.; Mulholland, N. P.; Vincent, J. L.; Bailey, A. M.; Willis, C. L.; Simpson, T. J.; Cox, R. J. Heterologous Production of Fungal Maleidrides Reveals the Cryptic Cyclization Involved in Their Biosynthesis. *Angewandte Chemie International Edition* **2016**, *55* (23), 6784–6788. https://doi.org/10.1002/anie.201511882.

(549) Han, Y. Bin; Bai, W.; Ding, C. X.; Liang, J.; Wu, S.-H.; Tan, R. X. Intertwined Biosynthesis of Skyrin and Rugulosin A Underlies the Formation of Cage-Structured Bisanthraquinones. *J Am Chem Soc* **2021**, *143* (35), 14218–14226. https://doi.org/10.1021/jacs.1c05421.

(550) Fukaya, M.; Ozaki, T.; Minami, A.; Oikawa, H. Biosynthetic Machineries of Anthraquinones and Bisanthraquinones in *Talaromyces Islandicus*. *Biosci Biotechnol Biochem* **2022**, *86* (4), 435–443. https://doi.org/10.1093/bbb/zbac009.

(551) Cao, Z.-Q.; Wang, G.-Q.; Luo, R.; Gao, Y.-H.; Lv, J.-M.; Qin, S.-Y.; Chen, G.-D.; Awakawa, T.; Bao, X.-F.; Mei, Q.-H.; Yao, X.-S.; Hu, D.; Abe, I.; Gao, H. Biosynthesis of Enfumafungin-Type Antibiotic Reveals an Unusual Enzymatic Fusion Pattern and Unprecedented C–C Bond Cleavage. *J Am Chem Soc* **2024**, *146* (18), 12723–12733. https://doi.org/10.1021/jacs.4c02415.

(552) Wang, G.-Q.; Chen, G.-D.; Qin, S.-Y.; Hu, D.; Awakawa, T.; Li, S.-Y.; Lv, J.-M.; Wang, C.-X.; Yao, X.-S.; Abe, I.; Gao, H. Biosynthetic Pathway for Furanosteroid Demethoxyviridin and Identification of an Unusual Pregnane Side-Chain Cleavage. *Nat Commun* **2018**, *9* (1), 1838. https://doi.org/10.1038/s41467-018-04298-2.

(553) Tang, J.; Matsuda, Y. Discovery of Fungal Onoceroid Triterpenoids through Domainless Enzyme-Targeted Global Genome Mining. *Nat Commun* **2024**, *15* (1), 4312. https://doi.org/10.1038/s41467-024-48771-7.

(554) Umemura, M.; Nagano, N.; Koike, H.; Kawano, J.; Ishii, T.; Miyamura, Y.; Kikuchi, M.; Tamano, K.; Yu, J.; Shin-ya, K.; Machida, M. Characterization of the Biosynthetic Gene Cluster for the Ribosomally Synthesized Cyclic Peptide Ustiloxin B in Aspergillus Flavus. *Fungal Genetics and Biology* **2014**, *68*, 23–30. https://doi.org/10.1016/j.fgb.2014.04.011.

(555) Ye, Y.; Ozaki, T.; Umemura, M.; Liu, C.; Minami, A.; Oikawa, H. Heterologous Production of Asperipin-2a: Proposal for Sequential Oxidative Macrocyclization by a Fungi-Specific DUF3328 Oxidase. *Org Biomol Chem* **2019**, *17* (1), 39–43. https://doi.org/10.1039/C8OB02824A.

(556) Shabani, S.; White, J. M.; Hutton, C. A. Total Synthesis of the Putative Structure of Asperipin-2a and Stereochemical Reassignment. *Org Lett* **2020**, *22* (19), 7730–7734. https://doi.org/10.1021/acs.orglett.0c02884.

(557) Jiang, Y.; Ozaki, T.; Liu, C.; Igarashi, Y.; Ye, Y.; Tang, S.; Ye, T.; Maruyama, J.; Minami, A.; Oikawa, H. Biosynthesis of Cyclochlorotine: Identification of the Genes Involved in Oxidative Transformations and Intramolecular *O* , *N* -Transacylation. *Org Lett* **2021**, *23* (7), 2616–2620. https://doi.org/10.1021/acs.orglett.1c00525.

(558) Yuan, Y.; Cheng, S.; Bian, G.; Yan, P.; Ma, Z.; Dai, W.; Chen, R.; Fu, S.; Huang, H.; Chi, H.; Cai, Y.; Deng, Z.; Liu, T. Efficient Exploration of Terpenoid Biosynthetic Gene Clusters in Filamentous Fungi. *Nature Catalysis 2022 5:4* **2022**, *5* (4), 277–287. https://doi.org/10.1038/s41929-022-00762-x.

(559) Behera, B. C. Citric Acid from *Aspergillus Niger*: A Comprehensive Overview. *Crit Rev Microbiol* **2020**, *46* (6), 727–749. https://doi.org/10.1080/1040841X.2020.1828815.

(560) Vassilev, NikolayB.; Vassileva, MariaCh.; Spassova, DimitrinkaI. Production of Gluconic Acid by Aspergillus Niger Immobilized on Polyurethane Foam. *Appl Microbiol Biotechnol* **1993**, *39* (3), 285–288. https://doi.org/10.1007/BF00192079.

(561) E., S.; N., D.-C.; J., F.; Dijck, P. van. On the Safety of Aspergillus Niger - a Review. *Appl Microbiol Biotechnol* **2002**, *59* (4–5), 426–435. https://doi.org/10.1007/s00253-002-1032-6.

(562) Lubertozzi, D.; Keasling, J. D. Developing Aspergillus as a Host for Heterologous Expression. *Biotechnol Adv* **2009**, *27* (1), 53–75. https://doi.org/10.1016/j.biotechadv.2008.09.001.

(563) Nikolaev, I.; Mathieu, M.; van de Vondervoort, P. J. I.; Visser, J.; Felenbok, B. Heterologous Expression of the Aspergillus Nidulans AlcR–AlcA System in Aspergillus Niger. *Fungal Genetics and Biology* **2002**, *37* (1), 89–97. https://doi.org/10.1016/S1087-1845(02)00037-3.

(564) Richter, L.; Wanka, F.; Boecker, S.; Storm, D.; Kurt, T.; Vural, Ö.; Süßmuth, R.; Meyer, V. Engineering of Aspergillus Niger for the Production of Secondary Metabolites. *Fungal Biol Biotechnol* **2014**, *1* (1), 4. https://doi.org/10.1186/s40694-014-0004-9.

(565) Boecker, S.; Grätz, S.; Kerwat, D.; Adam, L.; Schirmer, D.; Richter, L.; Schütze, T.; Petras, D.; Süssmuth, R. D.; Meyer, V. Aspergillus Niger Is a Superior Expression Host for the Production of Bioactive Fungal Cyclodepsipeptides. *Fungal Biol Biotechnol* **2018**, *5* (1), 4. https://doi.org/10.1186/s40694-018-0048-3.

(566) Kirchgaessner, L.; Wurlitzer, J. M.; Seibold, P. S.; Rakhmanov, M.; Gressler, M. A Genetic Tool to Express Long Fungal Biosynthetic Genes. *Fungal Biology and Biotechnology 2023 10:1* **2023**, *10* (1), 1–16. https://doi.org/10.1186/S40694-023-00152-3.

(567) Li, Y.; Li, C.; Fu, Y.; Zhang, Q.; Ma, J.; Zhou, J.; Li, J.; Du, G.; Liu, S. A CRISPR/Cas9-Based Visual Toolkit Enabling Multiplex Integration at Specific Genomic Loci in Aspergillus Niger. *Synth Syst Biotechnol* **2024**, *9* (2), 209–216. https://doi.org/10.1016/j.synbio.2024.01.014.

(568) Jarczynska, Z. D.; Rendsvig, J. K. H.; Pagels, N.; Viana, V. R.; Nødvig, C. S.; Kirchner, F. H.; Strucko, T.; Nielsen, M. L.; Mortensen, U. H. DIVERSIFY: A Fungal Multispecies Gene Expression Platform. *ACS Synth Biol* **2021**, *10* (3), 579–588. https://doi.org/10.1021/acssynbio.0c00587.

(569) Tran, V.-T.; Thai, H.-D.; Vu, T. X.; Vu, H. H.; Nguyen, G. T.; Trinh, M. T.; Tran, H. T. T.; Pham, H. T. T.; Le, N. T. H. An Efficient Agrobacterium-Mediated System Based on the PyrG Auxotrophic Marker for Recombinant Expression in the Filamentous Fungus Penicillium Rubens. *Biotechnol Lett* **2023**, *45* (5–6), 689–702. https://doi.org/10.1007/s10529-023-03374-y.

(570) Pohl, C.; Polli, F.; Schütze, T.; Viggiano, A.; Mózsik, L.; Jung, S.; de Vries, M.; Bovenberg, R. A. L.; Meyer, V.; Driessen, A. J. M. A Penicillium Rubens Platform Strain for Secondary Metabolite Production. *Sci Rep* **2020**, *10* (1), 7630. https://doi.org/10.1038/s41598-020-64893-6.

(571) Richardson, A. T.; Cameron, R. C.; Stevenson, L. J.; Singh, A. J.; Lukito, Y.; Berry, D.; Nicholson, M. J.; Parker, E. J. Biosynthesis of Nodulisporic Acids: A Multifunctional Monooxygenase Delivers a Complex and Highly Branched Array. *Angewandte Chemie International Edition* **2022**, *61* (49). https://doi.org/10.1002/anie.202213364.

(572) Bundela, R.; Cameron, R. C.; Singh, A. J.; McLellan, R. M.; Richardson, A. T.; Berry, D.; Nicholson, M. J.; Parker, E. J. Generation of Alternate Indole Diterpene Architectures in Two Species of *Aspergilli*. *J Am Chem Soc* **2023**, *145* (5), 2754–2758. https://doi.org/10.1021/jacs.2c11170.

(573) Hibbard, T. R.; McLellan, R. M.; Stevenson, L. J.; Richardson, A. T.; Nicholson, M. J.; Parker, E. J. Functional Crosstalk between Discrete Indole Terpenoid Gene Clusters in *Tolypocladium Album*. *Org Lett* **2023**, *25* (41), 7470–7475. https://doi.org/10.1021/acs.orglett.3c02412.

(574) Zhang, J.; Wu, C.; Wang, W.; Wang, W.; Wei, D. A Versatile Trichoderma Reesei Expression System for the Production of Heterologous Proteins. *Biotechnol Lett* **2018**, *40* (6), 965–972. https://doi.org/10.1007/s10529-018-2548-x.

(575) Cherry, J. R.; Fidantsef, A. L. Directed Evolution of Industrial Enzymes: An Update. *Curr Opin Biotechnol* **2003**, *14* (4), 438–443. https://doi.org/10.1016/S0958-1669(03)00099-5.

(576) Shenouda, M. L.; Ambilika, M.; Skellam, E.; Cox, R. J. Heterologous Expression of Secondary Metabolite Genes in Trichoderma Reesei for Waste Valorization. *Journal of Fungi* **2022**, *8* (4), 355. https://doi.org/10.3390/jof8040355.

(577) Kakule, T. B.; Zhang, S.; Zhan, J.; Schmidt, E. W. Biosynthesis of the Tetramic Acids Sch210971 and Sch210972. *Org Lett* **2015**, *17* (10), 2295–2297. https://doi.org/10.1021/acs.orglett.5b00715.

(578) Sirirungruang, S.; Markel, K.; Shih, P. M. Plant-Based Engineering for Production of High-Valued Natural Products. *Nat Prod Rep* **2022**, *39* (7), 1492–1509. https://doi.org/10.1039/D2NP00017B.

(579) Eidenberger, L.; Kogelmann, B.; Steinkellner, H. Plant-Based Biopharmaceutical Engineering. *Nature Reviews Bioengineering* **2023**, *1* (6), 426–439. https://doi.org/10.1038/s44222-023-00044-6.

(580) Bally, J.; Jung, H.; Mortimer, C.; Naim, F.; Philips, J. G.; Hellens, R.; Bombarely, A.; Goodin, M. M.; Waterhouse, P. M. The Rise and Rise of *Nicotiana Benthamiana*: A Plant for All Reasons. *Annu Rev Phytopathol* **2018**, *56* (1), 405–426. https://doi.org/10.1146/annurev-phyto-080417-050141.

(581) Lomonossoff, G. P.; D'Aoust, M.-A. Plant-Produced Biopharmaceuticals: A Case of Technical Developments Driving Clinical Deployment. *Science (1979)* **2016**, *353* (6305), 1237–1240. https://doi.org/10.1126/science.aaf6638.

(582) Reed, J.; Stephenson, M. J.; Miettinen, K.; Brouwer, B.; Leveau, A.; Brett, P.; Goss, R. J. M.; Goossens, A.; O'Connell, M. A.; Osbourn, A. A Translational Synthetic Biology Platform for Rapid Access to Gram-Scale Quantities of Novel Drug-like Molecules. *Metab Eng* **2017**, *42*. https://doi.org/10.1016/j.ymben.2017.06.012.

(583) Stephenson, M. J.; Reed, J.; Brouwer, B.; Osbourn, A. Transient Expression in <Em>Nicotiana Benthamiana</Em> Leaves for Triterpene Production at a Preparative Scale. *Journal of Visualized Experiments* **2018**, *2018* (138). https://doi.org/10.3791/58169.

(584) Li, J.; Zhang, X.; Renata, H. Asymmetric Chemoenzymatic Synthesis of (-)-Podophyllotoxin and Related Aryltetralin Lignans. *Angewandte Chemie International Edition* **2019**, *58* (34), 11657–11660. https://doi.org/10.1002/anie.201904102.

(585) Kim, S. S.; Wengier, D. L.; Ragland, C. J.; Sattely, E. S. Transcriptional Reactivation of Lignin Biosynthesis for the Heterologous Production of Etoposide Aglycone in *Nicotiana Benthamiana*. *ACS Synth Biol* **2022**, *11* (10), 3379–3387. https://doi.org/10.1021/acssynbio.2c00289.

(586) De La Peña, R.; Hodgson, H.; Liu, J. C.-T.; Stephenson, M. J.; Martin, A. C.; Owen, C.; Harkess, A.; Leebens-Mack, J.; Jimenez, L. E.; Osbourn, A.; Sattely, E. S. Complex Scaffold Remodeling in Plant Triterpene Biosynthesis. *Science (1979)* **2023**, *379* (6630), 361–368. https://doi.org/10.1126/science.adf1017.

(587) Hodgson, H.; De La Peña, R.; Stephenson, M. J.; Thimmappa, R.; Vincent, J. L.; Sattely, E. S.; Osbourn, A. Identification of Key Enzymes Responsible for Protolimonoid Biosynthesis in Plants: Opening the Door to Azadirachtin Production. *Proceedings of the National Academy of Sciences* **2019**, *116* (34), 17096–17104. https://doi.org/10.1073/pnas.1906083116.

(588) Hodgson, H.; Stephenson, M. J.; Kikuchi, S.; Martin, L. B. B.; Liu, J. C. T.; Casson, R.; Rejzek, M.; Sattely, E. S.; Osbourn, A. Plants Utilize a Protection/Deprotection Strategy in Limonoid Biosynthesis: A "Missing Link" Carboxylesterase Boosts Yields and Provides Insights into Furan Formation. *J Am Chem Soc* **2024**, *146* (43), 29305–29310. https://doi.org/10.1021/jacs.4c11213.

(589) Chen, X.; Hudson, G. A.; Mineo, C.; Amer, B.; Baidoo, E. E. K.; Crowe, S. A.; Liu, Y.; Keasling, J. D.; Scheller, H. V. Deciphering Triterpenoid Saponin Biosynthesis by Leveraging Transcriptome Response to Methyl Jasmonate Elicitation in Saponaria Vaccaria. *Nat Commun* **2023**, *14* (1), 7101. https://doi.org/10.1038/s41467-023-42877-0.

(590) Reed, J.; Orme, A.; El-Demerdash, A.; Owen, C.; Martin, L. B. B.; Misra, R. C.; Kikuchi, S.; Rejzek, M.; Martin, A. C.; Harkess, A.; Leebens-Mack, J.; Louveau, T.; Stephenson, M. J.; Osbourn, A. Elucidation of the Pathway for Biosynthesis of Saponin Adjuvants from the Soapbark Tree. *Science (1979)* **2023**, *379* (6638), 1252–1264. https://doi.org/10.1126/science.adf3727.

(591) Croteau, R.; Ketchum, R. E. B.; Long, R. M.; Kaspera, R.; Wildung, M. R. Taxol Biosynthesis and Molecular Genetics. *Phytochemistry Reviews* **2006**, *5* (1), 75–97. https://doi.org/10.1007/s11101-005-3748-2.

(592) Jennewein, S.; Long, R. M.; Williams, R. M.; Croteau, R. Cytochrome P450 Taxadiene 5α-Hydroxylase, a Mechanistically Unusual Monooxygenase Catalyzing the First Oxygenation Step of Taxol Biosynthesis. *Chem Biol* **2004**, *11* (3), 379–387. https://doi.org/10.1016/j.chembiol.2004.02.022.

(593) Jennewein, S.; Rithner, C. D.; Williams, R. M.; Croteau, R. B. Taxol Biosynthesis: Taxane 13α-Hydroxylase Is a Cytochrome P450-Dependent Monooxygenase. *Proceedings of the National Academy of Sciences* **2001**, *98* (24), 13595–13600. https://doi.org/10.1073/pnas.251539398.

(594) Schoendorf, A.; Rithner, C. D.; Williams, R. M.; Croteau, R. B. Molecular Cloning of a Cytochrome P450 Taxane 10β-Hydroxylase CDNA from *Taxus* and Functional Expression in Yeast. *Proceedings of the National Academy of Sciences* **2001**, *98* (4), 1501–1506. https://doi.org/10.1073/pnas.98.4.1501.

(595) Chau, M.; Croteau, R. Molecular Cloning and Characterization of a Cytochrome P450 Taxoid 2α-Hydroxylase Involved in Taxol Biosynthesis. *Arch Biochem Biophys* **2004**, *427* (1), 48–57. https://doi.org/10.1016/j.abb.2004.04.016.

(596) Chau, M.; Jennewein, S.; Walker, K.; Croteau, R. Taxol Biosynthesis. *Chem Biol* **2004**, *11* (5), 663–672. https://doi.org/10.1016/j.chembiol.2004.02.025.

(597) Walker, K.; Schoendorf, A.; Croteau, R. Molecular Cloning of a Taxa-4(20),11(12)-Dien-5α-Ol-O-Acetyl Transferase CDNA from Taxus and Functional Expression in Escherichia Coli. *Arch Biochem Biophys* **2000**, *374* (2), 371–380. https://doi.org/10.1006/abbi.1999.1609.

(598) Li, J.; Mutanda, I.; Wang, K.; Yang, L.; Wang, J.; Wang, Y. Chloroplastic Metabolic Engineering Coupled with Isoprenoid Pool Enhancement for Committed Taxanes Biosynthesis in Nicotiana Benthamiana. *Nat Commun* **2019**, *10* (1), 4850. https://doi.org/10.1038/s41467-019-12879-y.

(599) Liu, J. C.-T.; De La Peña, R.; Tocol, C.; Sattely, E. S. Reconstitution of Early Paclitaxel Biosynthetic Network. *Nat Commun* **2024**, *15* (1), 1419. https://doi.org/10.1038/s41467-024-45574-8.

(600) Zhang, Y.; Wiese, L.; Fang, H.; Alseekh, S.; Perez de Souza, L.; Scossa, F.; Molloy, J. J.; Christmann, M.; Fernie, A. R. Synthetic Biology Identifies the Minimal Gene Set Required for Paclitaxel Biosynthesis in a Plant Chassis. *Mol Plant* **2023**, *16* (12), 1951–1961. https://doi.org/10.1016/j.molp.2023.10.016.

(601) Zhao, Y.; Liang, F.; Xie, Y.; Duan, Y.-T.; Andeadelli, A.; Pateraki, I.; Makris, A. M.; Pomorski, T. G.; Staerk, D.; Kampranis, S. C. Oxetane Ring Formation in Taxol Biosynthesis Is Catalyzed by a Bifunctional Cytochrome P450 Enzyme. *J Am Chem Soc* **2024**, *146* (1), 801–810. https://doi.org/10.1021/jacs.3c10864.

(602) Jiang, B.; Gao, L.; Wang, H.; Sun, Y.; Zhang, X.; Ke, H.; Liu, S.; Ma, P.; Liao, Q.; Wang, Y.; Wang, H.; Liu, Y.; Du, R.; Rogge, T.; Li, W.; Shang, Y.; Houk, K. N.; Xiong, X.; Xie, D.; Huang, S.; Lei, X.; Yan, J. Characterization and Heterologous Reconstitution of *Taxus* Biosynthetic Enzymes Leading to Baccatin III. *Science (1979)* **2024**, *383* (6683), 622–629. https://doi.org/10.1126/science.adj3484.

(603) McClune, C. J.; Liu, J. C.-T.; Wick, C.; De La Peña, R.; Lange, B. M.; Fordyce, P. M.; Sattely, E. S. Multiplexed Perturbation of Yew Reveals Cryptic Proteins That Enable a Total Biosynthesis of Baccatin III and Taxol Precursors. *BioRxiv* **2024**. https://doi.org/10.1101/2024.11.06.622305. Accessed 2024-12-01.

(604) Hong, B.; Grzech, D.; Caputi, L.; Sonawane, P.; López, C. E. R.; Kamileen, M. O.; Hernández Lozada, N. J.; Grabe, V.; O'Connor, S. E. Biosynthesis of Strychnine. *Nature* **2022**, *607* (7919), 617–622. https://doi.org/10.1038/s41586-022-04950-4.

(605) Wang, Y.-J.; Huang, J.-P.; Tian, T.; Yan, Y.; Chen, Y.; Yang, J.; Chen, J.; Gu, Y.-C.; Huang, S.-X. Discovery and Engineering of the Cocaine Biosynthetic Pathway. *J Am Chem Soc* **2022**, *144* (48), 22000–22007. https://doi.org/10.1021/jacs.2c09091.

(606) Forman, V.; Luo, D.; Geu-Flores, F.; Lemcke, R.; Nelson, D. R.; Kampranis, S. C.; Staerk, D.; Møller, B. L.; Pateraki, I. A Gene Cluster in Ginkgo Biloba Encodes Unique Multifunctional Cytochrome P450s That Initiate Ginkgolide Biosynthesis. *Nat Commun* **2022**, *13* (1), 5143. https://doi.org/10.1038/s41467-022-32879-9.

(607) Nett, R. S.; Dho, Y.; Tsai, C.; Passow, D.; Martinez Grundman, J.; Low, Y.-Y.; Sattely, E. S. Plant Carbonic Anhydrase-like Enzymes in Neuroactive Alkaloid Biosynthesis. *Nature* **2023**, *624* (7990), 182–191. https://doi.org/10.1038/s41586-023-06716-y.

(608) Nett, R. S.; Dho, Y.; Low, Y.-Y.; Sattely, E. S. A Metabolic Regulon Reveals Early and Late Acting Enzymes in Neuroactive Lycopodium Alkaloid Biosynthesis. *Proceedings of the National Academy of Sciences* **2021**, *118* (24), e2102949118. https://doi.org/10.1073/pnas.2102949118.

(609) Polturak, G.; Misra, R. C.; El-Demerdash, A.; Owen, C.; Steed, A.; McDonald, H. P.; Wang, J.; Saalbach, G.; Martins, C.; Chartrain, L.; Wilkinson, B.; Nicholson, P.; Osbourn, A. Discovery of Isoflavone Phytoalexins in Wheat Reveals an Alternative Route to Isoflavonoid Biosynthesis. *Nat Commun* **2023**, *14* (1), 6977. https://doi.org/10.1038/s41467-023-42464-3.

(610) Mikkelsen, M. D.; Olsen, C. E.; Halkier, B. A. Production of the Cancer-Preventive Glucoraphanin in Tobacco. *Mol Plant* **2010**, *3* (4), 751–759. https://doi.org/10.1093/mp/ssq020.

(611) Crocoll, C.; Mirza, N.; Reichelt, M.; Gershenzon, J.; Halkier, B. A. Optimization of Engineered Production of the Glucoraphanin Precursor Dihomomethionine in Nicotiana Benthamiana. *Front Bioeng Biotechnol* **2016**, *4* (FEB), 178767. https://doi.org/10.3389/fbioe.2016.00014.

(612) Dias, S. L.; Chuang, L.; Liu, S.; Seligmann, B.; Brendel, F. L.; Chavez, B. G.; Hoffie, R. E.; Hoffie, I.; Kumlehn, J.; Bültemeier, A.; Wolf, J.; Herde, M.; Witte, C.-P.; D'Auria, J. C.; Franke, J. Biosynthesis of the Allelopathic Alkaloid Gramine in Barley by a Cryptic Oxidative Rearrangement. *Science (1979)* **2024**, *383* (6690), 1448–1454. https://doi.org/10.1126/science.adk6112.

(613) Mehta, N.; Meng, Y.; Zare, R.; Kamenetsky-Goldstein, R.; Sattely, E. A Developmental Gradient Reveals Biosynthetic Pathways to Eukaryotic Toxins in Monocot Geophytes. *Cell* **2024**, *187* (20), 5620-5637.e10. https://doi.org/10.1016/j.cell.2024.08.027.

(614) Grzech, D.; Smit, S. J.; Alam, R. M.; Boccia, M.; Nakamura, Y.; Hong, B.; Barbole, R.; Heinicke, S.; Kunert, M.; Seibt, W.; Grabe, V.; Caputi, L.; Lichman, B. R.; O'Connor, S. E.; Aharoni, A.; Sonawane, P. D. Incorporation of Nitrogen in Antinutritional Solanum Alkaloid Biosynthesis. *Nat Chem Biol* **2024**. https://doi.org/10.1038/s41589-024-01735-w.

(615) Chen, J.; Tan, J.; Duan, X.; Wang, Y.; Wen, J.; Li, W.; Li, Z.; Wang, G.; Xu, H. Plastidial Engineering with Coupled Farnesyl Diphosphate Pool Reconstitution and Enhancement for Sesquiterpene Biosynthesis in Tomato Fruit. *Metab Eng* **2023**, *77*, 41–52. https://doi.org/10.1016/j.ymben.2023.03.002.

(616) Schuurink, R.; Tissier, A. Glandular Trichomes: Micro-organs with Model Status? *New Phytologist* **2020**, *225* (6), 2251–2266. https://doi.org/10.1111/nph.16283.

(617) Wang, Y.; Wen, J.; Liu, L.; Chen, J.; Wang, C.; Li, Z.; Wang, G.; Pichersky, E.; Xu, H. Engineering of Tomato Type VI Glandular Trichomes for Trans-Chrysanthemic Acid Biosynthesis, the Acid Moiety of Natural Pyrethrin Insecticides. *Metab Eng* **2022**, *72*, 188–199. https://doi.org/10.1016/j.ymben.2022.03.007.

(618) Huchelmann, A.; Boutry, M.; Hachez, C. Plant Glandular Trichomes: Natural Cell Factories of High Biotechnological Interest. *Plant Physiol* **2017**, *175* (1), 6–22. https://doi.org/10.1104/pp.17.00727.

(619) Lee, S. G.; Roskoski, R.; Bauer, K.; Lipmann, F. Purification of the Polyenzymes Responsible for Tyrocidine Synthesis and Their Dissociation into Subunits. *Biochemistry* **1973**, *12* (3), 398–405. https://doi.org/10.1021/bi00727a006.

(620) Lipmann, F. Nonribosomal Polypeptide Synthesis on Polyenzyme Templates. *Acc Chem Res* **1973**, *6* (11), 361–367. https://doi.org/10.1021/ar50071a001.

(621) Whicher, J. R.; Dutta, S.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Dosey, A. M.; Narayan, A. R. H.; Håkansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. Structural Rearrangements of a Polyketide Synthase Module during Its Catalytic Cycle. *Nature* **2014**, *510* (7506), 560–564. https://doi.org/10.1038/nature13409.

(622) McCullough, T. M.; Dhar, A.; Akey, D. L.; Konwerski, J. R.; Sherman, D. H.; Smith, J. L. Structure of a Modular Polyketide Synthase Reducing Region. *Structure* **2023**, *31* (9), 1109-1120.e3. https://doi.org/10.1016/j.str.2023.05.019.

(623) Bonhomme, S.; Contreras-Martel, C.; Dessen, A.; Macheboeuf, P. Architecture of a PKS-NRPS Hybrid Megaenzyme Involved in the Biosynthesis of the Genotoxin Colibactin. *Structure* **2023**, *31* (6), 700-712.e4. https://doi.org/10.1016/j.str.2023.03.012.

(624) Cogan, D. P.; Zhang, K.; Li, X.; Li, S.; Pintilie, G. D.; Roh, S.-H.; Craik, C. S.; Chiu, W.; Khosla, C. Mapping the Catalytic Conformations of an Assembly-Line Polyketide Synthase Module. *Science (1979)* **2021**, *374* (6568), 729–734. https://doi.org/10.1126/science.abi8358.

(625) Bagde, S. R.; Mathews, I. I.; Fromme, J. C.; Kim, C.-Y. Modular Polyketide Synthase Contains Two Reaction Chambers That Operate Asynchronously. *Science (1979)* **2021**, *374* (6568), 723–729. https://doi.org/10.1126/science.abi8532.

(626) Chu Yuan Kee, M.-J.; Bharath, S. R.; Wee, S.; Bowler, M. W.; Gunaratne, J.; Pan, S.; Zhang, L.; Song, H. Structural Insights into the Substrate-Bound Condensation Domains of Non-Ribosomal Peptide Synthetase AmbB. *Sci Rep* **2022**, *12* (1), 5353. https://doi.org/10.1038/s41598-022-09188-8.

(627) Folger, I. B.; Frota, N. F.; Pistofidis, A.; Niquille, D. L.; Hansen, D. A.; Schmeing, T. M.; Hilvert, D. High-Throughput Reprogramming of an NRPS Condensation Domain. *Nat Chem Biol* **2024**, *20* (6), 761–769. https://doi.org/10.1038/s41589-023-01532-x.

(628) Keatinge-Clay, A. T. Polyketide Synthase Modules Redefined. *Angewandte Chemie International Edition* **2017**, *56* (17), 4658–4660. https://doi.org/10.1002/anie.201701281.

(629) Miyazawa, T.; Hirsch, M.; Zhang, Z.; Keatinge-Clay, A. T. An in Vitro Platform for Engineering and Harnessing Modular Polyketide Synthases. *Nat Commun* **2020**, *11* (1), 80. https://doi.org/10.1038/s41467-019-13811-0.

(630) Zhang, J.; Bista, R.; Miyazawa, T.; Keatinge-Clay, A. T. Boosting Titers of Engineered Triketide and Tetraketide Synthases to Record Levels through T7 Promoter Tuning. *Metab Eng* **2023**, *78*, 93–98. https://doi.org/10.1016/j.ymben.2023.05.008.

(631) Mabesoone, M. F. J.; Leopold-Messer, S.; Minas, H. A.; Chepkirui, C.; Chawengrum, P.; Reiter, S.; Meoded, R. A.; Wolf, S.; Genz, F.; Magnus, N.; Piechulla, B.; Walker, A. S.; Piel, J. Evolution-Guided Engineering of *Trans* -Acyltransferase Polyketide Synthases. *Science (1979)* **2024**, *383* (6689), 1312–1317. https://doi.org/10.1126/science.adj7621.

(632) Gay, D. C.; Wagner, D. T.; Meinke, J. L.; Zogzas, C. E.; Gay, G. R.; Keatinge-Clay, A. T. The LINKS Motif Zippers Trans-Acyltransferase Polyketide Synthase Assembly Lines into a Biosynthetic Megacomplex. *J Struct Biol* **2016**, *193* (3), 196–205. https://doi.org/10.1016/j.jsb.2015.12.011.

(633) Gay, D. C.; Gay, G.; Axelrod, A. J.; Jenner, M.; Kohlhaas, C.; Kampa, A.; Oldham, N. J.; Piel, J.; Keatinge-Clay, A. T. A Close Look at a Ketosynthase from a Trans-Acyltransferase Modular Polyketide Synthase. *Structure* **2014**, *22* (3), 444–451. https://doi.org/10.1016/j.str.2013.12.016.

(634) Yuzawa, S.; Deng, K.; Wang, G.; Baidoo, E. E. K.; Northen, T. R.; Adams, P. D.; Katz, L.; Keasling, J. D. Comprehensive *in Vitro* Analysis of Acyltransferase Domain Exchanges in Modular Polyketide Synthases and Its Application for Short-Chain Ketone Production. *ACS Synth Biol* **2017**, *6* (1), 139–147. https://doi.org/10.1021/acssynbio.6b00176.

(635) Englund, E.; Schmidt, M.; Nava, A. A.; Lechner, A.; Deng, K.; Jocic, R.; Lin, Y.; Roberts, J.; Benites, V. T.; Kakumanu, R.; Gin, J. W.; Chen, Y.; Liu, Y.; Petzold, C. J.; Baidoo, E. E. K.; Northen, T. R.; Adams, P. D.; Katz, L.; Yuzawa, S.; Keasling, J. D. Expanding Extender Substrate Selection for Unnatural Polyketide Biosynthesis by Acyltransferase Domain Exchange within a Modular Polyketide Synthase. *J Am Chem Soc* **2023**, *145* (16), 8822–8832. https://doi.org/10.1021/jacs.2c11027.

(636) Tanovic, A.; Samel, S. A.; Essen, L.-O.; Marahiel, M. A. Crystal Structure of the Termination Module of a Nonribosomal Peptide Synthetase. *Science (1979)* **2008**, *321* (5889), 659–663. https://doi.org/10.1126/science.1159850.

(637) Thong, W. L.; Zhang, Y.; Zhuo, Y.; Robins, K. J.; Fyans, J. K.; Herbert, A. J.; Law, B. J. C.; Micklefield, J. Gene Editing Enables Rapid Engineering of Complex Antibiotic Assembly Lines. *Nat Commun* **2021**, *12* (1), 6872. https://doi.org/10.1038/s41467-021-27139-1.

(638) Kries, H.; Niquille, D. L.; Hilvert, D. A Subdomain Swap Strategy for Reengineering Nonribosomal Peptides. *Chem Biol* **2015**, *22* (5), 640–648. https://doi.org/10.1016/j.chembiol.2015.04.015.

(639) Bozhüyük, K. A. J.; Linck, A.; Tietze, A.; Kranz, J.; Wesche, F.; Nowak, S.; Fleischhacker, F.; Shi, Y.-N.; Grün, P.; Bode, H. B. Modification and de Novo Design of Non-Ribosomal Peptide Synthetases Using Specific Assembly Points within Condensation Domains. *Nat Chem* **2019**, *11* (7), 653–661. https://doi.org/10.1038/s41557-019-0276-z.

(640) Bozhüyük, K. A. J.; Fleischhacker, F.; Linck, A.; Wesche, F.; Tietze, A.; Niesert, C.-P.; Bode, H. B. De Novo Design and Engineering of Non-Ribosomal Peptide Synthetases. *Nat Chem* **2018**, *10* (3), 275–281. https://doi.org/10.1038/nchem.2890.

(641) Calcott, M. J.; Owen, J. G.; Ackerley, D. F. Efficient Rational Modification of Non-Ribosomal Peptides by Adenylation Domain Substitution. *Nat Commun* **2020**, *11* (1), 4554. https://doi.org/10.1038/s41467-020-18365-0.

(642) Baunach, M.; Chowdhury, S.; Stallforth, P.; Dittmann, E. The Landscape of Recombination Events That Create Nonribosomal Peptide Diversity. *Mol Biol Evol* **2021**, *38* (5), 2116–2130. https://doi.org/10.1093/molbev/msab015.

(643) Booth, T. J.; Bozhüyük, K. A. J.; Liston, J. D.; Batey, S. F. D.; Lacey, E.; Wilkinson, B. Bifurcation Drives the Evolution of Assembly-Line Biosynthesis. *Nat Commun* **2022**, *13* (1), 3498. https://doi.org/10.1038/s41467-022-30950-z.

(644) Beck, C.; Garzón, J. F. G.; Weber, T. Recent Advances in Re-Engineering Modular PKS and NRPS Assembly Lines. *Biotechnology and Bioprocess Engineering* **2020**, *25* (6), 886–894. https://doi.org/10.1007/s12257-020-0265-5.

(645) Nguyen, K. T.; Ritz, D.; Gu, J.-Q.; Alexander, D.; Chu, M.; Miao, V.; Brian, P.; Baltz, R. H. Combinatorial Biosynthesis of Novel Antibiotics Related to Daptomycin. *Proceedings of the National Academy of Sciences* **2006**, *103* (46), 17462–17467. https://doi.org/10.1073/pnas.0608589103.

(646) Yan, F.; Burgard, C.; Popoff, A.; Zaburannyi, N.; Zipf, G.; Maier, J.; Bernauer, H. S.; Wenzel, S. C.; Müller, R. Synthetic Biology Approaches and Combinatorial Biosynthesis towards Heterologous Lipopeptide Production. *Chem Sci* **2018**, *9* (38), 7510–7519. https://doi.org/10.1039/C8SC02046A.

(647) Perrin, J.; Kulagina, N.; Unlubayir, M.; Munsch, T.; Carqueijeiro, I.; Dugé de Bernonville, T.; De Craene, J.-O.; Clastre, M.; St-Pierre, B.; Giglioli-Guivarc'h, N.; Gagneul, D.; Lanoue, A.; Courdavault, V.; Besseau, S. Exploiting Spermidine *N* -Hydroxycinnamoyltransferase Diversity and Substrate Promiscuity to Produce Various Trihydroxycinnamoyl Spermidines and Analogues in Engineered Yeast. *ACS Synth Biol* **2021**, *10* (2), 286–296. https://doi.org/10.1021/acssynbio.0c00391.

(648) Rütschlin, S.; Gunesch, S.; Böttcher, T. One Enzyme To Build Them All: Ring-Size Engineered Siderophores Inhibit the Swarming Motility of *Vibrio*. *ACS Chem Biol* **2018**, *13* (5), 1153–1158. https://doi.org/10.1021/acschembio.8b00084.

(649) Boddy, C. N.; Hotta, K.; Tse, M. L.; Watts, R. E.; Khosla, C. Precursor-Directed Biosynthesis of Epothilone in *Escherichia c Oli*. *J Am Chem Soc* **2004**, *126* (24), 7436–7437. https://doi.org/10.1021/ja048108s.

(650) BEHRENS, O. K.; CORSE, J. Biosynthesis of Penicillins; New Crystalline Biosynthetic Penicillins. *J Biol Chem* **1948**, *175* (2), 793–809.

(651) Jacobsen, J. R.; Hutchinson, C. R.; Cane, D. E.; Khosla, C. Precursor-Directed Biosynthesis of Erythromycin Analogs by an Engineered Polyketide Synthase. *Science (1979)* **1997**, *277* (5324), 367–369. https://doi.org/10.1126/science.277.5324.367.

(652) McCoy, E.; O'Connor, S. E. Directed Biosynthesis of Alkaloid Analogs in the Medicinal Plant *Catharanthus Roseus*. *J Am Chem Soc* **2006**, *128* (44), 14276–14277. https://doi.org/10.1021/ja066787w.

(653) Gotsbacher, M. P.; Codd, R. Azido-Desferrioxamine Siderophores as Functional Click-Chemistry Probes Generated in Culture upon Adding a Diazo-Transfer Reagent. *ChemBioChem* **2020**, *21* (10), 1433–1445. https://doi.org/10.1002/cbic.201900661.

(654) Chen, Y.; Paetz, C.; Schneider, B. Precursor-Directed Biosynthesis of Phenylbenzoisoquinolindione Alkaloids and the Discovery of a Phenylphenalenone-Based Plant Defense Mechanism. *J Nat Prod* **2018**, *81* (4), 879–884. https://doi.org/10.1021/acs.jnatprod.7b00885.

(655) Amrine, C. S. M.; Long, J. L.; Raja, H. A.; Kurina, S. J.; Burdette, J. E.; Pearce, C. J.; Oberlies, N. H. Engineering Fluorine into Verticillins (Epipolythiodioxopiperazine Alkaloids) via Precursor-Directed Biosynthesis. *J Nat Prod* **2019**, *82* (11), 3104–3110. https://doi.org/10.1021/acs.jnatprod.9b00711.

(656) Price, N. P. J.; Jackson, M. A.; Hartman, T. M.; Bannantine, J. P.; Naumann, T. A.; Vermillion, K. E.; Koch, A. A.; Kennedy, P. D. Precursor-Directed Biosynthesis and Biological Testing of *Omega* -Alicyclic- and *Neo* -Branched Tunicamycin *N* -Acyl Variants. *ACS Chem Biol* **2023**, *18* (10), 2267–2280. https://doi.org/10.1021/acschembio.3c00324.

(657) Rivera-Chávez, J.; Raja, H. A.; Graf, T. N.; Burdette, J. E.; Pearce, C. J.; Oberlies, N. H. Biosynthesis of Fluorinated Peptaibols Using a Site-Directed Building Block Incorporation Approach. *J Nat Prod* **2017**, *80* (6), 1883–1892. https://doi.org/10.1021/acs.jnatprod.7b00189.

(658) Neuenschwander, A.; Rocha, V. P. C.; Bastos, T. M.; Marcourt, L.; Morin, H.; da Rocha, C. Q.; Grimaldi, G. B.; de Sousa, K. A. F.; Borges, J. N.; Rivara-Minten, E.; Wolfender, J.-L.; Soares, M. B. P.; Queiroz, E. F. Production of Highly

Active Antiparasitic Compounds from the Controlled Halogenation of the *Arrabidaea Brachypoda* Crude Plant Extract. *J Nat Prod* **2020**, *83* (9), 2631–2640. https://doi.org/10.1021/acs.jnatprod.0c00433.

(659) Yan, Y.; Chen, J.; Zhang, L.; Zheng, Q.; Han, Y.; Zhang, H.; Zhang, D.; Awakawa, T.; Abe, I.; Liu, W. Multiplexing of Combinatorial Chemistry in Antimycin Biosynthesis: Expansion of Molecular Diversity and Utility. *Angewandte Chemie International Edition* **2013**, *52* (47), 12308–12312. https://doi.org/10.1002/anie.201305569.

(660) Zhang, N.; Liu, L.; Shan, G.; Cai, Q.; Lei, X.; Hong, B.; Wu, L.; Xie, Y.; Chen, R. Precursor-Directed Biosynthesis of New Sansanmycin Analogs Bearing Para-Substituted-Phenylalanines with High Yields. *J Antibiot (Tokyo)* **2016**, *69* (10), 765–768. https://doi.org/10.1038/ja.2016.2.

(661) Telfer, T. J.; Richardson-Sanchez, T.; Gotsbacher, M. P.; Nolan, K. P.; Tieu, W.; Codd, R. Analogues of Desferrioxamine B (DFOB) with New Properties and New Functions Generated Using Precursor-Directed Biosynthesis. *BioMetals* **2019**, *32* (3), 395–408. https://doi.org/10.1007/s10534-019-00175-7.

(662) Richardson-Sanchez, T.; Nolan, K. P.; Codd, R. Rubik's Cube of Siderophore Assembly Established from Mixed-Substrate Precursor-Directed Biosynthesis. *ACS Omega* **2018**, *3* (12), 18160–18169. https://doi.org/10.1021/acsomega.8b02803.

(663) Sester, A.; Winand, L.; Pace, S.; Hiller, W.; Werz, O.; Nett, M. Myxochelin- and Pseudochelin-Derived Lipoxygenase Inhibitors from a Genetically Engineered *Myxococcus Xanthus* Strain. *J Nat Prod* **2019**, *82* (9), 2544–2549. https://doi.org/10.1021/acs.jnatprod.9b00403.

(664) Dewa, A. A.; Elbanna, A. H.; Khalil, Z. G.; Capon, R. J. Neochrysosporazines: Precursor-Directed Biosynthesis Defines a Marine-Derived Fungal Natural Product P-Glycoprotein Inhibitory Pharmacophore. *J Med Chem* **2022**, *65* (3), 2610–2622. https://doi.org/10.1021/acs.jmedchem.1c01989.

(665) Birch, A. J. The Biosynthesis of Antibiotics. *Pure and Applied Chemistry* **1963**, *7* (4), 527–538. https://doi.org/10.1351/pac196307040527.

(666) Rinehart, K. L. Mutasynthesis of New Antibiotics. *Pure and Applied Chemistry* **1977**, *49* (9), 1361–1384. https://doi.org/10.1351/pac197749091361.

(667) Shier, W. T.; Rinehart, K. L.; Gottlieb, D. PREPARATION OF FOUR NEW ANTIBIOTICS FROM A MUTANT OF *Streptomyces Fradiae*. *Proceedings of the National Academy of Sciences* **1969**, *63* (1), 198–204. https://doi.org/10.1073/pnas.63.1.198.

(668) Winand, L.; Schneider, P.; Kruth, S.; Greven, N.-J.; Hiller, W.; Kaiser, M.; Pietruszka, J.; Nett, M. Mutasynthesis of Physostigmines in *Myxococcus Xanthus*. *Org Lett* **2021**, *23* (16), 6563–6567. https://doi.org/10.1021/acs.orglett.1c02374.

(669) Khatri, Y.; Hohlman, R. M.; Mendoza, J.; Li, S.; Lowell, A. N.; Asahara, H.; Sherman, D. H. Multicomponent Microscale Biosynthesis of Unnatural Cyanobacterial Indole Alkaloids. *ACS Synth Biol* **2020**, *9* (6), 1349–1360. https://doi.org/10.1021/acssynbio.0c00038.

(670) Mu, N.; Guo, H.; Zhang, E.; Yin, Y.; Wang, W.; Chen, D.; Wang, S.; Liu, W. Mutasynthesis Generates Antibacterial Benzothiophenic-Containing Nosiheptide Analogues. *J Nat Prod* **2022**, *85* (10), 2274–2281. https://doi.org/10.1021/acs.jnatprod.2c00273.

(671) Hennrich, O.; Weinmann, L.; Kulik, A.; Harms, K.; Klahn, P.; Youn, J.-W.; Surup, F.; Mast, Y. Biotransformation-Coupled Mutasynthesis for the Generation of Novel Pristinamycin Derivatives by Engineering the Phenylglycine Residue. *RSC Chem Biol* **2023**, *4* (12), 1050–1063. https://doi.org/10.1039/D3CB00143A.

(672) Schmidt, J. J.; Khatri, Y.; Brody, S. I.; Zhu, C.; Pietraszkiewicz, H.; Valeriote, F. A.; Sherman, D. H. A Versatile Chemoenzymatic Synthesis for the Discovery of Potent Cryptophycin Analogs. *ACS Chem Biol* **2020**, *15* (2), 524–532. https://doi.org/10.1021/acschembio.9b00998.

(673) Chen, L.; Li, Y.; Yue, Q.; Loksztejn, A.; Yokoyama, K.; Felix, E. A.; Liu, X.; Zhang, N.; An, Z.; Bills, G. F. Engineering of New Pneumocandin Side-Chain Analogues from *Glarea Lozoyensis* by Mutasynthesis and Evaluation of Their Antifungal Activity. *ACS Chem Biol* **2016**, *11* (10), 2724–2733. https://doi.org/10.1021/acschembio.6b00604.

(674) Hindra; Yang, D.; Teng, Q.; Dong, L.-B.; Crnovčić, I.; Huang, T.; Ge, H.; Shen, B. Genome Mining of *Streptomyces Mobaraensis* DSM40847 as a Bleomycin Producer Providing a Biotechnology Platform To Engineer Designer Bleomycin Analogues. *Org Lett* **2017**, *19* (6), 1386–1389. https://doi.org/10.1021/acs.orglett.7b00283.

(675) Klein, A. S.; Domröse, A.; Bongen, P.; Brass, H. U. C.; Classen, T.; Loeschcke, A.; Drepper, T.; Laraia, L.; Sievers, S.; Jaeger, K.-E.; Pietruszka, J. New Prodigiosin Derivatives Obtained by Mutasynthesis in *Pseudomonas Putida*. *ACS Synth Biol* **2017**, *6* (9), 1757–1765. https://doi.org/10.1021/acssynbio.7b00099.

(676) Dong, L.-B.; Rudolf, J. D.; Shen, B. A Mutasynthetic Library of Platensimycin and Platencin Analogues. *Org Lett* **2016**, *18* (18), 4606–4609. https://doi.org/10.1021/acs.orglett.6b02248.

(677) Valentová, K. Biotransformation of Natural Products and Phytochemicals: Metabolites, Their Preparation, and Properties. *Int J Mol Sci* **2023**, *24* (9), 8030. https://doi.org/10.3390/ijms24098030.

(678) van Beek, T. A.; Claassen, F. W.; Dorado, J.; Godejohann, M.; Sierra-Alvarez, R.; Wijnberg, J. B. P. A. Fungal Biotransformation Products of Dehydroabietic Acid. *J Nat Prod* **2007**, *70* (2), 154–159. https://doi.org/10.1021/np060325e.

(679) Kar, S.; Sanderson, H.; Roy, K.; Benfenati, E.; Leszczynski, J. Green Chemistry in the Synthesis of Pharmaceuticals. *Chem Rev* **2022**, *122* (3), 3637–3710. https://doi.org/10.1021/acs.chemrev.1c00631.

(680) Fernandes, P.; Cruz, A.; Angelova, B.; Pinheiro, H. M.; Cabral, J. M. S. Microbial Conversion of Steroid Compounds: Recent Developments. *Enzyme Microb Technol* **2003**, *32* (6), 688–705. https://doi.org/10.1016/S0141-0229(03)00029-2.

(681) Bhatti, H. N.; Khera, R. A. Biological Transformations of Steroidal Compounds: A Review. *Steroids* **2012**, *77* (12), 1267–1290. https://doi.org/10.1016/j.steroids.2012.07.018.

(682) Peterson, D. H.; Murray, H. C.; Eppstein, S. H.; Reineke, L. M.; Weintraub, A.; Meister, P. D.; Leigh, H. M. Microbiological Transformations of Steroids. [1] I. Introduction of Oxygen at Carbon-11 of Progesterone. *J Am Chem Soc* **1952**, *74* (23), 5933–5936. https://doi.org/10.1021/ja01143a033.

(683) Peterson, D. H.; Murray, H. C. MICROBIOLOGICAL OXYGENATION OF STEROIDS AT CARBON 11. *J Am Chem Soc* **1952**, *74* (7), 1871–1872. https://doi.org/10.1021/ja01127a531.

(684) Herráiz, I. Chemical Pathways of Corticosteroids, Industrial Synthesis from Sapogenins. In *Methods in Molecular Biology*; 2017; Vol. 1645, pp 15–27. https://doi.org/10.1007/978-1-4939-7183-1_2.

(685) Kamalakkannan, V.; Salim, A. A.; Capon, R. J. Microbiome-Mediated Biotransformation of Cane Toad Bufagenins. *J Nat Prod* **2017**, *80* (7), 2012–2017. https://doi.org/10.1021/acs.jnatprod.7b00134.

(686) Ekiz, G.; Yılmaz, S.; Yusufoglu, H.; Kırmızıbayrak, P. B.; Bedir, E. Microbial Transformation of Cycloastragenol and Astragenol by Endophytic Fungi Isolated from *Astragalus* Species. *J Nat Prod* **2019**, *82* (11), 2979–2985. https://doi.org/10.1021/acs.jnatprod.9b00336.

(687) Özçınar, Ö.; Özgür Tağ; Yusufoglu, H.; Kivçak, B.; Bedir, E. Biotransformation of Neoruscogenin by the Endophytic Fungus *Alternaria Eureka*. *J Nat Prod* **2018**, *81* (6), 1357–1367. https://doi.org/10.1021/acs.jnatprod.7b00898.

(688) Küçüksolak, M.; Üner, G.; Ballar Kırmızıbayrak, P.; Bedir, E. Neuroprotective Metabolites via Fungal Biotransformation of a Novel Sapogenin, Cyclocephagenol. *Sci Rep* **2022**, *12* (1), 18481. https://doi.org/10.1038/s41598-022-22799-5.

(689) Arunrattiyakorn, P.; Kuno, M.; Aree, T.; Laphookhieo, S.; Sriyatep, T.; Kanzaki, H.; Garcia Chavez, M. A.; Wang, Y. A.; Andersen, R. J. Biotransformation of β-Mangostin by an Endophytic Fungus of *Garcinia Mangostana* to Furnish

Xanthenes with an Unprecedented Heterocyclic Skeleton. *J Nat Prod* **2018**, *81* (10), 2244–2250. https://doi.org/10.1021/acs.jnatprod.8b00519.

(690) Chen, L.-X.; Zhao, Q.; Zhang, M.; Liang, Y.-Y.; Ma, J.-H.; Zhang, X.; Ding, L.-Q.; Zhao, F.; Qiu, F. Biotransformation of Curcumenol by *Mucor Polymorphosporus*. *J Nat Prod* **2015**, *78* (4), 674–680. https://doi.org/10.1021/np500845z.

(691) Han, F.; Xiao, Y.; Lee, I.-S. Microbial Conjugation Studies of Licochalcones and Xanthohumol. *Int J Mol Sci* **2021**, *22* (13), 6893. https://doi.org/10.3390/ijms22136893.

(692) Tronina, T.; Łużny, M.; Dymarska, M.; Urbaniak, M.; Kozłowska, E.; Piegza, M.; Stępień, Ł.; Janeczko, T. Glycosylation of Quercetin by Selected Entomopathogenic Filamentous Fungi and Prediction of Its Products' Bioactivity. *Int J Mol Sci* **2023**, *24* (14), 11857. https://doi.org/10.3390/ijms241411857.

(693) Fessner, N. D.; Grimm, C.; Srdič, M.; Weber, H.; Kroutil, W.; Schwaneberg, U.; Glieder, A. Natural Product Diversification by One-Step Biocatalysis Using Human P450 3A4. *ChemCatChem* **2022**, *14* (1), e202101564. https://doi.org/10.1002/cctc.202101564.

(694) Pieper, R.; Luo, G.; Cane, D. E.; Khosla, C. Remarkably Broad Substrate Specificity of a Modular Polyketide Synthase in a Cell-Free System. *J Am Chem Soc* **1995**, *117* (45), 11373–11374. https://doi.org/10.1021/ja00150a049.

(695) Wiesmann, K. E. H.; Cortés, J.; Brown, M. J. B.; Cutter, A. L.; Staunton, J.; Leadlay, P. F. Polyketide Synthesis in Vitro on a Modular Polyketide Synthase. *Chem Biol* **1995**, *2* (9), 583–589. https://doi.org/10.1016/1074-5521(95)90122-1.

(696) Abugrain, M. E.; Brumsted, C. J.; Osborn, A. R.; Philmus, B.; Mahmud, T. A Highly Promiscuous SS-Ketoacyl-ACP Synthase (KAS) III-like Protein Is Involved in Pactamycin Biosynthesis. *ACS Chem Biol* **2017**, *12* (2), 362–366. https://doi.org/10.1021/acschembio.6b01043.

(697) Kim, M. Il; Kwon, S. J.; Dordick, J. S. In Vitro Precursor-Directed Synthesis of Polyketide Analogues with Coenzyme A Regeneration for the Development of Antiangiogenic Agents. *Org Lett* **2009**, *11* (17), 3806–3809. https://doi.org/10.1021/ol901243e.

(698) Murli, S.; MacMillan, K. S.; Hu, Z.; Ashley, G. W.; Dong, S. D.; Kealey, J. T.; Reeves, C. D.; Kennedy, J. Chemobiosynthesis of Novel 6-Deoxyerythronolide B Analogues by Mutation of the Loading Module of 6-Deoxyerythronolide B Synthase 1. *Appl Environ Microbiol* **2005**, *71* (8), 4503–4509. https://doi.org/10.1128/AEM.71.8.4503-4509.2005.

(699) Gao, Z.; Wang, J.; Norquay, A. K.; Qiao, K.; Tang, Y.; Vederas, J. C. Investigation of Fungal Iterative Polyketide Synthase Functions Using Partially Assembled Intermediates. *J Am Chem Soc* **2013**, *135* (5), 1735–1738. https://doi.org/10.1021/ja4001823.

(700) Li, Z.; Martin, F. M.; Vederas, J. C. Biosynthetic Incorporation of Labeled Tetraketide Intermediates into Dehydrocurvularin, a Phytotoxin from Alternaria Cinerariae, with Assistance of .Beta.-Oxidation Inhibitors. *J Am Chem Soc* **1992**, *114* (4), 1531–1533. https://doi.org/10.1021/ja00030a088.

(701) Auclair, K.; Sutherland, A.; Kennedy, J.; Witter, D. J.; Van den Heever, J. P.; Hutchinson, C. R.; Vederas, J. C. Lovastatin Nonaketide Synthase Catalyzes an Intramolecular Diels - Alder Reaction of a Substrate Analogue [7]. *Journal of the American Chemical Society*. 2000. https://doi.org/10.1021/ja003216+.

(702) Zhou, H.; Zhan, J.; Watanabe, K.; Xie, X.; Tang, Y. A Polyketide Macrolactone Synthase from the Filamentous Fungus *Gibberella Zeae*. *Proceedings of the National Academy of Sciences* **2008**, *105* (17), 6249–6254. https://doi.org/10.1073/pnas.0800657105.

(703) Huitt-Roehl, C. R.; Hill, E. A.; Adams, M. M.; Vagstad, A. L.; Li, J. W.; Townsend, C. A. Starter Unit Flexibility for Engineered Product Synthesis by the Nonreducing Polyketide Synthase PksA. *ACS Chem Biol* **2015**, *10* (6), 1443–1449. https://doi.org/10.1021/acschembio.5b00005.

(704) Hansen, D. A.; Koch, A. A.; Sherman, D. H. Substrate Controlled Divergence in Polyketide Synthase Catalysis. *J Am Chem Soc* **2015**, *137* (11), 3735–3738. https://doi.org/10.1021/ja511743n.

(705) Hansen, D. A.; Koch, A. A.; Sherman, D. H. Identification of a Thioesterase Bottleneck in the Pikromycin Pathway through Full-Module Processing of Unnatural Pentaketides. *J Am Chem Soc* **2017**, *139* (38), 13450–13455. https://doi.org/10.1021/jacs.7b06432.

(706) Koch, A. A.; Hansen, D. A.; Shende, V. V.; Furan, L. R.; Houk, K. N.; Jiménez-Osés, G.; Sherman, D. H. A Single Active Site Mutation in the Pikromycin Thioesterase Generates a More Effective Macrocyclization Catalyst. *J Am Chem Soc* **2017**, *139* (38), 13456–13465. https://doi.org/10.1021/jacs.7b06436.

(707) Lowell, A. N.; DeMars, M. D.; Slocum, S. T.; Yu, F.; Anand, K.; Chemler, J. A.; Korakavi, N.; Priessnitz, J. K.; Park, S. R.; Koch, A. A.; Schultz, P. J.; Sherman, D. H. Chemoenzymatic Total Synthesis and Structural Diversification of Tylactone-Based Macrolide Antibiotics through Late-Stage Polyketide Assembly, Tailoring, and C—H Functionalization. *J Am Chem Soc* **2017**, *139* (23), 7913–7920. https://doi.org/10.1021/jacs.7b02875.

(708) Dardić, D.; Böhringer, N.; Plaza, A.; Zubeil, F.; Pohl, J.; Sommer, S.; Padva, L.; Becker, J.; Patras, M. A.; Bill, M.-K.; Kurz, M.; Toti, L.; Görgens, S. W.; Schuler, S. M. M.; Billion, A.; Schwengers, O.; Wohlfart, P.; Goesmann, A.; Tennagels, N.; Vilcinskas, A.; Hammann, P. E.; Schäberle, T. F.; Bauer, A. Antidiabetic Profiling of Veramycins, Polyketides Accessible by Biosynthesis, Chemical Synthesis and Precursor-Directed Modification. *Organic Chemistry Frontiers* **2022**, *9* (6), 1604–1615. https://doi.org/10.1039/D1QO01652K.

(709) Bravo-Rodriguez, K.; Ismail-Ali, A. F.; Klopries, S.; Kushnir, S.; Ismail, S.; Fansa, E. K.; Wittinghofer, A.; Schulz, F.; Sanchez-Garcia, E. Predicted Incorporation of Non-native Substrates by a Polyketide Synthase Yields Bioactive Natural Product Derivatives. *ChemBioChem* **2014**, *15* (13), 1991–1997. https://doi.org/10.1002/cbic.201402206.

(710) Tian, W.; Chen, X.; Zhang, J.; Zheng, M.; Wei, G.; Deng, Z.; Qu, X. Biosynthesis of Tetronates by a Nonribosomal Peptide Synthetase–Polyketide Synthase System. *Org Lett* **2023**, *25* (10), 1628–1632. https://doi.org/10.1021/acs.orglett.3c00103.

(711) An, J. H.; Kim, Y. S. A Gene Cluster Encoding Malonyl-CoA Decarboxylase (MatA), Malonyl-CoA Synthetase (MatB) and a Putative Dicarboxylate Carrier Protein (MatC) in *Rhizobium Trifolii*. *Eur J Biochem* **1998**, *257* (2), 395–402. https://doi.org/10.1046/j.1432-1327.1998.2570395.x.

(712) Koryakina, I.; McArthur, J.; Randall, S.; Draelos, M. M.; Musiol, E. M.; Muddiman, D. C.; Weber, T.; Williams, G. J. Poly Specific *Trans* -Acyltransferase Machinery Revealed *via* Engineered Acyl-CoA Synthetases. *ACS Chem Biol* **2013**, *8* (1), 200–208. https://doi.org/10.1021/cb3003489.

(713) Ad, O.; Thuronyi, B. W.; Chang, M. C. Y. Elucidating the Mechanism of Fluorinated Extender Unit Loading for Improved Production of Fluorine-Containing Polyketides. *Proceedings of the National Academy of Sciences* **2017**, *114* (5), E660–E668. https://doi.org/10.1073/pnas.1614196114.

(714) Grote, M.; Schulz, F. Exploring the Promiscuous Enzymatic Activation of Unnatural Polyketide Extender Units in Vitro and in Vivo for Monensin Biosynthesis. *ChemBioChem* **2019**, *20* (9), 1183–1189. https://doi.org/10.1002/cbic.201800734.

(715) Koryakina, I.; McArthur, J. B.; Draelos, M. M.; Williams, G. J. Promiscuity of a Modular Polyketide Synthase towards Natural and Non-Natural Extender Units. *Org Biomol Chem* **2013**, *11* (27), 4449. https://doi.org/10.1039/c3ob40633d.

(716) Sun, C.; Ma, B.; Li, G.; Tian, W.; Yang, L.; Peng, H.; Lin, Z.; Deng, Z.; Kong, X.; Qu, X. Engineering the Substrate Specificity of a P450 Dimerase Enables the Collective Biosynthesis of Heterodimeric Tryptophan-Containing Diketopiperazines. *Angewandte Chemie International Edition* **2023**, *62* (25), e202304994. https://doi.org/10.1002/anie.202304994.

(717) Musiol-Kroll, E. M.; Zubeil, F.; Schafhauser, T.; Härtner, T.; Kulik, A.; McArthur, J.; Koryakina, I.; Wohlleben, W.; Grond, S.; Williams, G. J.; Lee, S. Y.; Weber, T. Polyketide Bioderivatization Using the Promiscuous Acyltransferase KirCII. *ACS Synth Biol* **2017**, *6* (3), 421–427. https://doi.org/10.1021/acssynbio.6b00341.

(718) Koryakina, I.; Kasey, C.; McArthur, J. B.; Lowell, A. N.; Chemler, J. A.; Li, S.; Hansen, D. A.; Sherman, D. H.; Williams, G. J. Inversion of Extender Unit Selectivity in the Erythromycin Polyketide Synthase by Acyltransferase Domain Engineering. *ACS Chem Biol* **2017**, *12* (1), 114–123. https://doi.org/10.1021/acschembio.6b00732.

(719) Kalkreuter, E.; Bingham, K. S.; Keeler, A. M.; Lowell, A. N.; Schmidt, J. J.; Sherman, D. H.; Williams, G. J. Computationally-Guided Exchange of Substrate Selectivity Motifs in a Modular Polyketide Synthase Acyltransferase. *Nat Commun* **2021**, *12* (1), 2193. https://doi.org/10.1038/s41467-021-22497-2.

(720) Hughes, A. J.; Keatinge-Clay, A. Enzymatic Extender Unit Generation for In Vitro Polyketide Synthase Reactions: Structural and Func-Tional Showcasing of Streptomyces Coelicolor MatB. *Chem Biol* **2011**, *18* (2), 165–176. https://doi.org/10.1016/j.chembiol.2010.12.014.

(721) Zhang, F.; Ji, H.; Ali, I.; Deng, Z.; Bai, L.; Zheng, J. Structural and Biochemical Insight into the Recruitment of Acyl Carrier Protein-Linked Extender Units in Ansamitocin Biosynthesis. *ChemBioChem* **2020**, *21* (9), 1309–1314. https://doi.org/10.1002/cbic.201900628.

(722) Hughes, A. J.; Tibby, M. R.; Wagner, D. T.; Brantley, J. N.; Keatinge-Clay, A. T. Investigating the Reactivities of a Polyketide Synthase Module through Fluorescent Click Chemistry. *Chem. Commun.* **2014**, *50* (40), 5276–5278. https://doi.org/10.1039/C3CC47513A.

(723) Walker, M. C.; Thuronyi, B. W.; Charkoudian, L. K.; Lowry, B.; Khosla, C.; Chang, M. C. Y. Expanding the Fluorine Chemistry of Living Systems Using Engineered Polyketide Synthase Pathways. *Science (1979)* **2013**, *341* (6150), 1089–1094. https://doi.org/10.1126/science.1242345.

(724) Erb, T. J.; Berg, I. A.; Brecht, V.; Müller, M.; Fuchs, G.; Alber, B. E. Synthesis of C_5 -Dicarboxylic Acids from C_2 -Units Involving Crotonyl-CoA Carboxylase/Reductase: The Ethylmalonyl-CoA Pathway. *Proceedings of the National Academy of Sciences* **2007**, *104* (25), 10631–10636. https://doi.org/10.1073/pnas.0702791104.

(725) Li, C.; Roege, K. E.; Kelly, W. L. Analysis of the Indanomycin Biosynthetic Gene Cluster from *Streptomyces Antibioticus* NRRL 8167. *ChemBioChem* **2009**, *10* (6), 1064–1072. https://doi.org/10.1002/cbic.200800822.

(726) Mo, S.; Kim, D. H.; Lee, J. H.; Park, J. W.; Basnet, D. B.; Ban, Y. H.; Yoo, Y. J.; Chen, S.; Park, S. R.; Choi, E. A.; Kim, E.; Jin, Y.-Y.; Lee, S.-K.; Park, J. Y.; Liu, Y.; Lee, M. O.; Lee, K. S.; Kim, S. J.; Kim, D.; Park, B. C.; Lee, S.; Kwon, H. J.; Suh, J.-W.; Moore, B. S.; Lim, S.-K.; Yoon, Y. J. Biosynthesis of the Allylmalonyl-CoA Extender Unit for the FK506 Polyketide Synthase Proceeds through a Dedicated Polyketide Synthase and Facilitates the Mutasynthesis of Analogues. *J Am Chem Soc* **2011**, *133* (4), 976–985. https://doi.org/10.1021/ja108399b.

(727) Eustáquio, A. S.; McGlinchey, R. P.; Liu, Y.; Hazzard, C.; Beer, L. L.; Florova, G.; Alhamadsheh, M. M.; Lechner, A.; Kale, A. J.; Kobayashi, Y.; Reynolds, K. A.; Moore, B. S. Biosynthesis of the Salinosporamide A Polyketide Synthase Substrate Chloroethylmalonyl-Coenzyme A from *S* -Adenosyl- <scp>l</Scp> -Methionine. *Proceedings of the National Academy of Sciences* **2009**, *106* (30), 12295–12300. https://doi.org/10.1073/pnas.0901237106.

(728) Zhang, L.; Mori, T.; Zheng, Q.; Awakawa, T.; Yan, Y.; Liu, W.; Abe, I. Rational Control of Polyketide Extender Units by Structure-Based Engineering of a Crotonyl-CoA Carboxylase/Reductase in Antimycin Biosynthesis. *Angewandte Chemie International Edition* **2015**, *54* (45), 13462–13465. https://doi.org/10.1002/anie.201506899.

(729) Ray, L.; Valentic, T. R.; Miyazawa, T.; Withall, D. M.; Song, L.; Milligan, J. C.; Osada, H.; Takahashi, S.; Tsai, S.-C.; Challis, G. L. A Crotonyl-CoA Reductase-Carboxylase Independent Pathway for Assembly of Unusual Alkylmalonyl-CoA Polyketide Synthase Extender Units. *Nat Commun* **2016**, *7* (1), 13609. https://doi.org/10.1038/ncomms13609.

(730) Diacovich, L.; Peirú, S.; Kurth, D.; Rodríguez, E.; Podestá, F.; Khosla, C.; Gramajo, H. Kinetic and Structural Analysis of a New Group of Acyl-CoA Carboxylases Found in Streptomyces Coelicolor A3(2). *Journal of Biological Chemistry* **2002**, *277* (34), 31228–31236. https://doi.org/10.1074/jbc.M203263200.

(731) Zhang, J.; Zheng, M.; Yan, J.; Deng, Z.; Zhu, D.; Qu, X. A Permissive Medium Chain Acyl-CoA Carboxylase Enables the Efficient Biosynthesis of Extender Units for Engineering Polyketide Carbon Scaffolds. *ACS Catal* **2021**, *11* (19), 12179–12185. https://doi.org/10.1021/acscatal.1c03818.

(732) Zhu, X.; Liu, J.; Zhang, W. De Novo Biosynthesis of Terminal Alkyne-Labeled Natural Products. *Nat Chem Biol* **2015**, *11* (2), 115–120. https://doi.org/10.1038/nchembio.1718.

(733) Porterfield, W. B.; Poenateetai, N.; Zhang, W. Engineered Biosynthesis of Alkyne-Tagged Polyketides by Type I PKSs. *iScience* **2020**, *23* (3), 100938. https://doi.org/10.1016/j.isci.2020.100938.

(734) Carpenter, S. M.; Williams, G. J. Extender Unit Promiscuity and Orthogonal Protein Interactions of an Aminomalonyl-ACP Utilizing Trans-Acyltransferase from Zwittermicin Biosynthesis. *ACS Chem Biol* **2018**, *13* (12), 3361–3373. https://doi.org/10.1021/acschembio.8b00867.

(735) Rittner, A.; Joppe, M.; Schmidt, J. J.; Mayer, L. M.; Reiners, S.; Heid, E.; Herzberg, D.; Sherman, D. H.; Grininger, M. Chemoenzymatic Synthesis of Fluorinated Polyketides. *Nat Chem* **2022**, *14* (9), 1000–1006. https://doi.org/10.1038/s41557-022-00996-z.

(736) Sirirungruang, S.; Ad, O.; Privalsky, T. M.; Ramesh, S.; Sax, J. L.; Dong, H.; Baidoo, E. E. K.; Amer, B.; Khosla, C.; Chang, M. C. Y. Engineering Site-Selective Incorporation of Fluorine into Polyketides. *Nat Chem Biol* **2022**, *18* (8), 886–893. https://doi.org/10.1038/s41589-022-01070-y.

(737) TRABER, R.; HOFMANN, H.; KOBEL, H. Cyclosporins. New Analogues by Precursor Directed Biosynthesis. *J Antibiot (Tokyo)* **1989**, *42* (4), 591–597. https://doi.org/10.7164/antibiotics.42.591.

(738) Ragab, A. E.; Grüschow, S.; Rackham, E. J.; Goss, R. J. M. New Pacidamycins Biosynthetically: Probing N- and C-Terminal Substrate Specificity. *Org Biomol Chem* **2010**, *8* (14), 3128. https://doi.org/10.1039/c003964k.

(739) Jiang, G.; Zuo, R.; Zhang, Y.; Powell, M. M.; Zhang, P.; Hylton, S. M.; Loria, R.; Ding, Y. One-Pot Biocombinatorial Synthesis of Herbicidal Thaxtomins. *ACS Catal* **2018**, *8* (11), 10761–10768. https://doi.org/10.1021/acscatal.8b03317.

(740) Guo, H.; Schmidt, A.; Stephan, P.; Raguž, L.; Braga, D.; Kaiser, M.; Dahse, H.; Weigel, C.; Lackner, G.; Beemelmanns, C. Precursor-Directed Diversification of Cyclic Tetrapeptidic Pseudoxylallemycins. *ChemBioChem* **2018**, *19* (21), 2307–2311. https://doi.org/10.1002/cbic.201800503.

(741) Zhu, M.; Wang, L.; He, J. Chemical Diversification Based on Substrate Promiscuity of a Standalone Adenylation Domain in a Reconstituted NRPS System. *ACS Chem Biol* **2019**, *14* (2), 256–265. https://doi.org/10.1021/acschembio.8b00938.

(742) Moschny, J.; Lorenzen, W.; Hilfer, A.; Eckenstaler, R.; Jahns, S.; Enke, H.; Enke, D.; Schneider, P.; Benndorf, R. A.; Niedermeyer, T. H. J. Precursor-Directed Biosynthesis and Fluorescence Labeling of Clickable Microcystins. *J Nat Prod* **2020**, *83* (6), 1960–1970. https://doi.org/10.1021/acs.jnatprod.0c00251.

(743) Meyer, S.; Kehr, J.-C.; Mainz, A.; Dehm, D.; Petras, D.; Süssmuth, R. D.; Dittmann, E. Biochemical Dissection of the Natural Diversification of Microcystin Provides Lessons for Synthetic Biology of NRPS. *Cell Chem Biol* **2016**, *23* (4), 462–471. https://doi.org/10.1016/j.chembiol.2016.03.011.

(744) Kudo, F.; Takahashi, S.; Miyanaga, A.; Nakazawa, Y.; Nishino, K.; Hayakawa, Y.; Kawamura, K.; Ishikawa, F.; Tanabe, G.; Iwai, N.; Nagumo, Y.; Usui, T.; Eguchi, T. Mutational Biosynthesis of Hitachimycin Analogs Controlled by the β-Amino Acid–Selective Adenylation Enzyme HitB. *ACS Chem Biol* **2021**, *16* (3), 539–547. https://doi.org/10.1021/acschembio.1c00003.

(745) Yao, Z.; Sun, C.; Xia, Y.; Wang, F.; Fu, L.; Ma, J.; Li, Q.; Ju, J. Mutasynthesis of Antibacterial Halogenated Actinomycin Analogues. *J Nat Prod* **2021**, *84* (8), 2217–2225. https://doi.org/10.1021/acs.jnatprod.1c00294.

(746) Wang, Y.; He, J.; Alam, M. S.; Wang, F.; Shang, Z.; Chen, Y.; Sun, C.; Lu, Z.; Gao, Y.; Zhang, T.; Ju, J.; Ma, J. Efficient Mutasynthesis of "Non-Natural" Antitubercular Ilamycins with Low Cytotoxicity. *ACS Synth Biol* **2024**, *13* (3), 930–941. https://doi.org/10.1021/acssynbio.3c00730.

(747) Trauger, J. W.; Kohli, R. M.; Mootz, H. D.; Marahiel, M. A.; Walsh, C. T. Peptide Cyclization Catalysed by the Thioesterase Domain of Tyrocidine Synthetase. *Nature* **2000**, *407* (6801), 215–218. https://doi.org/10.1038/35025116.

(748) Beck, Z. Q.; Aldrich, C. C.; Magarvey, N. A.; Georg, G. I.; Sherman, D. H. Chemoenzymatic Synthesis of Cryptophycin/Arenastatin Natural Products. *Biochemistry* **2005**, *44* (41), 13457–13466. https://doi.org/10.1021/bi051140u.

(749) Grünewald, J.; Sieber, S. A.; Mahlert, C.; Linne, U.; Marahiel, M. A. Synthesis and Derivatization of Daptomycin: A Chemoenzymatic Route to Acidic Lipopeptide Antibiotics. *J Am Chem Soc* **2004**, *126* (51), 17025–17031. https://doi.org/10.1021/ja045455t.

(750) Lin, H.; Walsh, C. T. A Chemoenzymatic Approach to Glycopeptide Antibiotics. *J Am Chem Soc* **2004**, *126* (43), 13998–14003. https://doi.org/10.1021/ja045147v.

(751) Kobayashi, M.; Onozawa, N.; Matsuda, K.; Wakimoto, T. Chemoenzymatic Tandem Cyclization for the Facile Synthesis of Bicyclic Peptides. *Commun Chem* **2024**, *7* (1), 67. https://doi.org/10.1038/s42004-024-01147-w.

(752) Budimir, Z. L.; Patel, R. S.; Eggly, A.; Evans, C. N.; Rondon-Cordero, H. M.; Adams, J. J.; Das, C.; Parkinson, E. I. Biocatalytic Cyclization of Small Macrolactams by a Penicillin-Binding Protein-Type Thioesterase. *Nat Chem Biol* **2024**, *20* (1), 120–128. https://doi.org/10.1038/s41589-023-01495-z.

(753) Challis, G. L.; Ravel, J.; Townsend, C. A. Predictive, Structure-Based Model of Amino Acid Recognition by Nonribosomal Peptide Synthetase Adenylation Domains. *Chem Biol* **2000**, *7* (3), 211–224. https://doi.org/10.1016/S1074-5521(00)00091-0.

(754) Heard, S. C.; Winter, J. M. Structural, Biochemical and Bioinformatic Analyses of Nonribosomal Peptide Synthetase Adenylation Domains. *Nat Prod Rep* **2024**, *41* (7), 1180–1205. https://doi.org/10.1039/D3NP00064H.

(755) Thirlway, J.; Lewis, R.; Nunns, L.; Al Nakeeb, M.; Styles, M.; Struck, A.; Smith, C. P.; Micklefield, J. Introduction of a Non-Natural Amino Acid into a Nonribosomal Peptide Antibiotic by Modification of Adenylation Domain Specificity. *Angewandte Chemie International Edition* **2012**, *51* (29), 7181–7184. https://doi.org/10.1002/anie.201202043.

(756) Wang, D.; Miyanaga, A.; Chisuga, T.; Kudo, F.; Eguchi, T. Engineering the Substrate Specificity of (*S*)-β-Phenylalanine Adenylation Enzyme HitB. *ChemBioChem* **2024**, *25* (15), e202400383. https://doi.org/10.1002/cbic.202400383.

(757) Zhang, K.; Nelson, K. M.; Bhuripanyo, K.; Grimes, K. D.; Zhao, B.; Aldrich, C. C.; Yin, J. Engineering the Substrate Specificity of the DhbE Adenylation Domain by Yeast Cell Surface Display. *Chem Biol* **2013**, *20* (1), 92–101. https://doi.org/10.1016/j.chembiol.2012.10.020.

(758) Ishikawa, F.; Miyanaga, A.; Kitayama, H.; Nakamura, S.; Nakanishi, I.; Kudo, F.; Eguchi, T.; Tanabe, G. An Engineered Aryl Acid Adenylation Domain with an Enlarged Substrate Binding Pocket. *Angewandte Chemie International Edition* **2019**, *58* (21), 6906–6910. https://doi.org/10.1002/anie.201900318.

(759) Messenger, S. R.; McGuinniety, E. M. R.; Stevenson, L. J.; Owen, J. G.; Challis, G. L.; Ackerley, D. F.; Calcott, M. J. Metagenomic Domain Substitution for the High-Throughput Modification of Nonribosomal Peptides. *Nat Chem Biol* **2024**, *20* (2), 251–260. https://doi.org/10.1038/s41589-023-01485-1.

(760) Charlop-Powers, Z.; Owen, J. G.; Reddy, B. V. B.; Ternei, M. A.; Guimarães, D. O.; de Frias, U. A.; Pupo, M. T.; Seepe, P.; Feng, Z.; Brady, S. F. Global Biogeographic Sampling of Bacterial Secondary Metabolism. *Elife* **2015**, *4* (4), e05048. https://doi.org/10.7554/eLife.05048.

(761) Charlop-Powers, Z.; Owen, J. G.; Reddy, B. V. B.; Ternei, M. A.; Brady, S. F. Chemical-Biogeographic Survey of Secondary Metabolism in Soil. *Proceedings of the National Academy of Sciences* **2014**, *111* (10), 3757–3762. https://doi.org/10.1073/pnas.1318021111.

(762) Brown, A. S.; Calcott, M. J.; Owen, J. G.; Ackerley, D. F. Structural, Functional and Evolutionary Perspectives on Effective Re-Engineering of Non-Ribosomal Peptide Synthetase Assembly Lines. *Nat Prod Rep* **2018**, *35* (11), 1210–1228. https://doi.org/10.1039/C8NP00036K.

(763) Bloudoff, K.; Schmeing, T. M. Structural and Functional Aspects of the Nonribosomal Peptide Synthetase Condensation Domain Superfamily: Discovery, Dissection and Diversity. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2017**, *1865* (11), 1587–1604. https://doi.org/10.1016/j.bbapap.2017.05.010.

(764) Dekimpe, S.; Masschelein, J. Beyond Peptide Bond Formation: The Versatile Role of Condensation Domains in Natural Product Biosynthesis. *Nat Prod Rep* **2021**, *38* (10), 1910–1937. https://doi.org/10.1039/D0NP00098A.

(765) Luo, L.; Kohli, R. M.; Onishi, M.; Linne, U.; Marahiel, M. A.; Walsh, C. T. Timing of Epimerization and Condensation Reactions in Nonribosomal Peptide Assembly Lines: Kinetic Analysis of Phenylalanine Activating Elongation Modules of Tyrocidine Synthetase B. *Biochemistry* **2002**, *41* (29), 9184–9196. https://doi.org/10.1021/bi026047+.

(766) Stachelhaus, T.; Mootz, H. D.; Bergendahl, V.; Marahiel, M. A. Peptide Bond Formation in Nonribosomal Peptide Biosynthesis. *Journal of Biological Chemistry* **1998**, *273* (35), 22773–22781. https://doi.org/10.1074/jbc.273.35.22773.

(767) Bloudoff, K.; Alonzo, D. A.; Schmeing, T. M. Chemical Probes Allow Structural Insight into the Condensation Reaction of Nonribosomal Peptide Synthetases. *Cell Chem Biol* **2016**, *23* (3), 331–339. https://doi.org/10.1016/j.chembiol.2016.02.012.

(768) Ackerley, D. F.; Lamont, I. L. Characterization and Genetic Manipulation of Peptide Synthetases in Pseudomonas Aeruginosa PAO1 in Order to Generate Novel Pyoverdines. *Chem Biol* **2004**, *11* (7), 971–980. https://doi.org/10.1016/j.chembiol.2004.04.014.

(769) Doekel, S.; Marahiel, M. A. Dipeptide Formation on Engineered Hybrid Peptide Synthetases. *Chem Biol* **2000**, *7* (6), 373–384. https://doi.org/10.1016/S1074-5521(00)00118-6.

(770) Kaniusaite, M.; Tailhades, J.; Marschall, E. A.; Goode, R. J. A.; Schittenhelm, R. B.; Cryle, M. J. A Proof-Reading Mechanism for Non-Proteinogenic Amino Acid Incorporation into Glycopeptide Antibiotics. *Chem Sci* **2019**, *10* (41), 9466–9482. https://doi.org/10.1039/C9SC03678D.

(771) Ehmann, D. E.; Trauger, J. W.; Stachelhaus, T.; Walsh, C. T. Aminoacyl-SNACs as Small-Molecule Substrates for the Condensation Domains of Nonribosomal Peptide Synthetases. *Chem Biol* **2000**, *7* (10), 765–772. https://doi.org/10.1016/S1074-5521(00)00022-3.

(772) Belshaw, P. J.; Walsh, C. T.; Stachelhaus, T. Aminoacyl-CoAs as Probes of Condensation Domain Selectivity in Nonribosomal Peptide Synthesis. *Science (1979)* **1999**, *284* (5413), 486–489. https://doi.org/10.1126/science.284.5413.486.

(773) Kraas, F. I.; Giessen, T. W.; Marahiel, M. A. Exploring the Mechanism of Lipid Transfer during Biosynthesis of the Acidic Lipopeptide Antibiotic CDA. *FEBS Lett* **2012**, *586* (3), 283–288. https://doi.org/10.1016/j.febslet.2012.01.003.

(774) Camus, A.; Truong, G.; Mittl, P. R. E.; Markert, G.; Hilvert, D. Reprogramming Nonribosomal Peptide Synthetases for Site-Specific Insertion of α-Hydroxy Acids. *J Am Chem Soc* **2022**, *144* (38), 17567–17575. https://doi.org/10.1021/jacs.2c07013.

(775) Izoré, T.; Cryle, M. J. The Many Faces and Important Roles of Protein–Protein Interactions during Non-Ribosomal Peptide Synthesis. *Nat Prod Rep* **2018**, *35* (11), 1120–1139. https://doi.org/10.1039/C8NP00038G.

(776) Boder, E. T.; Wittrup, K. D. [25] Yeast Surface Display for Directed Evolution of Protein Expression, Affinity, and Stability. In *Methods in Enzymology*; 2000; Vol. 328, pp 430–444. https://doi.org/10.1016/S0076-6879(00)28410-3.

(777) Folger, I. B.; Frota, N. F.; Pistofidis, A.; Niquille, D. L.; Hansen, D. A.; Schmeing, T. M.; Hilvert, D. High-Throughput Reprogramming of an NRPS Condensation Domain. *Nat Chem Biol* **2024**, *20* (6), 761–769. https://doi.org/10.1038/s41589-023-01532-x.

(778) Chen, H.; Zhong, L.; Zhou, H.; Bai, X.; Sun, T.; Wang, X.; Zhao, Y.; Ji, X.; Tu, Q.; Zhang, Y.; Bian, X. Biosynthesis and Engineering of the Nonribosomal Peptides with a C-Terminal Putrescine. *Nat Commun* **2023**, *14* (1), 6619. https://doi.org/10.1038/s41467-023-42387-z.

(779) Nguyen, D. T.; Le, T. T.; Rice, A. J.; Hudson, G. A.; van der Donk, W. A.; Mitchell, D. A. Accessing Diverse Pyridine-Based Macrocyclic Peptides by a Two-Site Recognition Pathway. *J Am Chem Soc* **2022**, *144* (25), 11263–11269. https://doi.org/10.1021/jacs.2c02824.

(780) Zhao, Y.; Marschall, E.; Treisman, M.; McKay, A.; Padva, L.; Crüsemann, M.; Nelson, D. R.; Steer, D. L.; Schittenhelm, R. B.; Tailhades, J.; Cryle, M. J. Cytochrome P450 $_{Blt}$ Enables Versatile Peptide Cyclisation to Generate Histidine- and Tyrosine-Containing Crosslinked Tripeptide Building Blocks. *Angewandte Chemie International Edition* **2022**, *61* (37), e202204957. https://doi.org/10.1002/anie.202204957.

(781) Nam, H.; An, J. S.; Lee, J.; Yun, Y.; Lee, H.; Park, H.; Jung, Y.; Oh, K.-B.; Oh, D.-C.; Kim, S. Exploring the Diverse Landscape of Biaryl-Containing Peptides Generated by Cytochrome P450 Macrocyclases. *J Am Chem Soc* **2023**, *145* (40), 22047–22057. https://doi.org/10.1021/jacs.3c07140.

(782) Sugiyama, R.; Suarez, A. F. L.; Morishita, Y.; Nguyen, T. Q. N.; Tooh, Y. W.; Roslan, M. N. H. Bin; Lo Choy, J.; Su, Q.; Goh, W. Y.; Gunawan, G. A.; Wong, F. T.; Morinaka, B. I. The Biosynthetic Landscape of Triceptides Reveals Radical SAM Enzymes That Catalyze Cyclophane Formation on Tyr- and His-Containing Motifs. *J Am Chem Soc* **2022**, *144* (26), 11580–11593. https://doi.org/10.1021/jacs.2c00521.

(783) Schramma, K. R.; Seyedsayamdost, M. R. Lysine-Tryptophan-Crosslinked Peptides Produced by Radical SAM Enzymes in Pathogenic Streptococci. *ACS Chem Biol* **2017**, *12* (4), 922–927. https://doi.org/10.1021/acschembio.6b01069.

(784) Ma, S.; Chen, H.; Li, H.; Ji, X.; Deng, Z.; Ding, W.; Zhang, Q. Post-Translational Formation of Aminomalonate by a Promiscuous Peptide-Modifying Radical SAM Enzyme. *Angewandte Chemie International Edition* **2021**, *60* (36), 19957–19964. https://doi.org/10.1002/anie.202107192.

(785) Nguyen, N. A.; Cong, Y.; Hurrell, R. C.; Arias, N.; Garg, N.; Puri, A. W.; Schmidt, E. W.; Agarwal, V. A Silent Biosynthetic Gene Cluster from a Methanotrophic Bacterium Potentiates Discovery of a Substrate Promiscuous Proteusin Cyclodehydratase. *ACS Chem Biol* **2022**, *17* (6), 1577–1585. https://doi.org/10.1021/acschembio.2c00251.

(786) Zhao, X.; Kuipers, O. P. Nisin- and Ripcin-Derived Hybrid Lanthipeptides Display Selective Antimicrobial Activity against *Staphylococcus Aureus*. *ACS Synth Biol* **2021**, *10* (7), 1703–1714. https://doi.org/10.1021/acssynbio.1c00080.

(787) Arias-Orozco, P.; Inklaar, M.; Lanooij, J.; Cebrián, R.; Kuipers, O. P. Functional Expression and Characterization of the Highly Promiscuous Lanthipeptide Synthetase SyncM, Enabling the Production of Lanthipeptides with a Broad Range of Ring Topologies. *ACS Synth Biol* **2021**, *10* (10), 2579–2591. https://doi.org/10.1021/acssynbio.1c00224.

(788) Walker, J. A.; Hamlish, N.; Tytla, A.; Brauer, D. D.; Francis, M. B.; Schepartz, A. Redirecting RiPP Biosynthetic Enzymes to Proteins and Backbone-Modified Substrates. *ACS Cent Sci* **2022**, *8* (4), 473–482. https://doi.org/10.1021/acscentsci.1c01577.

(789) Korneli, M.; Fuchs, S. W.; Felder, K.; Ernst, C.; Zinsli, L. V.; Piel, J. Promiscuous Installation of <scp>d</Scp> -Amino Acids in Gene-Encoded Peptides. *ACS Synth Biol* **2021**, *10* (2), 236–242. https://doi.org/10.1021/acssynbio.0c00470.

(790) Sarkar, S.; Gu, W.; Schmidt, E. W. Applying Promiscuous RiPP Enzymes to Peptide Backbone *N* -Methylation Chemistry. *ACS Chem Biol* **2022**, *17* (8), 2165–2178. https://doi.org/10.1021/acschembio.2c00293.

(791) Matabaro, E.; Witte, L.; Gherlone, F.; Vogt, E.; Kaspar, H.; Künzler, M. Promiscuity of Omphalotin A Biosynthetic Enzymes Allows *de Novo* Production of Non-Natural Multiply Backbone N-Methylated Peptide Macrocycles in Yeast. *ChemBioChem* **2024**, *25* (3), e202300626. https://doi.org/10.1002/cbic.202300626.

(792) Zhang, Y.; Li, K.; Yang, G.; McBride, J. L.; Bruner, S. D.; Ding, Y. A Distributive Peptide Cyclase Processes Multiple Microviridin Core Peptides within a Single Polypeptide Substrate. *Nat Commun* **2018**, *9* (1), 1780. https://doi.org/10.1038/s41467-018-04154-3.

(793) Yang, X.; Lennard, K. R.; He, C.; Walker, M. C.; Ball, A. T.; Doigneaux, C.; Tavassoli, A.; van der Donk, W. A. A Lanthipeptide Library Used to Identify a Protein–Protein Interaction Inhibitor. *Nat Chem Biol* **2018**, *14* (4), 375–380. https://doi.org/10.1038/s41589-018-0008-5.

(794) Vinogradov, A. A.; Shimomura, M.; Goto, Y.; Ozaki, T.; Asamizu, S.; Sugai, Y.; Suga, H.; Onaka, H. Minimal Lactazole Scaffold for in Vitro Thiopeptide Bioengineering. *Nat Commun* **2020**, *11* (1), 2272. https://doi.org/10.1038/s41467-020-16145-4.

(795) Ozaki, T.; Yamashita, K.; Goto, Y.; Shimomura, M.; Hayashi, S.; Asamizu, S.; Sugai, Y.; Ikeda, H.; Suga, H.; Onaka, H. Dissection of Goadsporin Biosynthesis by in Vitro Reconstitution Leading to Designer Analogues Expressed in Vivo. *Nat Commun* **2017**, *8* (1), 14207. https://doi.org/10.1038/ncomms14207.

(796) Fleming, S. R.; Bartges, T. E.; Vinogradov, A. A.; Kirkpatrick, C. L.; Goto, Y.; Suga, H.; Hicks, L. M.; Bowers, A. A. Flexizyme-Enabled Benchtop Biosynthesis of Thiopeptides. *J Am Chem Soc* **2019**, *141* (2), 758–762. https://doi.org/10.1021/jacs.8b11521.

(797) Li, K.; Condurso, H. L.; Li, G.; Ding, Y.; Bruner, S. D. Structural Basis for Precursor Protein–Directed Ribosomal Peptide Macrocyclization. *Nat Chem Biol* **2016**, *12* (11), 973–979. https://doi.org/10.1038/nchembio.2200.

(798) Burkhart, B. J.; Kakkar, N.; Hudson, G. A.; van der Donk, W. A.; Mitchell, D. A. Chimeric Leader Peptides for the Generation of Non-Natural Hybrid RiPP Products. *ACS Cent Sci* **2017**, *3* (6), 629–638. https://doi.org/10.1021/acscentsci.7b00141.

(799) Tianero, Ma. D. B.; Donia, M. S.; Young, T. S.; Schultz, P. G.; Schmidt, E. W. Ribosomal Route to Small-Molecule Diversity. *J Am Chem Soc* **2012**, *134* (1), 418–425. https://doi.org/10.1021/ja208278k.

(800) Sgambelluri, R. M.; Smith, M. O.; Walton, J. D. Versatility of Prolyl Oligopeptidase B in Peptide Macrocyclization. *ACS Synth Biol* **2018**, *7* (1), 145–152. https://doi.org/10.1021/acssynbio.7b00264.

(801) Ludewig, H.; Czekster, C. M.; Oueis, E.; Munday, E. S.; Arshad, M.; Synowsky, S. A.; Bent, A. F.; Naismith, J. H. Characterization of the Fast and Promiscuous Macrocyclase from Plant PCY1 Enables the Use of Simple Substrates. *ACS Chem Biol* **2018**, *13* (3), 801–811. https://doi.org/10.1021/acschembio.8b00050.

(802) Czekster, C. M.; Ludewig, H.; McMahon, S. A.; Naismith, J. H. Characterization of a Dual Function Macrocyclase Enables Design and Use of Efficient Macrocyclization Substrates. *Nat Commun* **2017**, *8* (1), 1045. https://doi.org/10.1038/s41467-017-00862-4.

(803) Schmitt, S.; Montalbán-López, M.; Peterhoff, D.; Deng, J.; Wagner, R.; Held, M.; Kuipers, O. P.; Panke, S. Analysis of Modular Bioengineered Antimicrobial Lanthipeptides at Nanoliter Scale. *Nat Chem Biol* **2019**, *15* (5), 437–443. https://doi.org/10.1038/s41589-019-0250-5.

(804) Precord, T. W.; Mahanta, N.; Mitchell, D. A. Reconstitution and Substrate Specificity of the Thioether-Forming Radical *S* -Adenosylmethionine Enzyme in Freyrasin Biosynthesis. *ACS Chem Biol* **2019**, *14* (9), 1981–1989. https://doi.org/10.1021/acschembio.9b00457.

(805) Eastman, K. A. S.; Kincannon, W. M.; Bandarian, V. Leveraging Substrate Promiscuity of a Radical *S* -Adenosyl-<scp>L</Scp> -Methionine RiPP Maturase toward Intramolecular Peptide Cross-Linking Applications. *ACS Cent Sci* **2022**, *8* (8), 1209–1217. https://doi.org/10.1021/acscentsci.2c00501.

(806) Groß, S.; Panter, F.; Pogorevc, D.; Seyfert, C. E.; Deckarm, S.; Bader, C. D.; Herrmann, J.; Müller, R. Improved Broad-Spectrum Antibiotics against Gram-Negative Pathogens *via* Darobactin Biosynthetic Pathway Engineering. *Chem Sci* **2021**, *12* (35), 11882–11893. https://doi.org/10.1039/D1SC02725E.

(807) Woodard, A. M.; Peccati, F.; Navo, C. D.; Jiménez-Osés, G.; Mitchell, D. A. Darobactin Substrate Engineering and Computation Show Radical Stability Governs Ether versus C–C Bond Formation. *J Am Chem Soc* **2024**, *146* (20), 14328–14340. https://doi.org/10.1021/jacs.4c03994.

(808) Morinaka, B. I.; Lakis, E.; Verest, M.; Helf, M. J.; Scalvenzi, T.; Vagstad, A. L.; Sims, J.; Sunagawa, S.; Gugger, M.; Piel, J. Natural Noncanonical Protein Splicing Yields Products with Diverse β-Amino Acid Residues. *Science (1979)* **2018**, *359* (6377), 779–782. https://doi.org/10.1126/science.aao0157.

(809) Scott, T. A.; Verest, M.; Farnung, J.; Forneris, C. C.; Robinson, S. L.; Ji, X.; Hubrich, F.; Chepkirui, C.; Richter, D. U.; Huber, S.; Rust, P.; Streiff, A. B.; Zhang, Q.; Bode, J. W.; Piel, J. Widespread Microbial Utilization of Ribosomal β-

Amino Acid-Containing Peptides and Proteins. *Chem* **2022**, *8* (10), 2659–2677. https://doi.org/10.1016/j.chempr.2022.09.017.

(810) Lakis, E.; Magyari, S.; Piel, J. In Vivo Production of Diverse B-Amino Acid-Containing Proteins. *Angewandte Chemie International Edition* **2022**, *61* (29), e202202695. https://doi.org/10.1002/anie.202202695.

(811) McGeady, P.; Pyun, H.-J.; Coates, R. M.; Croteau, R. Biosynthesis of Monoterpenes: Inhibition of (+)-Pinene and (-)-Pinene Cyclases by Thia and Aza Analogs of the 4R- and 4S-α-Terpinyl Carbocation. *Arch Biochem Biophys* **1992**, *299* (1), 63–72. https://doi.org/10.1016/0003-9861(92)90244-Q.

(812) Rajaonarivony, J. I. M.; Gershenzon, J.; Croteau, R. Characterization and Mechanism of (4S)-Limonene Synthase, A Monoterpene Cyclase from the Glandular Trichomes of Peppermint (Mentha x Piperita). *Arch Biochem Biophys* **1992**, *296* (1), 49–57. https://doi.org/10.1016/0003-9861(92)90543-6.

(813) Oberhauser, C.; Harms, V.; Seidel, K.; Schröder, B.; Ekramzadeh, K.; Beutel, S.; Winkler, S.; Lauterbach, L.; Dickschat, J. S.; Kirschning, A. Exploiting the Synthetic Potential of Sesquiterpene Cyclases for Generating Unnatural Terpenoids. *Angewandte Chemie International Edition* **2018**, *57* (36), 11802–11806. https://doi.org/10.1002/anie.201805526.

(814) Quan, Z.; Awakawa, T.; Wang, D.; Hu, Y.; Abe, I. Multidomain P450 Epoxidase and a Terpene Cyclase from the Ascochlorin Biosynthetic Pathway in *Fusarium* Sp. *Org Lett* **2019**, *21* (7), 2330–2334. https://doi.org/10.1021/acs.orglett.9b00616.

(815) Gu, B.; Dickschat, J. S. Ruptenes: A Family of Terpene Analogs Give Insight into Cyclisation Mechanisms by Cascade Disruption. *Angewandte Chemie International Edition* **2023**, *62* (32), e202307006. https://doi.org/10.1002/anie.202307006.

(816) Tran, C. D.; Dräger, G.; Struwe, H. F.; Siedenberg, L.; Vasisth, S.; Grunenberg, J.; Kirschning, A. Cyclopropylmethyldiphosphates Are Substrates for Sesquiterpene Synthases: Experimental and Theoretical Results. *Org Biomol Chem* **2022**, *20* (39), 7833–7839. https://doi.org/10.1039/D2OB01279K.

(817) Harms, V.; Schröder, B.; Oberhauser, C.; Tran, C. D.; Winkler, S.; Dräger, G.; Kirschning, A. Methyl-Shifted Farnesyldiphosphate Derivatives Are Substrates for Sesquiterpene Cyclases. *Org Lett* **2020**, *22* (11), 4360–4365. https://doi.org/10.1021/acs.orglett.0c01345.

(818) Struwe, H.; Schrödter, F.; Spinck, H.; Kirschning, A. Sesquiterpene Backbones Generated by Sesquiterpene Cyclases: Formation of *Iso* -Caryolan-1-Ol and an Isoclovane. *Org Lett* **2023**, *25* (48), 8575–8579. https://doi.org/10.1021/acs.orglett.3c03383.

(819) Gu, B.; Goldfuss, B.; Schnakenburg, G.; Dickschat, J. S. Subrutilane—A Hexacyclic Sesterterpene from *Streptomyces Subrutilus*. *Angewandte Chemie International Edition* **2023**, *62* (48), e202313789. https://doi.org/10.1002/anie.202313789.

(820) Struwe, H.; Li, H.; Schrödter, F.; Höft, L.; Fohrer, J.; Dickschat, J. S.; Kirschning, A. Telescoping a Prenyltransferase and a Diterpene Synthase to Transform Unnatural FPP Derivatives to Diterpenoids. *Org Lett* **2024**, *26* (28), 5888–5892. https://doi.org/10.1021/acs.orglett.4c01670.

(821) Drummond, L.; Kschowak, M. J.; Breitenbach, J.; Wolff, H.; Shi, Y.-M.; Schrader, J.; Bode, H. B.; Sandmann, G.; Buchhaupt, M. Expanding the Isoprenoid Building Block Repertoire with an IPP Methyltransferase from *Streptomyces Monomycini*. *ACS Synth Biol* **2019**, *8* (6), 1303–1313. https://doi.org/10.1021/acssynbio.8b00525.

(822) Schooley, D. A.; Judy, K. J.; Bergot, B. J.; Hall, M. S.; Siddall, J. B. Biosynthesis of the Juvenile Hormones of *Manduca Sexta:* Labeling Pattern from Mevalonate, Propionate, and Acetate. *Proceedings of the National Academy of Sciences* **1973**, *70* (10), 2921–2925. https://doi.org/10.1073/pnas.70.10.2921.

(823) Eiben, C. B.; de Rond, T.; Bloszies, C.; Gin, J.; Chiniquy, J.; Baidoo, E. E. K.; Petzold, C. J.; Hillson, N. J.; Fiehn, O.; Keasling, J. D. Mevalonate Pathway Promiscuity Enables Noncanonical Terpene Production. *ACS Synth Biol* **2019**, *8* (10), 2238–2247. https://doi.org/10.1021/acssynbio.9b00230.

(824) Ignea, C.; Raadam, M. H.; Koutsaviti, A.; Zhao, Y.; Duan, Y.-T.; Harizani, M.; Miettinen, K.; Georgantea, P.; Rosenfeldt, M.; Viejo-Ledesma, S. E.; Petersen, M. A.; Bredie, W. L. P.; Staerk, D.; Roussis, V.; Ioannou, E.; Kampranis, S. C. Expanding the Terpene Biosynthetic Code with Non-Canonical 16 Carbon Atom Building Blocks. *Nat Commun* **2022**, *13* (1), 5188. https://doi.org/10.1038/s41467-022-32921-w.

(825) Reuter, T.; Dieminger, L.; Steidle, S.; Zoller, K.; Holocher, M.; Zhou, L.; Hanauska, D.; Racz, K.; Barra, L. Non-canonical C16 Homoterpene Biosynthesis Widespread in Actinobacteria. *Angewandte Chemie International Edition* **2024**, e202418613. https://doi.org/10.1002/anie.202418613.

(826) Saalim, M.; Liu, S.; Bennett, S. D.; Zaleta-Pinet, D. A.; Poulin, R. X.; Clark, B. R. Precursor-Directed Biosynthesis of Antialgal Fluorinated Bacillamide Derivatives in *Bacillus Atrophaeus*. *J Nat Prod* **2024**, *87* (2), 388–395. https://doi.org/10.1021/acs.jnatprod.3c01178.

(827) Roy, A. D.; Grüschow, S.; Cairns, N.; Goss, R. J. M. Gene Expression Enabling Synthetic Diversification of Natural Products: Chemogenetic Generation of Pacidamycin Analogs. *J Am Chem Soc* **2010**, *132* (35), 12243–12245. https://doi.org/10.1021/ja1060406.

(828) Sharma, S. V.; Tong, X.; Pubill-Ulldemolins, C.; Cartmell, C.; Bogosyan, E. J. A.; Rackham, E. J.; Marelli, E.; Hamed, R. B.; Goss, R. J. M. Living GenoChemetics by Hyphenating Synthetic Biology and Synthetic Chemistry in Vivo. *Nat Commun* **2017**, *8* (1), 229. https://doi.org/10.1038/s41467-017-00194-3.

(829) Runguphan, W.; O'Connor, S. E. Diversification of Monoterpene Indole Alkaloid Analogs through Cross-Coupling. *Org Lett* **2013**, *15* (11), 2850–2853. https://doi.org/10.1021/ol401179k.

(830) Lai, H.-E.; Obled, A. M. C.; Chee, S. M.; Morgan, R. M.; Lynch, R.; Sharma, S. V.; Moore, S. J.; Polizzi, K. M.; Goss, R. J. M.; Freemont, P. S. GenoChemetic Strategy for Derivatization of the Violacein Natural Product Scaffold. *ACS Chem Biol* **2021**, *16* (11), 2116–2123. https://doi.org/10.1021/acschembio.1c00483.

(831) Boccia, M.; Grzech, D.; Lopes, A. A.; O'Connor, S. E.; Caputi, L. Directed Biosynthesis of New to Nature Alkaloids in a Heterologous Nicotiana Benthamiana Expression Host. *Front Plant Sci* **2022**, *13*, 919443. https://doi.org/10.3389/fpls.2022.919443.

(832) Dror, M. J.; Misa, J.; Yee, D. A.; Chu, A. M.; Yu, R. K.; Chan, B. B.; Aoyama, L. S.; Chaparala, A. P.; O'Connor, S. E.; Tang, Y. Engineered Biosynthesis of Plant Heteroyohimbine and Corynantheine Alkaloids in *Saccharomyces Cerevisiae*. *J Ind Microbiol Biotechnol* **2024**, *51*, kuad047. https://doi.org/10.1093/jimb/kuad047.

(833) Zehner, S.; Kotzsch, A.; Bister, B.; Süssmuth, R. D.; Méndez, C.; Salas, J. A.; van Pée, K.-H. A Regioselective Tryptophan 5-Halogenase Is Involved in Pyrroindomycin Biosynthesis in Streptomyces Rugosporus LL-42D005. *Chem Biol* **2005**, *12* (4), 445–452. https://doi.org/10.1016/j.chembiol.2005.02.005.

(834) Domergue, J.; Erdmann, D.; Fossey-Jouenne, A.; Petit, J.-L.; Debard, A.; de Berardinis, V.; Vergne-Vaxelaire, C.; Zaparucha, A. XszenFHal, a Novel Tryptophan 5-Halogenase from Xenorhabdus Szentirmaii. *AMB Express* **2019**, *9* (1), 175. https://doi.org/10.1186/s13568-019-0898-y.

(835) Lee, J.; Kim, J.; Kim, H.; Kim, E.; Jeong, H.; Choi, K.; Kim, B. Characterization of a Tryptophan 6-Halogenase from *Streptomyces Albus* and Its Regioselectivity Determinants. *ChemBioChem* **2020**, *21* (10), 1446–1452. https://doi.org/10.1002/cbic.201900723.

(836) Seibold, C.; Schnerr, H.; Rumpf, J.; Kunzendorf, A.; Hatscher, C.; Wage, T.; Ernyei, A. J.; Dong, C.; Naismith, J. H.; Van Pée, K.-H. A Flavin-Dependent Tryptophan 6-Halogenase and Its Use in Modification of Pyrrolnitrin Biosynthesis. *Biocatal Biotransformation* **2006**, *24* (6), 401–408. https://doi.org/10.1080/10242420601033738.

(837) Keller, S.; Wage, T.; Hohaus, K.; Hölzer, M.; Eichhorn, E.; van Pée, K.-H. Purification and Partial Characterization of Tryptophan 7-Halogenase (PrnA) FromPseudomonas Fluorescens. *Angewandte Chemie International Edition* **2000**, *39* (13), 2300–2302. https://doi.org/10.1002/1521-3773(20000703)39:13<2300::AID-ANIE2300>3.0.CO;2-I.

(838) Yeh, E.; Garneau, S.; Walsh, C. T. Robust *in Vitro* Activity of RebF and RebH, a Two-Component Reductase/Halogenase, Generating 7-Chlorotryptophan during Rebeccamycin Biosynthesis. *Proceedings of the National Academy of Sciences* **2005**, *102* (11), 3960–3965. https://doi.org/10.1073/pnas.0500755102.

(839) Latham, J.; Henry, J.-M.; Sharif, H. H.; Menon, B. R. K.; Shepherd, S. A.; Greaney, M. F.; Micklefield, J. Integrated Catalysis Opens New Arylation Pathways via Regiodivergent Enzymatic C–H Activation. *Nat Commun* **2016**, *7* (1), 11873. https://doi.org/10.1038/ncomms11873.

(840) Runguphan, W.; Qu, X.; O'Connor, S. E. Integrating Carbon–Halogen Bond Formation into Medicinal Plant Metabolism. *Nature* **2010**, *468* (7322), 461–464. https://doi.org/10.1038/nature09524.

(841) Sánchez, C.; Zhu, L.; Braña, A. F.; Salas, A. P.; Rohr, J.; Méndez, C.; Salas, J. A. Combinatorial Biosynthesis of Antitumor Indolocarbazole Compounds. *Proceedings of the National Academy of Sciences* **2005**, *102* (2), 461–466. https://doi.org/10.1073/pnas.0407809102.

(842) Banerjee, T.; DuHadaway, J. B.; Gaspari, P.; Sutanto-Ward, E.; Munn, D. H.; Mellor, A. L.; Malachowski, W. P.; Prendergast, G. C.; Muller, A. J. A Key in Vivo Antitumor Mechanism of Action of Natural Product-Based Brassinins Is Inhibition of Indoleamine 2,3-Dioxygenase. *Oncogene* **2008**, *27* (20), 2851–2857. https://doi.org/10.1038/sj.onc.1210939.

(843) Calgaro-Kozina, A.; Vuu, K. M.; Keasling, J. D.; Loqué, D.; Sattely, E. S.; Shih, P. M. Engineering Plant Synthetic Pathways for the Biosynthesis of Novel Antifungals. *ACS Cent Sci* **2020**, *6* (8), 1394–1400. https://doi.org/10.1021/acscentsci.0c00241.

(844) Reed, K. B.; Brooks, S. M.; Wells, J.; Blake, K. J.; Zhao, M.; Placido, K.; d'Oelsnitz, S.; Trivedi, A.; Gadhiyar, S.; Alper, H. S. A Modular and Synthetic Biosynthesis Platform for de Novo Production of Diverse Halogenated Tryptophan-Derived Molecules. *Nat Commun* **2024**, *15* (1), 3188. https://doi.org/10.1038/s41467-024-47387-1.

(845) Corr, M. J.; Sharma, S. V.; Pubill-Ulldemolins, C.; Bown, R. T.; Poirot, P.; Smith, D. R. M.; Cartmell, C.; Abou Fayad, A.; Goss, R. J. M. Sonogashira Diversification of Unprotected Halotryptophans, Halotryptophan Containing Tripeptides; and Generation of a New to Nature Bromo-Natural Product and Its Diversification in Water. *Chem Sci* **2017**, *8* (3), 2039–2046. https://doi.org/10.1039/C6SC04423A.

(846) Renault, Y. J. G.; Lynch, R.; Marelli, E.; Sharma, S. V.; Pubill-Ulldemolins, C.; Sharp, J. A.; Cartmell, C.; Cárdenas, P.; Goss, R. J. M. Buchwald Hartwig Diversification of Unprotected Halotryptophans, Halotryptophan Containing Tripeptides and the Natural Product Barettin in Aqueous Conditions. *Chemical Communications* **2019**, *55* (91), 13653–13656. https://doi.org/10.1039/C9CC02554E.

(847) Pubill-Ulldemolins, C.; Sharma, S. V.; Cartmell, C.; Zhao, J.; Cárdenas, P.; Goss, R. J. M. Heck Diversification of Indole-Based Substrates under Aqueous Conditions: From Indoles to Unprotected Halo-tryptophans and Halo-tryptophans in Natural Product Derivatives. *Chemistry – A European Journal* **2019**, *25* (46), 10866–10875. https://doi.org/10.1002/chem.201901327.

(848) Cantoni, G. L. THE NATURE OF THE ACTIVE METHYL DONOR FORMED ENZYMATICALLY FROM L-METHIONINE AND ADENOSINETRIPHOSPHATE [1,2]. *J Am Chem Soc* **1952**, *74* (11), 2942–2943. https://doi.org/10.1021/ja01131a519.

(849) Zhang, C.; Weller, R. L.; Thorson, J. S.; Rajski, S. R. Natural Product Diversification Using a Non-Natural Cofactor Analogue of *S* -Adenosyl- <scp>l</Scp> -Methionine. *J Am Chem Soc* **2006**, *128* (9), 2760–2761. https://doi.org/10.1021/ja056231t.

(850) Peng, J.; Liao, C.; Bauer, C.; Seebeck, F. P. Fluorinated *S* -Adenosylmethionine as a Reagent for Enzyme-Catalyzed Fluoromethylation. *Angewandte Chemie International Edition* **2021**, *60* (52), 27178–27183. https://doi.org/10.1002/anie.202108802.

(851) Wang, W.; Zhao, H.; Yu, N.; Chen, F.; Dong, M. Stable *S* -Adenosylmethionine Analogue for Enzymatic Fluoromethylation. *ACS Catal* **2023**, *13* (20), 13729–13734. https://doi.org/10.1021/acscatal.3c03313.

(852) Neti, S. S.; Wang, B.; Iwig, D. F.; Onderko, E. L.; Booker, S. J. Enzymatic Fluoromethylation Enabled by the *S* -Adenosylmethionine Analog *Te* -Adenosyl- *L* -(Fluoromethyl)Homotellurocysteine. *ACS Cent Sci* **2023**, *9* (5), 905–914. https://doi.org/10.1021/acscentsci.2c01385.

(853) Yu, N.; Zhao, H.; Wang, W.; Dong, M. Enzymatic Fluoroethylation by a Fluoroethyl Selenium Analogue of S-Adenosylmethionine. *ACS Catal* **2024**, *14* (8), 6211–6216. https://doi.org/10.1021/acscatal.4c01112.

(854) Jockmann, E.; Subrizi, F.; Mohr, M. K. F.; Carter, E. M.; Hebecker, P. M.; Popadić, D.; Hailes, H. C.; Andexer, J. N. Expanding the Substrate Scope of *N* - and *O* -Methyltransferases from Plants for Chemoselective Alkylation**. *ChemCatChem* **2023**, *15* (22), e202300930. https://doi.org/10.1002/cctc.202300930.

(855) Stecher, H.; Tengg, M.; Ueberbacher, B. J.; Remler, P.; Schwab, H.; Griengl, H.; Gruber-Khadjawi, M. Biocatalytic Friedel–Crafts Alkylation Using Non-natural Cofactors. *Angewandte Chemie International Edition* **2009**, *48* (50), 9546–9548. https://doi.org/10.1002/anie.200905095.

(856) McKean, I. J. W.; Sadler, J. C.; Cuetos, A.; Frese, A.; Humphreys, L. D.; Grogan, G.; Hoskisson, P. A.; Burley, G. A. *S* -Adenosyl Methionine Cofactor Modifications Enhance the Biocatalytic Repertoire of Small Molecule *C* -Alkylation. *Angewandte Chemie International Edition* **2019**, *58* (49), 17583–17588. https://doi.org/10.1002/anie.201908681.

(857) Ding, W.; Zhou, M.; Li, H.; Li, M.; Qiu, Y.; Yin, Y.; Pan, L.; Yang, W.; Du, Y.; Zhang, X.; Tang, Z.; Liu, W. Biocatalytic Fluoroalkylation Using Fluorinated *S* -Adenosyl- <scp>l</Scp> -Methionine Cofactors. *Org Lett* **2023**, *25* (30), 5650–5655. https://doi.org/10.1021/acs.orglett.3c02028.

(858) Law, B. J. C.; Struck, A.-W.; Bennett, M. R.; Wilkinson, B.; Micklefield, J. Site-Specific Bioalkylation of Rapamycin by the RapM 16-O-Methyltransferase. *Chem Sci* **2015**, *6* (5), 2885–2892. https://doi.org/10.1039/C5SC00164A.

(859) Herbert, A. J.; Shepherd, S. A.; Cronin, V. A.; Bennett, M. R.; Sung, R.; Micklefield, J. Engineering Orthogonal Methyltransferases to Create Alternative Bioalkylation Pathways. *Angewandte Chemie International Edition* **2020**, *59* (35), 14950–14956. https://doi.org/10.1002/anie.202004963.

(860) Salcedo, R. G.; Olano, C.; Fernández, R.; Braña, A. F.; Méndez, C.; de la Calle, F.; Salas, J. A. Elucidation of the Glycosylation Steps during Biosynthesis of Antitumor Macrolides PM100117 and PM100118 and Engineering for Novel Derivatives. *Microb Cell Fact* **2016**, *15* (1), 187. https://doi.org/10.1186/s12934-016-0591-7.

(861) OMURA, S.; SASAKI, Y.; IWAI, Y.; TAKESHIMA, H. Staurosporine, a Potentially Important Gift from a Microorganism. *J Antibiot (Tokyo)* **1995**, *48* (7), 535–548. https://doi.org/10.7164/antibiotics.48.535.

(862) Di Marco, A.; Gaetani, M.; Scarpinato, B. Adriamycin (NSC-123,127): A New Antibiotic with Antitumor Activity. *Cancer chemotherapy reports. Part 1* **1969**, *53* (1), 33–37.

(863) MCCORMICK, M. H.; MCGUIRE, J. M.; PITTENGER, G. E.; PITTENGER, R. C.; STARK, W. M. Vancomycin, a New Antibiotic. I. Chemical and Biologic Properties. *Antibiot Annu* **1955**, *3*, 606–611.

(864) Wolf, H. Avilamycin, an Inhibitor of the 30 S Ribosomal Subunits Function. *FEBS Lett* **1973**, *36* (2), 181–186. https://doi.org/10.1016/0014-5793(73)80364-3.

(865) Singh, M. P.; Petersen, P. J.; Weiss, W. J.; Kong, F.; Greenstein, M. Saccharomicins, Novel Heptadecaglycoside Antibiotics Produced by *Saccharothrix Espanaensis*: Antibacterial and Mechanistic Activities. *Antimicrob Agents Chemother* **2000**, *44* (8), 2154–2159. https://doi.org/10.1128/AAC.44.8.2154-2159.2000.

(866) Thibodeaux, C. J.; Melançon, C. E.; Liu, H. Natural-Product Sugar Biosynthesis and Enzymatic Glycodiversification. *Angewandte Chemie International Edition* **2008**, *47* (51), 9814–9859. https://doi.org/10.1002/anie.200801204.

(867) Han, A. R.; Park, J. W.; Lee, M. K.; Ban, Y. H.; Yoo, Y. J.; Kim, E. J.; Kim, E.; Kim, B.-G.; Sohng, J. K.; Yoon, Y. J. Development of a Streptomyces Venezuelae-Based Combinatorial Biosynthetic System for the Production of Glycosylated Derivatives of Doxorubicin and Its Biosynthetic Intermediates. *Appl Environ Microbiol* **2011**, *77* (14), 4912–4923. https://doi.org/10.1128/AEM.02527-10.

(868) Shepherd, M. D.; Liu, T.; Méndez, C.; Salas, J. A.; Rohr, J. Engineered Biosynthesis of Gilvocarcin Analogues with Altered Deoxyhexopyranose Moieties. *Appl Environ Microbiol* **2011**, *77* (2), 435–441. https://doi.org/10.1128/AEM.01774-10.

(869) Shinde, P. B.; Han, A. R.; Cho, J.; Lee, S. R.; Ban, Y. H.; Yoo, Y. J.; Kim, E. J.; Kim, E.; Song, M.-C.; Park, J. W.; Lee, D. G.; Yoon, Y. J. Combinatorial Biosynthesis and Antibacterial Evaluation of Glycosylated Derivatives of 12-Membered Macrolide Antibiotic YC-17. *J Biotechnol* **2013**, *168* (2), 142–148. https://doi.org/10.1016/j.jbiotec.2013.05.014.

(870) Blanco, G.; Patallo, E. P.; Braña, A. F.; Trefzer, A.; Bechthold, A.; Rohr, J.; Méndez, C.; Salas, J. A. Identification of a Sugar Flexible Glycosyltransferase from Streptomyces Olivaceus, the Producer of the Antitumor Polyketide Elloramycin. *Chem Biol* **2001**, *8* (3), 253–263. https://doi.org/10.1016/S1074-5521(01)00010-2.

(871) Rodríguez, L.; Aguirrezabalaga, I.; Allende, N.; Braña, A. F.; Méndez, C.; Salas, J. A. Engineering Deoxysugar Biosynthetic Pathways from Antibiotic-Producing Microorganisms. *Chem Biol* **2002**, *9* (6), 721–729. https://doi.org/10.1016/S1074-5521(02)00154-0.

(872) Pérez, M.; Lombó, F.; Zhu, L.; Gibson, M.; Braña, A. F.; Rohr, J.; Salas, J. A.; Méndez, C. Combining Sugar Biosynthesis Genes for the Generation of <scp>l</Scp> - and <scp>d</Scp> -Amicetose and Formation of Two Novel Antitumor Tetracenomycins. *Chem. Commun.* **2005**, No. 12, 1604–1606. https://doi.org/10.1039/B417815G.

(873) Pérez, M.; Lombó, F.; Baig, I.; Braña, A. F.; Rohr, J.; Salas, J. A.; Méndez, C. Combinatorial Biosynthesis of Antitumor Deoxysugar Pathways in *Streptomyces Griseus*: Reconstitution of "Unnatural Natural Gene Clusters" for the Biosynthesis of Four 2,6- <scp>d</Scp> -Dideoxyhexoses. *Appl Environ Microbiol* **2006**, *72* (10), 6644–6652. https://doi.org/10.1128/AEM.01266-06.

(874) Eric Nybo, S.; Shabaan, K. A.; Kharel, M. K.; Sutardjo, H.; Salas, J. A.; Méndez, C.; Rohr, J. Ketoolivosyl-Tetracenomycin C: A New Ketosugar Bearing Tetracenomycin Reveals New Insight into the Substrate Flexibility of Glycosyltransferase ElmGT. *Bioorg Med Chem Lett* **2012**, *22* (6), 2247–2250. https://doi.org/10.1016/j.bmcl.2012.01.094.

(875) Tirkkonen, H.; Brown, K. V.; Niemczura, M.; Faudemer, Z.; Brown, C.; Ponomareva, L. V.; Helmy, Y. A.; Thorson, J. S.; Nybo, S. E.; Metsä-Ketelä, M.; Shaaban, K. A. Engineering BioBricks for Deoxysugar Biosynthesis and Generation of New Tetracenomycins. *ACS Omega* **2023**, *8* (23), 21237–21253. https://doi.org/10.1021/acsomega.3c02460.

(876) Langenhan, J. M.; Griffith, B. R.; Thorson, J. S. Neoglycorandomization and Chemoenzymatic Glycorandomization: Two Complementary Tools for Natural Product Diversification. *J Nat Prod* **2005**, *68* (11), 1696–1711. https://doi.org/10.1021/np0502084.

(877) Albermann, C.; Soriano, A.; Jiang, J.; Vollmer, H.; Biggins, J. B.; Barton, W. A.; Lesniak, J.; Nikolov, D. B.; Thorson, J. S. Substrate Specificity of NovM: Implications for Novobiocin Biosynthesis and Glycorandomization. *Org Lett* **2003**, *5* (6), 933–936. https://doi.org/10.1021/ol0341086.

(878) Zhang, C.; Griffith, B. R.; Fu, Q.; Albermann, C.; Fu, X.; Lee, I.-K.; Li, L.; Thorson, J. S. Exploiting the Reversibility of Natural Product Glycosyltransferase-Catalyzed Reactions. *Science (1979)* **2006**, *313* (5791), 1291–1294. https://doi.org/10.1126/science.1130028.

(879) Williams, G. J.; Yang, J.; Zhang, C.; Thorson, J. S. Recombinant *E. Coli* Prototype Strains for *in Vivo* Glycorandomization. *ACS Chem Biol* **2011**, *6* (1), 95–100. https://doi.org/10.1021/cb100267k.

(880) Zhang, C.; Albermann, C.; Fu, X.; Thorson, J. S. The in Vitro Characterization of the Iterative Avermectin Glycosyltransferase AveBI Reveals Reaction Reversibility and Sugar Nucleotide Flexibility. *J Am Chem Soc* **2006**, *128* (51), 16420–16421. https://doi.org/10.1021/ja065950k.

(881) Yamase, H.; Zhao, L.; Liu, H. Engineering a Hybrid Sugar Biosynthetic Pathway: Production of <scp>l</Scp> -Rhamnose and Its Implication on Dihydrostreptose Biosynthesis. *J Am Chem Soc* **2000**, *122* (49), 12397–12398. https://doi.org/10.1021/ja003246c.

(882) Zhao, L.; Sherman, D. H.; Liu, H. Biosynthesis of Desosamine: Construction of a New Methymycin/Neomethymycin Analogue by Deletion of a Desosamine Biosynthetic Gene. *J Am Chem Soc* **1998**, *120* (39), 10256–10257. https://doi.org/10.1021/ja981782n.

(883) Fu, X.; Albermann, C.; Jiang, J.; Liao, J.; Zhang, C.; Thorson, J. S. Antibiotic Optimization via in Vitro Glycorandomization. *Nat Biotechnol* **2003**, *21* (12), 1467–1469. https://doi.org/10.1038/nbt909.

(884) Hoffmeister, D.; Yang, J.; Liu, L.; Thorson, J. S. Creation of the First Anomeric <scp>d</Scp> / <scp>l</Scp> -Sugar Kinase by Means of Directed Evolution. *Proceedings of the National Academy of Sciences* **2003**, *100* (23), 13184–13189. https://doi.org/10.1073/pnas.100.23.13184.

(885) Yang, J.; Fu, X.; Liao, J.; Liu, L.; Thorson, J. S. Structure-Based Engineering of E. Coli Galactokinase as a First Step toward In Vivo Glycorandomization. *Chem Biol* **2005**, *12* (6), 657–664. https://doi.org/10.1016/j.chembiol.2005.04.009.

(886) Fu, X.; Albermann, C.; Zhang, C.; Thorson, J. S. Diversifying Vancomycin via Chemoenzymatic Strategies. *Org Lett* **2005**, *7* (8), 1513–1515. https://doi.org/10.1021/ol0501626.

(887) Malmierca, M. G.; Pérez-Victoria, I.; Martín, J.; Reyes, F.; Méndez, C.; Salas, J. A.; Olano, C. New Sipanmycin Analogues Generated by Combinatorial Biosynthesis and Mutasynthesis Approaches Relying on the Substrate Flexibility of Key Enzymes in the Biosynthetic Pathway. *Appl Environ Microbiol* **2020**, *86* (3), e02453-19. https://doi.org/10.1128/AEM.02453-19.

(888) Zhang, H.; Gong, N.; Zhang, H.; Li, Q.; Ma, J.; Wei, X.; Li, W.; Ju, J. Characterization of the Glycosyltransferase and Methyltransferase Encoded Remotely from the Actinopyrone Biosynthetic Gene Cluster Enables Access to Diverse Analogues. *Org Lett* **2022**, *24* (49), 9065–9070. https://doi.org/10.1021/acs.orglett.2c03707.

(889) Yan, P.; Wang, G.; Huang, M.; Liu, Z.; Dai, C.; Hu, B.; Gu, M.; Deng, Z.; Liu, R.; Wang, X.; Liu, T. Combinatorial Biosynthesis Creates a Novel Aglycone Polyether with High Potency and Low Side Effects Against Bladder Cancer. *Advanced Science* **2024**, *11* (32), 2404668. https://doi.org/10.1002/advs.202404668.

(890) Huang, M.; Liu, B.; Liu, R.; Li, J.; Chen, J.; Jiang, F.; Ding, H.; Deng, Z.; Liu, T. Aglycone Polyether Nanchangmycin and Its Homologues Exhibit Apoptotic and Antiproliferative Activities against Cancer Stem Cells. *ACS Pharmacol Transl Sci* **2018**, *1* (2), 84–95. https://doi.org/10.1021/acsptsci.8b00007.

(891) Jian, X.; Wang, C.; Wu, S.; Sun, G.; Huang, C.; Qiu, C.; Liu, Y.; Leadlay, P. F.; Liu, D.; Deng, Z.; Zhou, F.; Sun, Y. Glycodiversification of Gentamicins through in Vivo Glycosyltransferase Swapping Enabled the Creation of Novel Hybrid Aminoglycoside Antibiotics with Potent Activity and Low Ototoxicity. *Acta Pharm Sin B* **2024**, *14* (9), 4149–4163. https://doi.org/10.1016/j.apsb.2024.04.030.

(892) Jiménez-Osés, G.; Osuna, S.; Gao, X.; Sawaya, M. R.; Gilson, L.; Collier, S. J.; Huisman, G. W.; Yeates, T. O.; Tang, Y.; Houk, K. N. The Role of Distant Mutations and Allosteric Regulation on LovD Active Site Dynamics. *Nat Chem Biol* **2014**, *10* (6), 431–436. https://doi.org/10.1038/nchembio.1503.

(893) Wu, Z.; Xia, Z.; Tang, Z.; Li, J.; Liu, W. Mutasynthesis Generates Nine New Pyrroindomycins. *Org Biomol Chem* **2024**, *22* (14), 2813–2818. https://doi.org/10.1039/D4OB00239C.

(894) Seidel, K.; Balakrishnan, A.; Alexiou, C.; Janko, C.; Komoll, R.; Wang, L.; Kirschning, A.; Ott, M. Synthesis of Magnetic-Nanoparticle/Ansamitocin Conjugates—Inductive Heating Leads to Decreased Cell Proliferation In Vitro and Attenuation Of Tumour Growth In Vivo. *Chemistry – A European Journal* **2017**, *23* (50), 12326–12337. https://doi.org/10.1002/chem.201701491.

(895) Pollier, J.; Moses, T.; Goossens, A. Combinatorial Biosynthesis in Plants: A (p)Review on Its Potential and Future Exploitation. *Nat Prod Rep* **2011**, *28* (12), 1897. https://doi.org/10.1039/c1np00049g.

(896) Weissman, K. J.; Leadlay, P. F. Combinatorial Biosynthesis of Reduced Polyketides. *Nature Reviews Microbiology*. 2005, pp 925–936. https://doi.org/10.1038/nrmicro1287.

(897) Hopwood, D. A.; Malpartida, F.; Kieser, H. M.; Ikeda, H.; Duncan, J.; Fujii, I.; Rudd, B. A. M.; Floss, H. G.; Ōmura, S. Production of 'Hybrid' Antibiotics by Genetic Engineering. *Nature* **1985**, *314* (6012), 642–644. https://doi.org/10.1038/314642a0.

(898) Liu, Y.; Wang, H.; Song, R.; Chen, J.; Li, T.; Li, Y.; Du, L.; Shen, Y. Targeted Discovery and Combinatorial Biosynthesis of Polycyclic Tetramate Macrolactam Combamides A–E. *Org Lett* **2018**, *20* (12), 3504–3508. https://doi.org/10.1021/acs.orglett.8b01285.

(899) Yan, Y.; Wang, H.; Song, Y.; Zhu, D.; Shen, Y.; Li, Y. Combinatorial Biosynthesis of Oxidized Combamides Using Cytochrome P450 Enzymes from Different Polycyclic Tetramate Macrolactam Pathways. *ACS Synth Biol* **2021**, *10* (10), 2434–2439. https://doi.org/10.1021/acssynbio.1c00178.

(900) Guo, Z.; Li, P.; Chen, G.; Li, C.; Cao, Z.; Zhang, Y.; Ren, J.; Xiang, H.; Lin, S.; Ju, J.; Chen, Y. Design and Biosynthesis of Dimeric Alboflavusins with Biaryl Linkages via Regiospecific C–C Bond Coupling. *J Am Chem Soc* **2018**, *140* (51), 18009–18015. https://doi.org/10.1021/jacs.8b10136.

(901) Schotte, C.; Li, L.; Wibberg, D.; Kalinowski, J.; Cox, R. J. Synthetic Biology Driven Biosynthesis of Unnatural Tropolone Sesquiterpenoids. *Angewandte Chemie International Edition* **2020**, *59* (52), 23870–23878. https://doi.org/10.1002/anie.202009914.

(902) Rao, L.; Yuan, G.-Y.; Chen, X.-Y.; Ran, J.-L.; Zou, Y. Reshaping the Diversity of Oxidized Polyquinane Sesquiterpenoids by Cytochrome P450s. *Org Lett* **2023**, *25* (18), 3276–3280. https://doi.org/10.1021/acs.orglett.3c01024.

(903) Mafu, S.; Jia, M.; Zi, J.; Morrone, D.; Wu, Y.; Xu, M.; Hillwig, M. L.; Peters, R. J. Probing the Promiscuity of *Ent* -Kaurene Oxidases via Combinatorial Biosynthesis. *Proceedings of the National Academy of Sciences* **2016**, *113* (9), 2526–2531. https://doi.org/10.1073/pnas.1512096113.

(904) Wei, Q.; Zeng, H.-C.; Zou, Y. Divergent Biosynthesis of Fungal Dioxafenestrane Sesquiterpenes by the Cooperation of Distinctive Baeyer–Villiger Monooxygenases and α-Ketoglutarate-Dependent Dioxygenases. *ACS Catal* **2021**, *11* (2), 948–957. https://doi.org/10.1021/acscatal.0c05319.

(905) Lin, G.-M.; Voigt, C. A. Design of a Redox-Proficient Escherichia Coli for Screening Terpenoids and Modifying Cytochrome P450s. *Nat Catal* **2023**, *6* (11), 1016–1029. https://doi.org/10.1038/s41929-023-01049-5.

(906) Tao, H.; Mori, T.; Chen, H.; Lyu, S.; Nonoyama, A.; Lee, S.; Abe, I. Molecular Insights into the Unusually Promiscuous and Catalytically Versatile Fe(II)/α-Ketoglutarate-Dependent Oxygenase SptF. *Nat Commun* **2022**, *13* (1), 95. https://doi.org/10.1038/s41467-021-27636-3.

(907) Frey, M.; Bathe, U.; Meink, L.; Balcke, G. U.; Schmidt, J.; Frolov, A.; Soboleva, A.; Hassanin, A.; Davari, M. D.; Frank, O.; Schlagbauer, V.; Dawid, C.; Tissier, A. Combinatorial Biosynthesis in Yeast Leads to over 200 Diterpenoids. *Metab Eng* **2024**, *82*, 193–200. https://doi.org/10.1016/j.ymben.2024.02.006.

(908) Tang, M.-C.; Shen, C.; Deng, Z.; Ohashi, M.; Tang, Y. Combinatorial Biosynthesis of Terpenoids through Mixing-and-Matching Sesquiterpene Cyclase and Cytochrome P450 Pairs. *Org Lett* **2022**, *24* (26), 4783–4787. https://doi.org/10.1021/acs.orglett.2c01785.

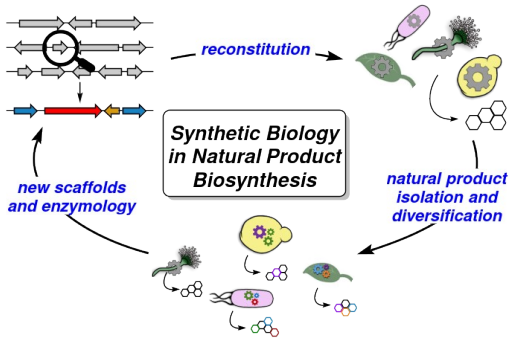
(909) Zhao, Y.; Duan, Y.; Zang, J.; Raadam, M. H.; Pateraki, I.; Miettinen, K.; Staerk, D.; Kampranis, S. C. Structure-Agnostic Bioactivity-Driven Combinatorial Biosynthesis Reveals New Antidiabetic and Anticancer Triterpenoids. *Angewandte Chemie International Edition* **2024**, e202416218. https://doi.org/10.1002/anie.202416218.

(910) Tsukada, K.; Shinki, S.; Kaneko, A.; Murakami, K.; Irie, K.; Murai, M.; Miyoshi, H.; Dan, S.; Kawaji, K.; Hayashi, H.; Kodama, E. N.; Hori, A.; Salim, E.; Kuraishi, T.; Hirata, N.; Kanda, Y.; Asai, T. Synthetic Biology Based Construction of Biological Activity-Related Library of Fungal Decalin-Containing Diterpenoid Pyrones. *Nat Commun* **2020**, *11* (1), 1830. https://doi.org/10.1038/s41467-020-15664-4.

(911) Sun, Y.; Gerke, J.; Becker, K.; Kuhnert, E.; Verwaaijen, B.; Wibberg, D.; Kalinowski, J.; Stadler, M.; Cox, R. J. Rapid Discovery of Terpene Tailoring Enzymes for Total Biosynthesis. *Chem Sci* **2023**, *14* (46), 13463–13467. https://doi.org/10.1039/D3SC04172G.

(912) Huang, J.; Quest, A.; Cruz-Morales, P.; Deng, K.; Pereira, J. H.; Van Cura, D.; Kakumanu, R.; Baidoo, E. E. K.; Dan, Q.; Chen, Y.; Petzold, C. J.; Northen, T. R.; Adams, P. D.; Clark, D. S.; Balskus, E. P.; Hartwig, J. F.; Mukhopadhyay, A.; Keasling, J. D. Complete Integration of Carbene-Transfer Chemistry into Biosynthesis. *Nature* **2023**, *617* (7960), 403–408. https://doi.org/10.1038/s41586-023-06027-2.

TOC Figure


reconstitution
Synthetic Biology
in Natural Product
Biosynthesis
new scaffolds
and enzymology
natural product
isolation and
diversification