Published in final edited form as:

Chem Rev. 2017 April 26; 117(8): 5334–5366. doi:10.1021/acs.chemrev.6b00683.

The Uncommon Enzymology of Cis-Acyltransferase Assembly Lines

Adrian T. Keatinge-Clay*

Department of Molecular Biosciences, The University of Texas at Austin, Austin, Texas 78712, United States

Abstract

The enzymology of 135 assembly lines containing primarily cis-acyltransferase modules is comprehensively analyzed, with greater attention paid to less common phenomena. Diverse online transformations, in which the substrate and/or product of the reaction is an acyl chain bound to an acyl carrier protein, are classified so that unusual reactions can be compared and underlying assembly-line logic can emerge. As a complement to the chemistry surrounding the loading, extension, and offloading of assembly lines that construct primarily polyketide products, structural aspects of the assembly-line machinery itself are considered. This review of assembly-line phenomena, covering the literature up to 2017, should thus be informative to the modular polyketide synthase novice and expert alike.

Graphical Abstract

1. INTRODUCTION TO ASSEMBLY LINES

Polyketide synthase (PKS) modules are the principal carbon–carbon bond-forming components of enzymatic assembly lines that biosynthesize many important medicines.^{1–8} We still have only a surface-level understanding of the structures and functions of these molecular machines and must look deeper to learn how they truly operate.

*Corresponding Author adriankc@utexas.edu.

Notes

The author declares no competing financial interest.

Within the same assembly line, PKS modules often collaborate with nonribosomal peptide synthetase (NRPS) modules. Usually, PKS modules add carboxylic acids and NRPS modules add amino acids during the extension and processing of an acyl chain into a mature, bioactive compound. This review focuses on assembly lines primarily composed of PKS modules; other recent reviews have centered on NRPS machinery.^{9,10}

The number of assembly lines whose sequences and natural products are known is growing exponentially, thanks to the decreased cost of sequencing. To organize this knowledge, a revolutionary database named Minimum Information about a Biosynthetic Gene Cluster (MIBiG) was recently created.¹¹ Researchers can now readily navigate between the architecture of an assembly line (aided by the program antiSMASH),¹² its gene cluster, all associated nucleic and amino acid sequences, and the literature proposing biosynthetic models for how chemistry is performed on intermediates bound to the assembly line as well as assembly-line products. By connecting diverse assembly lines, MIBiG has converted big data from a liability into an asset.

This review is focused on assembly lines primarily containing modules in which acyltransferases (ATs) are embedded, known as cis-AT modules.⁴ Two of the earliest assembly lines to be studied, those producing the antibacterial erythromycin and the immunosuppressant rapamycin, contain cis-AT modules (as well as one NRPS module in the rapamycin assembly line).¹³ Other, more recently discovered assembly lines, such as mupirocin synthase, contain PKS modules in which ATs are not embedded, known as trans-AT modules.^{5,6} The biosynthetic research community is still working out the basic logic of the diverse trans-AT modules, whereas it has a strong foothold with cis-AT modules.

While the fundamental patterns of cis-AT assembly lines are taught in basic biochemistry courses, many phenomena in cis-AT assembly lines remain to be understood. Through comparison of large numbers of diverse cis-AT assembly lines, deeper architectural and functional patterns of assembly lines have a chance of emerging. Making these connections will allow researchers studying an uncommon architecture or chemistry in one assembly line to benefit from the studies of similar features in other assembly lines. This review will cover the enzymology surrounding transformations conducted on acyl intermediates covalently bound to the assembly lines, as well as that surrounding their loading and offloading, but will not cover the diverse post-assembly-line tailoring reactions.

The well-documented, nonredundant assembly lines in the MIBiG repository primarily composed of cis-AT modules are comprehensively analyzed here (as well as a few assembly lines largely composed of NRPS modules but whose cis-AT modules display intriguing enzymology). These 135 assembly lines collectively contain 1129 PKS modules and mediate the biosynthesis of abyssomicin,¹⁴ ajudazol,¹⁵ akaeolide,¹⁶ althiomycin,¹⁷ ambruticin,¹⁸ amphotericin,¹⁹ anatoxin,²⁰ annimycin,²¹ ansamitocin,²² antimycin,²³ apoptolidin,²⁴ aurafuron,²⁵ aureothin,²⁶ avermectin,²⁷ bafilomycin,²⁸ barbamide,²⁹ BE-14106,³⁰ bengamide,³¹ borrelidin,³² calcimycin,³³ candicidin,³⁴ chalcomycin,³⁵ chaxamycin,³⁶ chlorizidine,³⁷ chlorothricin,³⁸ chondramide,³⁹ chondrochloren,⁴⁰ coelimycin,⁴¹ concanamycin,⁴² conglobatin,⁴³ cremimycin,⁴⁴ crocacin,⁴⁵ cryptophycin,⁴⁶ curacin,⁴⁷ cyclizidine,⁴⁸ cylindrospermopsin,⁴⁹ cystothiazole,⁵⁰ disciformycin,⁵¹ divergolide,⁵²

DKxanthene,⁵³ E-492,⁵⁴ E-837,⁵⁴ ebelactone,⁵⁵ ECO-02301,⁵⁶ elaiophylin,⁵⁷ epothilone,⁵⁸ erythromycin,⁵⁹ FD-891,⁶⁰ filipin,⁶¹ FK520,⁶² fostriecin,⁶³ geldanamycin,⁶⁴ gephyronic acid,⁶⁵ guadinomine,⁶⁵ gulmirecin,⁶⁶ halstoctacosanolide,⁶⁷ haprolid,⁶⁸ hectochlorin,⁶⁹ herbimycin,⁶⁴ herboxidiene,⁷⁰ hitachimycin,⁷¹ hygrocine,⁷² incednine,⁷³ indanomycin,⁷⁴ jamaicamide,⁷⁵ jerangolid,¹⁸ kendomycin,⁷⁶ kijanimicin,⁷⁷ lankamycin,⁷⁸ lasalocid,⁷⁹ leupyrrin,⁸⁰ lipomycin,⁸¹ lobophorin,⁸² lobosamide,⁸³ lorneic acid,¹⁶ macbecin,⁸⁴ maklamicin,⁸⁵ meilingmycin,⁸⁶ melithiazol,⁸⁷ meridamycin,⁸⁸ microcystin,⁸⁹ microsclerodermin,⁹⁰ ML-449,⁹¹ monensin,⁹² mycinamicin,⁹³ mycolactone,⁹⁴ myxalamid,⁹⁴ myxothiazol,⁸⁷ nanchangmycin,⁹⁵ nannocystin,⁹⁶ naphthomycin,²² neoauoreothin,⁹⁷ niddamycin,⁹⁸ nigericin,⁹⁹ nocardiopeptin,¹⁰⁰ oligomycin,⁶¹ nystatin,¹⁰¹ pellasoren,¹⁰² phenylannolone,¹⁰³ phoslactomycin,¹⁰⁴ piericidin,¹⁰⁵ pimaricin,¹⁰⁶ pikromycin,²⁴⁶ pladienolide,¹⁰⁷ puwainaphycin,¹⁰⁸ pyoluteorin,¹⁰⁹ quartromycin,¹¹⁰ rapamycin,¹¹¹ reveromycin,¹¹² rifamycin,¹¹³ rubradirin,¹¹⁴ salinilactam,¹¹⁵ salinomycin,¹¹⁶ sanglifehrin,¹¹⁷ soraphen,¹¹⁸ spinosyn,¹¹⁹ spirangien,¹²⁰ stambomycin,¹²¹ stigmatellin,¹²² streptazone,¹²³ streptolydigin,¹²⁴ tautomycetin,¹²⁵ tautomycin,¹²⁶ tetrocarcin,¹²⁷ tetronasin,¹²⁸ tetronomycin,¹²⁹ thuggacin (*Sorangium cellulosum*),¹³⁰ tiacumicin,¹³¹ tirandamycin,¹³² tubulysin,¹³³ tylosin,¹³⁴ versipelostatin,¹³⁵ vicensin,¹³⁶ yersiniabactin,¹³⁷ and zwittermycin.¹³⁸

The order of discussion is from more common to less common enzymology, with greater attention paid to the less common. The enzymes involved in extension will be analyzed first, in section 2, followed by the enzymes that load an acyl chain onto the first module, in section 3, and then the enzymes that offload the grown acyl chain, in section 4.

2. EXTENSION

Cis-acyltransferase modules minimally contain three domains: an acyltransferase (AT) that selects an α -carboxyacyl extender unit to be added to the acyl chain intermediate, an acyl carrier protein (ACP) to which the AT transfers the extender unit, and a ketosynthase (KS) that performs the carbon-carbon bond-forming reaction between the intermediate and the extender unit-bound ACP (Figure 1). KS and ACP are located at the N- and C-terminal ends of the module, respectively. This definition of a module may soon need updating, as it has been discovered that, during the evolution of assembly lines, the groups of domains that migrate together possess a KS at their C-terminal end, not their N-terminal end.^{139,140}

If an extension module contains only KS, AT, and ACP, it is referred to as an α -module. If it contains a ketoreductase (KR), which reduces the β -keto group formed by KS, it is referred to as a β -module. If it contains both KR and a dehydratase (DH), which eliminates the β -hydroxyl group and an α -hydrogen to form a double bond, it is referred to as a γ -module. If it contains KR, DH, and an enoylreductase (ER), which reduces the double bond to a single bond, it is referred to as a δ -module. This nomenclature holds even if enzymatic domains are inactive. While these four types of extender modules are by far the most common, others exist. The processing enzymes are C-terminal to the KS and AT domains, usually in the order DH + KR_s + ER + KR_c [ER is most commonly inserted between the ketoreductase structural (KR_s) and catalytic (KR_c) subdomains].¹⁴¹ Of the cis-AT modules analyzed in this review, 6% are α -type, 30% are β -type, 47% are γ -type, and 16% are δ -type.

Acyltransferases most commonly select malonyl or (2*S*)-methylmalonyl extender units from the corresponding acyl-CoA. Usually, malonyl-specific ATs possess GHS(I/V)G and HAFH motifs (catalytic residues are shown in boldface type), while methylmalonyl-specific ATs possess GHSQG and YASH motifs.^{142,143} Of the 1129 modules analyzed, 55% specifically incorporate malonyl groups and 39% specifically incorporate methylmalonyl groups. ATs possess an α/β -hydrolase fold and insert into a 120-residue flanking subdomain usually associated with KSs (~90 residues N-terminal to the AT domain and ~30 residues C-terminal to the AT domain; also known as the KS/AT adapter or linker; PDB 2HG4).^{144,145} A loading acyltransferase (AT_L), usually N-terminal to the first module, either selects an extender unit that is converted through decarboxylation into a primer unit or directly selects a primer unit for its assembly line (PDB 4RL1).¹⁴⁶

Ketosynthases form carbon–carbon bonds by a reaction that proceeds through an inversion of configuration, such that D- α -substituents are produced from the L- α -substituents of extender units.^{147,148} Many are thought to serve as gatekeepers, distinguishing between intermediates (e.g., hydrated/dehydrated, epimerized/unepimerized).^{149–152} Ketosynthases possess a thiolase fold and become acylated on a conserved cysteine (TACSSS), which is activated by a conserved histidine (KSNIGHT)¹⁴⁴ (PDB 2HG4). Extension of a polyketide by an α -carboxyacyl-ACP is facilitated by another conserved histidine (HGTGT, ~40 residues N-terminal to the KSNIGHT histidine) that helps catalyze decarboxylative carbon–carbon bond formation. The loading enzyme KS_L (also referred to as KS^Q or KS^S) lacks the cysteine of condensation-competent KSs (a glutamine or serine residue most commonly replaces the cysteine) but catalyzes the decarboxylation of α -carboxyacyl-ACP_L to prepare a primer unit for the first module.¹⁵³ The KSs of cis-AT modules are associated with an ~120-residue flanking subdomain whose function is unknown, but it is also associated with most KSs from trans-AT modules.¹⁴⁵

Ketoreductases perform several catalytic functions. The most common reduction (NADPH-mediated) they perform generates a D- β -hydroxyl group, and a KR possessing this stereoselectivity is referred to as B-type (the D/L system is preferred for polyketide intermediates over the *R/S* system since stereochemical assignments are independent of neighboring groups).^{148,154–157} When a KR generates an L- β -hydroxyl substituent, it is referred to as A-type. The situation is more complex when a KR reduces an α -substituted β -keto intermediate. When the reduction generates a D- α -substituent, a 1 is appended, and when it generates an L- α -substituent, a 2 is appended. Thus, A1-, A2-, B1-, and B2-type KRs are present in cis-AT modules, each with its own amino acid fingerprint.^{155,158} For an L- α -substituent to be generated, an epimerization of the α -substituted β -ketoacyl intermediate must be catalyzed. This is usually also mediated by KR.¹⁵⁴ Ketoreductases that are reductase-incompetent are referred to as C-type, with those that catalyze epimerization being referred to as C2-type¹⁵⁹ (e.g., in pikromycin module 3)¹⁵⁹ and those that do not have any catalytic function being referred to as C1-type. The KR type can be indicated by a subscript (e.g., KR_A, KR_B). Both KR subdomains (KR_s and KR_c) possess a short-chain dehydrogenase/reductase (SDR) fold (PDB 2FR0).¹⁴¹ Usually, a tyrosine (YAAANA) helps catalyze the reduction reaction.

Dehydratases most commonly generate α/β -trans double bonds from D- β -hydroxyacyl (or D- α -substituted D- β -hydroxyacyl) intermediates.^{148,160} Dehydratases are thought to generate cis double bonds when they dehydrate L- β -hydroxy intermediates generated by KR_A (as is thought for DH in annimycin module 4, disciformycin module 1, fostriecin modules 2 and 3, gulmirecin module 1, incednine module 5, lobosamide module 5, phenylannolone module 4, and phoslactomycin module 2). From KR fingerprint analysis, halstoctacosanolide module 15 and phoslactomycin module 1 appear to be generating cis double bonds from intermediates produced by KR_B; however, in the case of KR of phoslactomycin module 1, closer structural and functional analysis revealed that it is actually a KR_A.¹⁶¹ Hydrated and dehydrated intermediates have similar energies; thus a downstream enzyme, such as KS, ER, TE, or a NRPS condensation (C) domain, may be selective for the dehydrated intermediate. Some DHs are hypothesized to help mediate the epimerization of α -substituted β -ketoacyl chains (e.g., DH of FK520 module 4 and rapamycin module 6). Dehydratases possess a double-hotdog fold, contain a catalytic histidine (HxxxGxxxxP) and active-site aspartate (HPALLD), and are dimeric (PDB 3EL6).^{162,163}

Enoylreductases reduce α/β -trans double bonds with NADPH to yield a single bond between the α - and β -carbons. If present, an α -substituent is set in the L-orientation when an L-type ER reduces it and in the D-orientation when a D-type ER reduces it. The presence of a signature tyrosine indicates an L-type ER.¹⁶⁴ Enoylreductases possess a medium-chain dehydrogenase/reductase (MDR) fold (PDB 5DP1)¹⁵⁹ and usually insert into the KR domain between KR_s and KR_c (PDB 3SLK).¹⁵⁸ Biophysical studies of ER from spinosyn module 2 suggest that it is monomeric. Most ERs do not possess the motif employed by MDR enzymes to dimerize.^{165,166}

Modules frequently contain enzymatic domains that do not perform the catalysis expected of them. KR⁰, DH⁰, and ER⁰ are typical, while AT⁰ and KS⁰ are less common (see section 2.1.5 and section 2.1.6). KR⁰, DH⁰, and ER⁰ domains may be catalytically active but not presented with an appropriate substrate, catalytically active with a different function, or catalytically inactive. Why domains that do not perform catalysis during polyketide synthesis have been evolutionarily retained is unclear. One hypothesis is that they are solely performing a structural role; however, many apparently catalytically inactive domains seem to have evolutionarily resisted substitution of their catalytic residues.

Acyl carrier proteins are four-helix domains that possess a serine (DSL) that is post-translationally modified with an ~18-Å phosphopantetheinyl arm by a phosphopantetheinyl transferase (PDB 2LIW).^{167–170} In an assembly line, they shuttle polyketides between the cognate enzymes within a module as well as to the next module or offloading enzyme. Sometimes they must also dock to a trans enzyme for chemistry to be performed on the acyl chain. A loading acyl carrier protein (ACP_L) can either be embedded in the assembly line N-terminal to the first module or encoded on a separate polypeptide and dock to the first module to transfer a primer acyl unit. The high-energy thioester linkage between the growing acyl chain and either ACP or KS enables the polyketide to be readily passed through assembly lines via transthioesterification and eventually offloaded. Nearly all the enzymes discussed here operate on an acyl-ACP substrate. Acyl groups carried by assembly-

line ACPs are not thought to be sequestered within ACP but instead interact with the surface of ACP.¹⁷¹

The crystal structures that have been obtained help in generating a model for homodimeric assembly lines composed of cis-AT modules (Figure 1). The ACPs must diffuse to become charged with an extender unit at AT, receive the acyl intermediate at KS, present the extended intermediate to the processing enzymes (either embedded or not embedded in the assembly line), and transfer the processed intermediate to the next module. Flexible polypeptide linkers N- and C-terminal to ACP provide the freedom for this. From the available crystal structures, KS and DH are dimeric (the DH dimer interface is relatively small) while AT, KR, and ACP are monomeric. ER is apparently monomeric in the majority of modules.¹⁶⁶

While most PKS domains have been structurally characterized, they are commonly represented as circles in biosynthetic model illustrations (here, ATs and ERs are shown branching off to indicate their insertion into KS and ER, respectively). The intermediates accepted by the assembly-line enzymes downstream of a module or by a nonembedded offloading enzyme are shown bound to the phosphopantetheinyl arm of ACP (a mixture of intermediates may exist on the ACP until one is selected by a downstream enzyme). The most common module numbering system is used here: downstream of the loading machinery, both PKS and NRPS modules are counted, whether or not they catalyze an extension of the acyl chain.

2.1. Uncommon Architecture

While the vast majority of cis-AT PKS modules are canonical α -, β -, γ -, or δ -type, exceptions occur, and much can be learned from them (Figure 2).

2.1.1. Unusual Domain Organizations.

2.1.1.1. Methyltransferase Insertion.: Embedded methyltransferases (MTs) are sometimes inserted into α -, β -, γ -, and δ -modules. α -Carbon methyltransferases (CMTs), which methylate the α -carbon of β -ketoacyl intermediates, usually insert into the KR domain after the first β -strand, $\beta 1$ (as observed in anatoxin module 3, cryptophycin module 1, curacin module 7, cylindospermopsin module 1, epothilone module 8, gephyronic acid modules 1–6, hectochlorin module 1, jamaicamide module 2, leupyrrin modules 3 and 5, microcystin modules 1 and 2, puwainaphycin module 1, and yersiniabactin module 3; exceptions are the CMT inserted after ACP of microcystin module 4 and the CMT inserted into KR_C of tubulysin module 5, as discussed in section 2.1.1.5).¹⁴¹ Oxygen methyltransferases (OMTs), which methylate β -hydroxyl and β -keto groups, insert immediately N-terminal to KR if KR is present (as observed in ajudazol module 1, barbamide module 1, crocacin modules 1 and 2, curacin module 9, cystothiazole modules 4 and 5, haprolid module 1, jamaicamide module 6, melithiazol modules 4 and 5, myxothiazol modules 6 and 7, nannocystin module 4, and stigmatellin modules 4 and 5). The OMT positioned after bengamide module 3 methylates an α -hydroxyl group, and the OMT positioned after chondrochloren module 5 may methylate the α -hydroxyl groups of intermediates bound to both chondrochloren modules 4 and 5. Of the cis-AT modules

analyzed here, 1.7% contain CMTs, while another 1.3% contain OMTs. Carbon and oxygen methyltransferases are class I methyltransferases.¹⁶⁶ The structurally characterized CMT from curacin module 7 possesses an equivalent fold to the CMT domain of the mammalian fatty acid synthase (FAS) (PDB 5THZ and 2VZ9; see section 2.3.2).^{172,173} Interestingly, CMTs that have inserted into α -type modules anatoxin module 3 and gephyronic acid module 5 retain KR_s but not KR_c, indicating domain–domain interactions are present, perhaps similar to those observed in the mammalian FAS. In contrast, the ~300-residue OMTs inserted between AT and ACP of the α -type modules cystathiazole module 6 and jamaidamide module 6 are not associated with KR_s.

2.1.1.2. Loading Acetyltransferase Insertion into the First Module.: The loading enzyme AT_L can reside within the first module. From the 135 assembly lines analyzed, each of the 15 examples comes from myxobacteria (ajudazol, aurafuron, bengamide, chondramide, disciformycin, gulmirecin, haprolid, myxalamide, myxothiazol, nannocystin, pellasoren, soraphen, spirangien, stigmatellin, and thuggacin assembly lines). AT_L, composed of both the AT domain and the ~120-residue flanking subdomain, inserts immediately C-terminal to the KS domain into the flexible polypeptide connecting KS and the flanking subdomain of the first module. The location of the insertion lies toward the N-terminal end of the assembly line, appropriate for transferring primer units to ACP_L which is located N-terminal to the KS (also see section 3.1.2.2).

2.1.1.3. Ketosynthase + Acyltransferase + Dehydratase + Acyl Carrier

Protein.: Dehydratase is sometimes present in a module that does not contain other processing enzymes. Of the cis-AT modules analyzed in this review, 1.0% possess KS + AT + DH + ACP architecture (ajudazol module 13, chaxamycin module 2, chondrochloren module 3, cremimycin module 4, curacin module 2, FK520 module 4, naphthomycin module 2, nigericin module 1, salinomycin module 10, stigmatellin module 7, and thuggacin module 2). Dehydratase⁰ appears in the KS_L + AT_L + DH⁰ + ACP_L loading machinery of amphotericin and nystatin assembly lines.

Dehydratase⁰ of FK520 module 4 apparently serves as an epimerase to help convert the orientation (from D to L) of the α -ethyl substituent of a β -ketoacyl intermediate. Similarly, DH⁰ of nigericin module 1 may help epimerize the α -methyl substituent of a β -ketoacyl chain (DH⁰ domains of the γ -modules monensin module 1 and nanchangmycin module 1 apparently also possess this function, as the neighboring KR⁰ domains appear to be catalytically incompetent). The DH of ajudazol module 13 produces the β -keto γ/δ -cis-alkene functionality necessary for the subsequent offloading reaction, in which an isochromanone ring is generated (see section 2.3.4).¹⁵

2.1.1.4. Halogenase within a Module.: For an enzyme besides KS, AT, MT, KR, DH, and ER to be located within the boundaries of a cis-AT module is extremely uncommon. However, in both curacin module 1 and jamaicamide module 1, a halogenase (HAL) is positioned between KS + AT and the three ACPs of these modules. A mysterious 150-residue region is located between KS + AT and what has been crystallographically observed as the start of the 320-residue HAL domain of curacin module 1 (PDB 3NNM) (the

mysterious region is 110 residues in jamaicamide module 1).¹⁷⁴ HAL does possess a small dimeric interface and so could be positioned on the 2-fold axis of the assembly line.

2.1.1.5. Tubulysin Ordering.: Tubulysin modules 3 and 5 show different domain ordering than any of the other cis-AT modules analyzed in this review. In both of these modules, the KR domain is N-terminal to the DH domain. Also in both of these modules, ER appears outside the KR domain boundaries, C-terminal to KR and DH, and appears to be dimeric, possessing an intact MDR dimerization motif.¹⁶⁶ KR and DH may both be monomeric, such that the processing enzymes in these modules do not occupy too much of the 2-fold axis, which could impede the ACPs from accessing their cognate enzymes. A CMT in tubulysin module 5 is also inserted into an unusual location: it is inserted into KR_C rather than into KR_S, immediately after the NADPH-binding motif.

2.1.1.6. Enoylreductase Inserted into Dehydratase.: An unusual δ -module is present in FD-891 module 8. Instead of ER being inserting into the KR domain, it is inserted into the DH domain (N-terminal to the HPALLD-containing helix). It lacks an intact MDR dimerization motif, indicating that it is monomeric.¹⁶⁶ If it is assumed that DH is in a dimeric state, as observed from other cis-AT modules,¹⁶² each inserted ER domain must be properly oriented from DH such that they are accessible to their cognate ACPs.

2.1.1.7. Enzymatic Domains between Modules.: Diverse enzymes are positioned between the modules of an assembly line. Several are C-terminal to a cis-AT module and N-terminal to an NRPS module: an OMT resides between bengamide modules 3 and 4 as well as chondrochloren modules 5 and 6; an aminotransferase (AMT) is present between guadinomine modules 4 and 5 as well as microcystin modules 4 and 5; and both AMT and α -oxidase (OX) are located between microsclerodermin modules 4 and 5 as well as puwainaphycin modules 2 and 3. Enzymes can also be located between PKS modules: an enoyl-CoA-hydrolase-like enzyme (ECH2) involved in β -branching and an ER are present between curacin modules 1 and 2 as well as jamaicamide modules 1 and 2, and two mysterious domains are positioned between ambruticin modules 8 and 9, the first with homology to a thioesterase (TE) and the second with homology to a pyridoxal phosphate (PLP)-dependent AMT (see section 2.4.8).

2.1.2. Dimeric Enoylreductase.—Enoylreductases are known to be dimeric in iterative polyketide synthases, such as mycocerosic acid synthase, and to possess intact MDR dimerization motifs (PDB 5BP4).^{166,175} Some ERs located at the C-terminal ends of polypeptides possess the MDR dimerization motif (ajudazol modules 7 and 9, epothilone module 6, and the epothilone loading machinery). The nucleotide-binding subdomain of ER from phthiocerol module crystallized as a dimer (PDB 1PQW),^{176,177} indicating that this ER is dimeric. With so much space along the 2-fold axis being occupied by the dimeric DH and dimeric ER in these modules, the ACP at the terminus of the polypeptide needs to be free on its C-terminal end to move away from a downstream module (similar to iterative polyketide synthases that possess C-terminal ACPs or monomeric, C-terminal TEs). If it were covalently attached to the next module, it would not be able to access each of its cognate

enzymes unless the flexible linkers were very long.¹⁶⁶ As mentioned in section 2.1.1.5, the ERs of tubulysin modules 3 and 5 are also likely dimeric.

2.1.3. Split Modules.—In contrast to trans-AT PKS modules, cis-AT PKS modules are rarely encoded across two polypeptides.^{178,179} Typically, one cis-AT module associates with its downstream cis-AT module through a 70-residue, three-helix C-terminal docking domain (CDD) located C-terminal to the first module and a 35-residue, one-helix N-terminal docking domain (NDD) located N-terminal to the second module.¹⁸⁰ However, exceptions have been reported (those that appear to be sequencing artifacts are not discussed).

2.1.3.1. Split after Ketosynthase + Acetyltransferase.: A split between KS + AT and DH + ER + KR + ACP occurs in myxalamid module 7. While ~20 residues C-terminal to AT could be involved in docking, no docking residues are apparent N-terminal to DH.

The final portion of the monensin, nanchangmycin, and salinomycin assembly lines is composed of KS⁰ + AT⁰ and a separately encoded ACP. KS⁰ is thought to transfer the polyketide to ACP to facilitate several ensuing tailoring reactions. Docking motifs do not appear to be present C-terminal to AT⁰ or N-terminal to ACP.

Split modules present in barbamide module 1 and jamaicamide module 6 apparently divide KS + AT from OMT + ACP. The sequences C-terminal to KS + AT (~100 residues) and the sequences N-terminal to OMT (~40 residues) align well between these split modules.

2.1.3.2. Split between Ketoreductase and Acyl Carrier Protein.: A unique disconnection between KR and ACP is observed near the junction of a cis-AT module with a trans-AT module (ambruticin modules 7 and 8, respectively). While several residues C-terminal to KR of ambruticin module 7 could play a role in docking, no residues at the N-terminal end of ACP are available. However, ~120 residues between this ACP and KS of module 8 show a repetitive sequence that may be involved in docking the AmbE and AmbF polypeptides to one another. Whether a relationship exists between the split nature of ambruticin module 7 and its iterative behavior is unknown (see section 2.2.2).

2.1.3.3. Split between Ketosynthase + Acyltransferase⁰ + Acyl Carrier Protein and Acyltransferase + Acyl Carrier Protein.: Conglobatin module 1 is highly unusual; however, comparison with the oxazolomycin trans-AT assembly line helps decipher the function of each of its domains.^{6,43,181} KS + AT⁰ + ACP in each assembly line is encoded on one polypeptide (CongB, OzmQ). Equivalent NRPS loading machinery (CongA, OzmO) loads the KS with a formylglycine. In the oxazolomycin assembly line, the oxazoline ring is thought to be formed by the action of the trans-cyclodehydratase OzmP after ACP extends the formylglycine with a malonyl unit. The equivalent trans-cyclodehydratase, CongE, may also operate on the equivalent ACP (the first ACP of conglobatin module 1). While the cyclized intermediate is transferred to the next KS in the oxazolomycin assembly line, it is apparently transferred to the AT in the conglobatin assembly line. This AT resembles an AT_L, without an extender unit-recognizing arginine or the fingerprint residues of a malonyl-specific AT (see section 3.1.2). The second ACP of this unusual module may then transfer the cyclized intermediate to the KS of conglobatin module 2. How this split module is

noncovalently united is unclear. While the ACP from the first polypeptide is not followed by extra residues, ~270 residues with similarity to a flanking subdomain and the portions of KS that contact the flanking subdomain precede AT of the second polypeptide.

2.1.3.4. Split within Dehydratase.: In the BE-14106, cremimycin, and ML-449 assembly lines, the two hotdog subdomains of the DH double-hotdog (BE-14106 module 3, cremimycin module 3, and ML-449 module 4) translated on separate polypeptides apparently noncovalently reassemble into active DHs. Whether this split serves a purpose in each of these assembly lines is unclear; the split does occur in the same module in which a trans enzyme β -aminates and hydrolyzes the acyl chain from the synthase (see section 4.2.4).

2.1.3.5. Split within Ketosynthase.: KS_L of the cyclizidine assembly line is split between two polypeptides.⁴⁸ The split occurs at a surface loop, so the catalytic machinery of this heterodimeric KS_L is apparently not affected.

2.1.4. Incorporation of Trans-Acyltransferase Modules.—A few trans-AT modules have been observed within primarily cis-AT assembly lines. Pyoluteorin module 2 is a trans-AT module, as is ambruticin module 8 (a Favorskii reaction is thought to occur at this module; see section 2.4.8). The neocarzinil assembly line is composed of four modules, in which modules 1 and 3 are cis-AT and modules 2 and 4 are trans-AT.¹⁸² The $KS + AT^0 + ACP$ portion of conglobatin module 1 is equivalent to the first module of the oxazolomycin trans-AT assembly line.^{43,181}

2.1.5. Acyltransferase⁰.—Acyltransferase⁰ sometimes appears where a catalytically competent AT is expected. These domains are mysterious because if they had no function, they would have likely been whittled away through genetic deletions. They still possess a GxxxG sequence in place of the GHSxG motif, but are on average ~130 residues shorter than their catalytically active counterparts (but not the AT^0 from conglobatin module 3, which is full-length). Interestingly, their sequence conservation with catalytically active ATs falls off immediately prior to the ~65-residue ferredoxin-like small subdomain, suggesting that AT^0 retains only the large subdomain of AT. Acyltransferase⁰ is present within bengamide module 3, conglobatin modules 1 and 3, DKxanthene module 1, gephyronic acid modules 2 and 4, piericidin module 4, and the loading machinery of the fostriecin assembly line, each of which must obtain extender units. In these locations, a trans-AT likely fills the role of transferring an extender unit to the neighboring ACP. Indeed, AT^0 -containing zwittermicin module 4 receives an aminomalonyl extender unit from the trans-AT ZmaF. The final modules of several related assembly lines (monensin module 13, nanchangmycin module 14, nigericin module 15, and salinomycin module 15) also possess AT^0 . These nonextending modules apparently transfer acyl chains to an ACP for subsequent reactions and release (see section 2.4.1).¹⁸³ The terminal modules of both the aureothin and neo-aureothin as well as the indanomycin assembly lines are also nonextending and contain an AT^0 .^{74,97} Acyltransferase⁰ may be playing a structural role.¹⁸⁴

The skipped-module maklamicin module 1 is also thought to contain an AT^0 .⁸⁵ This AT^0 is larger than canonical ATs (see section 2.2.3). Acyltransferase⁰ from the skipped-module

curacin module 2 does not appear to operate during assembly-line biosynthesis, yet by sequence inspection it seems no different than a functional malonyl-specific AT.⁴⁷ The AT⁰ domains in the skipped-module microsclerodermin modules 2 and 4 do not possess catalytic serines. In contrast to the usual mode of acquiring KS-bound acyl chains through extender units, the ACPs of maklamicin module 1 and curacin module 2 must collect the KS-bound acyl chain directly through transthioesterification. As microsclerodermin module 2 does not possess an ACP, the ACP of microsclerodermin module 3 may acquire the KS-bound acyl chain through transthioesterification; this interaction of an ACP with KS from the previous module is unique among the discussed assembly lines.

2.1.6. Ketosynthase⁰.—Ketosynthase⁰ is common in assembly lines primarily comprised of trans-AT modules, where it is thought to act as a gatekeeper and/or an enzyme that enables an acyl chain to be transferred from one type of carrier protein to another.^{5,6} These KS⁰ domains lack the first histidine (HGTGT) necessary for chain extension.

A KS⁰ is part of the loading machinery of the DKxanthene assembly line,⁵³ but it is believed not to help in the loading of ACP_L; instead, the loading adenylation (A_L) domain, DkxA, is hypothesized to transfer an L-prolyl group directly to ACP_L C-terminal to KS⁰. This KS⁰ is also mysterious because it is one of the few KSs associated with primarily cis-AT assembly lines that does not contain an AT within its associated flanking subdomain. A KS⁰ may be present in curacin module 2, which is not thought to perform condensation. KS⁰ domains in the loading machinery for cystothiazole and melithiazol assembly lines do not catalyze decarboxylation but do retain the Cys-His-His catalytic machinery. Their function is unclear, since the neighboring AT_L is thought to acquire isobutyryl primer units from isobutyryl-CoA (as expected for AT_Ls, both lack the fingerprint arginine of extender-unit specific ATs; see section 3.1.2), which are passed to the KS of the first module by ACP_L.

Ketosynthase⁰ often has a role in the final module of an assembly line. Ketosynthase⁰ in the final modules of polyether assembly lines (monensin module 13, nanchangmycin module 14, nigericin module 15, and salinomycin module 15) is thought to transfer the fully grown acyl chain to a separately encoded ACP on which several enzymes operate (see section 2.4.1).¹⁸³ Mysterious KS⁰ domains are also present near the C-terminal ends of the ebelactone and indanomycin assembly lines (see section 4.3).^{55,74} Of this large group of C-terminal KS⁰s, the only one that does not possess the Cys-His-His catalytic machinery of a condensation-competent KS is the nanchangmycin KS⁰, in which a glutamine replaces the first histidine. Why these apparent KS⁰ domains usually retain the histidine that facilitates decarboxylation in condensation-competent KSs is unclear.

Intriguing KS⁰ domains are present in microsclerodermin modules 2 and 4.⁹⁰ One follows a module that iterates with malonyl extender units three times to generate a triene moiety, while the other follows a module that iterates with a hydroxymalonyl unit two times to generate an $\alpha,\beta,\gamma,\delta$ -tetrahydroxy moiety (see section 2.2.2). These KS⁰ domains are likely gatekeepers, selecting only the properly iterated intermediate for transthioesterification to the KS of the downstream module.

2.1.7. Ketoreductase⁰ in Loading Machinery.—Embedded between the A_L and ACP_L domains of the loading machinery for the ansamitocin, chaxamycin, divergolide, geldanamycin, herbimycin, hygrocine, kendomycin, macbecin, microsclerodermin, naphthomycin, and rubradirin assembly lines is a KR⁰. No catalytic activity is expected for these domains, and they lack the tyrosine usually present in catalytically competent KRs. The loading machinery of the rapamycin and FK520 assembly lines each include a whittled-down KR⁰ that contains an active ER_L between its KR_s and KR_c subdomains. One hypothesis is that these are playing a structural role, perhaps not so different from many catalytically inactive KR⁰ domains within cis-AT modules.

2.1.8. Tandem Acyl Carrier Proteins.—While tandem ACPs are not unusual in trans-AT assembly lines, they are quite rare in cis-AT assembly lines. Extra ACPs have been hypothesized to increase the flux through that part of the assembly line.^{185,186} The curacin and jamaicamide assembly lines each contain three ACPs in their first module. Each of these ACPs must interact with eight enzymes (two not embedded in the assembly line) during the catalytic cycle. Tandem ACPs collaborate with a KS_L and a trans-AT to supply the first module of the fostriecin assembly line with acetyl primer units. Ambruticin module 8 (a trans-AT module) contains two ACPs, as does chaxamycin module 6. The tandem ACPs within the discussed ambruticin, chaxamycin, curacin, fostriecin, and jamaicamide modules are 80–95% identical to one another.

2.2. Uncommon Module Behavior

2.2.1. Use of Processing Enzyme outside Module Boundaries.—Aside from downstream KS, TE, or NPRS C domains, acyl-ACP substrates can sometimes access embedded enzymes located outside the boundaries of their cis-AT module (Figure 3).

In many assembly lines, a plausible hypothesis had been that β -hydroxy intermediates are dehydrated by DH of the next module (ajudazol modules 4 and 13, curacin modules 5 and 7, epothilone module 5, jamaicamide module 4, myxalamid module 6, spirangien modules 11 and 13, and thuggacin module 7). However, recent studies of the curacin pathway have revealed that dehydration is mediated by DH when the hydroxyl group is in the δ -position. An α/β -trans double bond or β -keto group (ajudazol module 13) may facilitate deprotonation of the L- γ -hydrogen (see section 2.3.4).

2.2.1.1. Enzymes between Modules. The between-module enzymes discussed in section 2.1.1.7 can be accessed by ACPs of the upstream module (OMT between bengamide modules 3 and 4 as well as chondrochloren modules 5 and 6, AMT between guadinomine modules 4 and 5 as well as microcystin modules 4 and 5, AMT and OX between microsclerodermin modules 4 and 5 as well as puwainaphycin modules 2 and 3, ECH2 and ER between curacin modules 1 and 2 as well as jamaicamide modules 1 and 2, and two mysterious domains between ambruticin modules 8 and 9). Many of these enzymes are positioned at polypeptide junctions and may have more freedom to access nearby modules. The α -hydroxyl groups of the acyl chains bound by the ACPs of both chondrochloren modules 4 and 5 may be methylated by the OMT that is embedded C-terminal to

chondrochloren module 5, which is the most C-terminal domain of its polypeptide⁴⁰ (see section 2.3.2.4).

2.2.1.2. Enoylreductase of Previous Module.: Tetrocarcin module 9 (γ -type) does not contain an ER, yet for a [4 + 2]-cycloaddition reaction to occur after offloading, the intermediate passed to tetrocarcin module 10 is expected to be fully reduced at the α - and β -positions (see section 2.3.5).¹²⁷ Instead of proposing a trans-ER activity, the authors hypothesize that ER from tetrocarcin module 8 performs the reduction of the α/β -unsaturated intermediate bound to ACP of tetrocarcin module 9. Phthiocerol module 4 (γ -type) also does not contain an ER but is expected to reduce the β -carbon of the acyl chain it processes to a methylene. Either a trans-ER is responsible or the dimeric ER of the previous module is being accessed.^{166,177}

2.2.2. Iteration.—While this review focuses on processive rather than iterative polyketide synthases, several modules within assembly lines are known to act in an iterative fashion.¹ The first modules of some assembly lines iterate several cycles to provide the lipophilic portion of a lipopeptide such as myxochromide¹⁸⁷ or dihydromaltophylin,¹⁸⁸ while other modules embedded in assembly lines perform either two or three catalytic cycles. Aureothin module 1 and neo-aureothin module 1 operate twice to add a diene to the acyl chain.^{26,97} Microsclerodermin module 3 operates twice to generate an $\alpha,\beta,\gamma,\delta$ -tetrahydroxy moiety.⁹⁰ Ambruticin module 7, DKxanthene module 4, and microsclerodermin module 1 each operates three times to add a triene to the acyl chain.^{18,53,90} Borrelidin module 5 operates three times, each time adding a methylmalonyl group and utilizing each of its processing enzymes (KR, DH, and ER).³² Either stigmatellin module 8 or 9 stutters to help generate a tetraketoacyl chain for the offloading cyclase to cyclize into a chromone ring (see section 4.3.4).¹²² The mechanisms that control the number of iterations of stuttering are not yet well understood, but downstream KS specificity is surely involved, as made apparent by the dedicated KS⁰ domains of the microsclerodermin assembly line (see section 2.1.6).⁹⁰

Crocacin module 3, a γ -module in CroC, was proposed to iterate, first yielding a double bond and then a β -keto intermediate.⁴⁵ However, the ensuing CroD encodes an apparently functional KS and AT that could perform the second extension. While this putative module is likely to extend the once-extended product of crocacin module 3, the location of its ACP is unclear. This ACP would present the β -keto substrate to β -branching enzymes and is predicted to possess a fingerprint common to such ACPs (see section 2.4.7).¹⁸⁶ One more extender unit is incorporated into nanchangmycin than anticipated from its assembly line; one of its final modules, but not its terminal module that contains a KS⁰ and AT⁰, may stutter.⁹⁵

2.2.3. Module Skipping.—Some modules in cis-AT assembly lines seem to be completely skipped such that they neither extend nor process a polyketide. Maklamicin module 1 is apparently skipped.⁸⁵ Its AT⁰ contains an unusual 62-residue insertion and is thought to be inactive (see section 2.1.5). Perhaps the ACP of this module cannot bring an extender unit to KS but can accept the growing polyketide chain from KS as well as transfer it to maklamicin module 2. Curacin module 2 is also skipped.⁴⁷ In the microsclerodermin

assembly line, modules 2 and 4 are skipped and their $KS^0 + AT^0$ architecture is suggestive that their only function is to select the proper intermediate from the preceding iterative modules.⁹⁰ Interestingly, microscloerodermin module 2 does not possess an ACP and may need to interact with the ACP of the following module.

An intriguing example of module skipping comes from the biosynthesis of quartromicin.¹¹⁰ Two products are generated by the five-module assembly line: one generated through the operation of each module and the other generated by each of the modules except the fourth. Both the shorter and longer acyl chain products are necessary to form quartromicin, which is composed of two of each chain. At least two possibilities for this exist: ACP of the fourth module is able to pick up the polyketide from KS of the fourth module without an extender unit half of the time and transfer it to the fifth module, or the polypeptide encoding the fourth module, QmnA2, is excluded from half of the quartromicin assembly lines. Skipping of NRPS modules has also been observed, as in the myxochromide assembly line.¹⁸⁹

2.3. Uncommon Activities Not Requiring Trans Enzymes

Many uncommon phenomena within cis-AT modules result from embedded enzymes that possess unusual activities or are spontaneous (Figures 4 and 5).

2.3.1. Unusual Acyltransferase Function.

2.3.1.1. Uncommon Extender Units.: The most common extender units are malonyl (R = H on the α -carbon) and (2*S*)-methylmalonyl groups (R = CH₃), which are selected by 55% and 39% of the embedded ATs analyzed here, respectively. (2*S*)-Ethylmalonyl (R = CH₂CH₃) and (2*R*)-methoxymalonyl (R = OCH₃) are next most common, comprising 2.4% and 1.6% of embedded ATs analyzed here, respectively. (2*R*)-Hydroxymalonyl (R = OH) and (2*S*)-aminomalonyl (R = NH₂) are quite rare, accounting for 0.5% and 0.1% of embedded ATs analyzed here, respectively. A (2*R*)-hydroxymalonyl group is incorporated by the ATs of bengamide module 2 (how a hydroxymalonyl group is acquired by the following module that contains an AT⁰ is unclear), chondrochloren modules 4 and 5, microscloerodermin module 3, and zwittermicin module 5, while a (2*S*)-aminomalonyl group is incorporated by guadinomine module 4 (the aminomalonyl group incorporated by zwittermicin module 4 is delivered by the trans-AT ZmaF).

Less common extender units are incorporated by the ATs of akaeolide module 6 (R = CH₂CH₂CH₃), reveromycin module 4 (R = CH₂CH₂CH₂CH₃), divergolide module 6 [R = CH₂CH(CH₃)₂], and sanglifehrin module 13 (R = CH₂CH₂COCH₃), as well as filipin module 13 and thuggacin module 11 (R = CH₂CH₂CH₂CH₂CH₂CH₃). The AT of antimycin module 3 is tolerant to a wide variety of extender units.²³ The AT of leupyrrin module 2 selects a 2-carboxy-3-hydroxy-5-methylhexanoyl group, and the AT of chlorizidine module 2 takes a very unusual extender unit containing a dichloropyrrole moiety.³⁷ Extender units are biosynthesized by diverse pathways, although many uncommon extender units are generated via α/β -trans-unsaturated acyl-CoAs through the action of crotonyl-CoA carboxylase/reductases.^{190,191}

2.3.1.2. Promiscuous Acyltransferases.: Some ATs are known to transfer both malonyl and methylmalonyl extender units to their cognate ACPs (elaiophylin module 7, epothilone module 4, naphthomycin module 13, spinosyn module 8, and stambomycin module 12). Other ATs are known to transfer both methylmalonyl and ethylmalonyl extender units to their cognate ACPs (monensin module 5 and phenylannolone module 1). This promiscuity enables the generation of more than one product by the same assembly line and can have an effect on the bioactivity of the natural products generated.⁵⁷

2.3.1.3. Acyltransferases That Appear To Defy the Predictive Model.: The tetrocarcin AT_L and the ATs of tetrocarcin modules 1 and 2 contain the GHSQG and YASH motifs considered to be predictive for methylmalonyl extender units; however, the cognate ACPs appear to be incorporating malonyl units into the tetrocarcin product.¹²⁷ The kijanamicin AT_L and kijanamicin modules 1 and 2 also possess the GHSQG and YASH motifs; the ACP_L does supply a methylmalonyl group for the construction of kijanimicin, but the ACPs of modules 1 and 2 supply malonyl units.⁷⁷ AT from the tetronomycin module 3 also possesses the GHSQG and YASH motifs, yet its cognate ACP supplies a malonyl group for the construction of tetronomycin.¹²⁹ At the sequence level, this AT is quite different from the other methylmalonyl-specific ATs in the tetronomycin assembly line.¹²⁹ One explanation to account for each of these cases is that a trans-AT is supplying malonyl units to ACPs that are not charged by their embedded, cognate AT, as with trans-ATs functioning in modules containing AT⁰ domains (see section 2.1.5). More difficult to explain is the acceptance of a methylmalonyl unit by a module containing an AT predicted to be malonyl-specific. This appears to be the case in DKxanthene module 6, the AT of which contains GHSLG and HAFH motifs.⁵³ One possibility is that a trans-MT adds the methyl group after extension by a malonyl group. Another is that the leucine after the catalytic serine prevents malonyl groups from entering AT (this residue is usually an isoleucine or valine).

Conglobatin module 1 extends an acyl chain with a malonyl unit, yet its AT does not possess the fingerprint residues of a malonyl-specific AT (see section 2.1.3.3).⁴³ The malonyl group may actually be delivered by a trans-AT, as occurs in an equivalent trans-AT module in the oxazolomycin assembly line.^{6,181} The real function of the AT may prove to be unique. By sequence, it possesses the features of a primer-unit selective AT_L and could accept the acyl intermediate that was likely just transformed into an oxazoline by the transcyclodehydratase CongE (see section 3.1). If this is correct, the AT could be playing the role of a gatekeeper.

2.3.2. Methyltransferase.—Embedded MTs are much more common in trans-AT modules; however, occasionally an S-adenosylmethionine (SAM)-dependent MT is incorporated within a cis-AT module (see section 2.1.1.1).¹⁹² Methyltransferases that methylate the α -carbon (CMTs) clade separately from those that methylate hydroxyl/keto substituents (OMTs). While relatively few studies have been conducted on these enzymes, the structure of the CMT from curacin module 7 is known (PDB 5THZ).^{172,193} Of the cis-AT modules analyzed here, 1.9% of modules contain CMTs, while 1.5% contain OMTs (10 operating on β -hydroxyl groups and 6 on β -keto groups); two modules are followed by an OMT that operates on α -hydroxyl groups.

2.3.2.1. α -Carbon Methylation.: Embedded CMTs are present in anatoxin module 3, cryptophycin module 1, curacin module 7, cylindrospermopsin module 1, epothilone module 8, gephyronic acid modules 1–6, hectochlorin module 1, jamaicamide module 2, leupyrrin modules 3 and 5, microcystin modules 1, 2, and 4, puwainaphycin module 1, tubulysin module 5, and yersiniabactin module 3. In vitro studies support that the CMTs of the gephyronic acid modules and curacin module 7 operate on β -ketoacyl substrates;^{172,193} the sequence homology of these CMTs with the aforementioned CMTs is suggestive that they all operate at this stage.

Dimethylation may occur in epothilone module 8, gephyronic acid module 5, and yersiniabactin module 3. The sequences of these CMTs do not diverge significantly from CMTs that only methylate once. Gephyronic acid module 5 is expected to methylate twice since AT of module 5 is predicted to be malonyl-specific. The specificities of the ATs in the epothilone and yersiniabactin modules containing MTs are not clear from their sequence, and epothilone module 8 has been proposed to generate one of the geminal methyl groups from a methylmalonyl extender unit and the other from SAM.¹⁹⁴ The epothilone and yersiniabactin CMTs have also been hypothesized to dimethylate malonyl extender units prior to KS-mediated condensation.¹⁹⁵

In the assembly lines discussed here, 10 out of 16 KRs that function after an MT (curacin module 7, cylindrospermopsin module 1, gephyronic acid modules 1–4, jamaicamide module 2, leupyrrin modules 3 and 5, and puwainaphycin module 1) possess a histidine residue instead of a catalytic tyrosine. Incredibly, KR from microcystin module 1, which is hypothesized to be reductase-competent and contains the expected B-type LDD motif, possesses an alanine instead of a catalytic tyrosine or histidine. A mechanism may need to be in place within these modules to slow ketoreduction relative to methyltransfer; otherwise the methyl group might not be added to the growing chain.

α -Hydrogens whose pK_a s are not decreased by a neighboring β -keto moiety may also be methylated by a special class of CMT. The loading machinery in the saxitoxin pathway contains a CMT_L hypothesized to convert acetyl-ACP to propionyl-ACP.¹⁹⁶ A highly homologous CMT_L in the loading machinery of the gephyronic acid assembly line may dimethylate acetyl-ACP_L to generate an isovaleryl primer unit.⁶⁵ Another possibility is that, in both the saxitoxin and gephyronic acid loading machineries, malonyl-ACP is the substrate for methylation and the product is decarboxylated by the neighboring GCN5-related *N*-acetyltransferase (GNAT_L; see section 3.3).

2.3.2.2. β -Hydroxyl Methylation.: Embedded OMTs that operate on β -hydroxyl substituents reside in crocacin modules 1 and 2, curacin module 9, cystothiazole module 4, haprolid module 1, melithiazol module 4, myxothiazol module 6, nannocystin module 4, and stigmatellin modules 4 and 5. From this small list, L- β -hydroxyacyl substrates are about as common as D- β -hydroxyacyl substrates (crocacin module 2, curacin module 9, haprolid module 1, and stigmatellin module 5 methylate D- β -hydroxyacyl chains, while the others methylate L- β -hydroxyacyl chains). No clear fingerprint is obvious for OMTs that operate on D- versus L- β -hydroxyacyl substrates.

2.3.2.3. β -Keto Methylation. Embedded MTs that operate on β -keto groups reside in ajudazol module 1, barbamide module 1, cystothiazole module 5, jamaicamide module 6, melithiazol module 5, and myxothiazol module 7. As of now, the mechanism for these enzymes is unclear; however, since these OMTs are so similar in sequence to the OMTs that methylate hydroxyl groups (no fingerprint is apparent that differentiates them), they may simply methylate the enol tautomer and not actively participate in the tautomerization. Regardless of the tautomerization mechanism, all but one (ajudazol module 1) of these OMTs appear to trap the cis form of the enol.

2.3.2.4. α -Hydroxyl Methylation. α -Hydroxyl groups incorporated by the hydroxymalonyl-specific ATs of bengamide module 3 and chondrochloren modules 4 and 5 become methylated. The OMT embedded C-terminal to chondrochloren module 5 may methylate the D- α - and D- γ -hydroxyl groups (but not the D- γ -hydroxyl group of the intermediate bound to module 5) or it may methylate the D- α -hydroxyl group of the intermediate bound to module 4 as well as the intermediate bound to module 5.⁴⁰ The position of this enzyme at the C-terminal end of a polypeptide could help it access intermediates bound to the ACPs of both modules (see section 2.2.1.1).

2.3.3. Double-Bond Shift.—Instead of the α/β -trans unsaturated products expected from γ -modules containing a KR_B, β/γ -trans unsaturated products are apparently generated by ambruticin module 4, chaxamycin module 4, chondrochloren module 4, crocacin module 5, cryptophycin module 1, disciformycin module 6, divergolide module 8, gephyronic acid module 1, hygrocin module 8, myxothiazol modules 1 and 2, naphthomycin module 4, rifamycin module 4, rubradirin module 4, and vicanistatin module 5. Whether the final module or tailoring enzymes in the divergolide, hygrocin, and microsclerodermin pathways performs the double-bond shifts observed in the final products is unclear.^{52,72,90} In type II bacterial fatty acid biosynthesis, the dehydratase FabA interconverts β -hydroxydecanoyl-ACP with β/γ -cis-decenoyl-ACP.¹⁹⁷ In the DHs of the aforementioned cis-AT modules, whether β/γ double bonds are generated via regular α/β double-bond formation and subsequent isomerization or directly through an unusual dehydration is unknown. If an isomerization occurs, a downstream enzyme may need to be selective for the isomerized product over the unisomerized substrate with which it is in equilibrium. The KS likely plays this role in most systems; however, a trans-MT takes on this responsibility after a double-bond shift mediated by the DH from ambruticin module 4.¹⁹⁸ Many trans-AT assembly lines possess DHs that both dehydrate and isomerize as well as dedicated enzymes related to DH that perform only the isomerization (referred to as enoyl-isomerase or EI).^{199,200} In trans-AT assembly lines, KSs with unique fingerprints are downstream of the isomerizing enzyme and are hypothesized to be specific for the isomerized product.^{5,6} In the primarily cis-AT assembly lines, not enough sequences are available to detect fingerprints for DHs or downstream KSs suspected to be involved in β/γ -trans double-bond formation. The shifted β/γ -trans- δ/ϵ -trans-diene moiety generated by ansamitocin module 3 is thought to be produced through a DH-catalyzed vinylogous syn dehydration of a β -hydroxy γ/δ -trans-alkene substrate.^{201,202}

The stereochemistry of α -substituents may help decipher how β/γ -trans double bonds are formed. Ambruticin module 4, chaxamycin module 4, disciformycin module 6, gephyronic acid module 1, myxothiazol module 2, naphthomycin module 4, rifamycin module 4, and rubradirin module 4 produce L- α -substituents, while chondrochloren module 4 and cryptophycin module 1 produce D- α -substituents. If α/β -trans double-bonded intermediates are not generated as intermediates during the formation of β/γ -trans double-bonded products, then the orientation of the α -substituent should have been set by the neighboring KRs, which in each of the discussed modules is predicted to be B1-type and yield a D- α -substituent. Since the majority possess L- α -substituents, it seems that most, if not all, of these DHs are catalyzing an isomerization reaction in addition to an α/β -dehydration reaction as observed for FabA.¹⁹⁷

2.3.4. Dehydration of δ -Hydroxy Intermediates.—Like the type B dehydrating bimodules of assembly lines containing primarily trans-AT modules,^{5,6} myxalamid modules 5 and 6, spirangien modules 10 and 11, spirangien modules 12 and 13, and thuggacin modules 6 and 7 possess a KR_A in the first module and a KR_B and DH in the second module. In trans-AT dehydrating bimodules, the DH of the second module may perform two dehydrations on a D- β -L- δ -dihydroxyacyl substrate to yield an α/β -trans- γ/δ -cis-diene. The first dehydration would form an α/β -trans-olefin and facilitate a vinylogous dehydration to yield a γ/δ -cis-olefin. Vinylogous dehydration can be catalyzed by the DHs of curacin modules 5 and 7.²⁰³ An alternative route for δ -hydroxy dehydration, which does not involve a second oxyanion hole, is from the β -keto δ -hydroxy intermediate, as in ajudazol module 13.¹⁵

In the epothilone assembly line, the L-hydroxyl group produced by KR_A of module 4 is dehydrated by DH of module 5 to generate a cis double bond.²⁰⁴ Similar cooperation to form a cis double bond is also observed by ajudazol modules 12 and 13 as well as curacin modules 4 and 5. Ajudazol modules 3 and 4 and jamaicamide modules 3 and 4 cooperate to yield a trans double bond from the D-hydroxyl group generated by the predicted KR_B of the first modules. Both of the double bonds generated through the cooperation of curacin modules 6 and 7 and stigmatellin modules 2 and 3 are trans, with the KR of both modules predicted to be KR_B.

In each of these examples, the stereochemistry of the double bond correlates with the orientation of the hydroxyl group predicted from the KR type.¹⁴⁸ However, KR_B is predicted in both ajudazol modules 5 and 6, yet the product from them, based on the structure of ajudazol, is a diene in which both double bonds possess cis geometry.¹⁵ Either both KRs are actually KR_A (as in the case of phoslactomycin module 1)¹⁶¹ or an uncharacterized isomerase activity is responsible. The DHs of some trans-AT modules may possess such an isomerase activity (macrolactin module 11 is expected to generate a α/β -trans- γ/δ -cis-diene but actually yields a α/β -cis- γ/δ -trans-diene).²⁰⁵ The fingerprint HxxYGxxxxPxxx(H/Y) seems to flag DH domains that perform dehydration of δ -hydroxy intermediates, with the signature residues near the catalytic histidine (PDB 3KG7 and 3KG8).²⁰³

A unique dehydration is mediated by DH in the terminal module of the ajudazol assembly line.¹⁵ A β -ketone may facilitate the deprotonation of the L- γ -hydrogen of a β -keto L- δ -

hydroxy intermediate. The generated β -keto γ/δ -cis-alkene is then cyclized by TE into an isochromanone ring.

2.3.5. Carbocyclization.—All-carbon rings can be produced online through diverse mechanisms. A [4 + 2]-cycloaddition had been hypothesized to occur online during the biosynthesis of the spirotetronates chlorothricin, kijanimicin, lobophorin, maklamicin, tetrocarcin, and versipelostatin.⁸⁵ However, in the chlorothricin pathway and related pyrroindanomycin pathway, the equivalent [4 + 2]-cycloaddition, as well as a subsequent [4 + 2]-cycloaddition, appears to be mediated offline by tailoring enzymes.²⁰⁶ A [4 + 2]-cycloaddition takes place during biosynthesis of the unrelated natural product indanomycin; however, the dienophile may not be generated online (DH of indanomycin module 3 is not expected to yield the required trans double-bond geometry from the KR_A -generated L- β -hydroxyacyl intermediate thought to be produced by indanomycin module 2; see section 2.3.4).⁷⁴ A current, comprehensive review of known [4 + 2]-cycloadditions is available in this issue.²⁵²

2.3.5.1. Cyclohexane through a Cascade Reaction.: A carbocyclization is thought to occur on the final module of tetronasin and tetronomycin assembly lines.^{128,129} Like most other polyketide carbocyclizations, the transformation is enabled by the inherent reactivity of the polyketide chain. The cyclization is part of a cascade reaction (a six-membered pyran is also formed) and is thought to be enzyme-catalyzed. The final modules of both the tetronasin and tetronomycin assembly lines are β -modules that contain KS, AT, KR^0 , and ACP. The KR^0 domains are candidates for catalyzing the reactions. The catalytic YAAANA motif of KR is replaced by a QxxxHA motif in both KR^0 domains, and the LDD motif that helps KR_B orient β -ketoacyl-ACPs is intact in the tetronomycin KR^0 .¹⁴⁸ Cyclization reactions are catalyzed by reductase-incompetent thioester reductase (R^*) domains at the C-terminal end of assembly lines such as the chlorizidine assembly line (see section 4.1.2).³⁷ Alternatively, enzymes involved in tetronate formation (see section 4.2.1) may mediate cyclohexane ring generation.

2.3.5.2. Cyclopropanation by Enoylreductase⁰: A cyclopropyl moiety is formed early in the assembly-line synthesis of curacin. A combination of β -branching enzymes and HAL create a α/β -trans unsaturated γ -chlorinated intermediate that is presented to an embedded ER^0 . The enolate intermediate generated during ER^0 -catalyzed NADPH-mediated reduction intramolecularly attacks the γ -carbon to displace the chloride and form a cyclopropane ring.²⁰⁷

2.3.6. β -Keto to β -Amine.—Embedded pyridoxal phosphate (PLP)-dependent AMTs sometimes reside after a cis-AT module and before an NRPS module. An AMT is present between guadinomine modules 4 and 5 and between microcystin modules 4 and 5 to transaminate a β -keto group into a β -amine.²⁰⁸ Since an aminomalonyl group is incorporated by guadinomine module 4, AMT operates on an α -amino β -keto substrate to generate a highly unusual α,β -diamino product.

Embedded OX domains have been observed immediately C-terminal to an AMT domain, as is the case between microsclerodermin modules 4 and 5⁹⁰ and between puwainaphycin

modules 2 and 3.¹⁰⁸ The AT puwainaphycin module 2 is predicted to be malonyl-specific, so a β -ketoacyl substrate would be generated for OX to α -hydroxylate and AMT to transaminate.

2.3.7. Pyran Formation.—During assembly-line synthesis of the related polyketides ambruticin and jerangolid, six-membered pyrans are formed. In vitro studies of DH from ambruticin module 3 confirm that, in addition to dehydration, it can catalyze pyran formation by enabling the attack of a ζ -hydroxyl group on the β -carbon of an α/β -trans double bond.²⁰⁹ The evolutionarily related jerangolid module 3 likely catalyzes pyran formation in the same manner.¹⁸ DH of herboxidiene module 8 may also catalyze pyran formation.⁷⁰ No sequence fingerprints have been detected to differentiate DHs capable of catalyzing pyran formation. Trans-AT assembly lines often possess pyran synthase (PS) domains that catalyze both furan and pyran formation but not dehydration, even though they are evolutionarily related to DH.²¹⁰ While PSs possess a histidine that corresponds to the catalytic histidine of DH (HxxxGxxxxP), they do not possess an aspartate that corresponds to the active-site aspartate of DH (HPALLD).²¹⁰ This implies that histidine, not aspartate, plays a catalytic role in pyran formation by pyran-generating DHs, potentially mediating stereoselective transfer of the ζ -hydroxy proton to the α -carbon.

During the concerted reaction that occurs in the final module of the tetronasin and tetronomycin assembly lines (see section 2.3.5.1), a hydroxyl substituent may attack a trans double bond to form a pyran ring in the same reaction that yields a cyclohexane ring.¹²⁹ In these final modules, KR⁰ could be responsible for catalyzing these reactions. A pyran is formed from a β,ζ -dihydroxyacyl substrate by the final module of the indanomycin assembly line, which includes KS⁰, AT⁰, and an enzymatic domain termed Cyc.⁷⁴ This enzyme resembles SalBIII, which in the salinomycin pathway forms a pyran ring from a β,ζ -dihydroxyacyl chain through α/β -dehydration followed by Michael addition.²¹¹

2.3.8. Halogenation.—When natural product biosynthesis involves an assembly line, halogenation reactions are most often mediated by tailoring enzymes that react after assembly-line synthesis. However, HAL domains embedded between KS + AT and ACP in the first modules of the curacin and jamaicamide assembly lines introduce chlorine early in acyl chain growth (see section 2.1.1.4).^{47,75} The curacin and jamaicamide HALs are non-heme iron α -ketoglutarate-dependent enzymes that monochlorinate hydroxymethylglutaryl-like intermediates bound to one of three ACPs in the first module; the structure of the curacin HAL has been determined (PDB 3NNM).²⁰⁷ While in the biosynthesis of curacin the chlorine is displaced in an ensuing cyclopropanation reaction (see section 2.3.5.2), the chlorine is retained in the jamaicamide final product. While nonembedded HALs often act on nonembedded acyl-ACP_Ls, a HAL in the hectochlorin pathway is N-terminally fused to the hectochlorin ACP_L; it dichlorinates C5 of a hexanoyl chain bound to ACP_L.⁶⁹

2.3.9. α -Oxidation.— α -Oxidation of acyl intermediates is thought to be catalyzed by an embedded flavin-dependent α -oxidase (OX) domain, which appears C-terminal to an embedded AMT domain between cis-AT and NRPS modules in both the microsclerodermin and puwainaphycin assembly lines.^{90,108} The OX between microsclerodermin modules 4 and 5 may operate on an acyl chain that already possesses an α -hydroxyl group. Embedded

OXs are observed without an accompanying AMT domain in trans-AT assembly lines (e.g., OX embedded in the patellazole assembly line, 41% identical to the microsclerodermin OX).²¹² Whether they oxidize β -keto intermediates is not yet clear; all reported OXs collaborate with modules (either cis-AT or trans-AT) that possess other active processing enzymes.^{5,6}

2.4. Uncommon Activities Requiring Trans Enzymes

Many uncommon phenomena within cis-AT modules result from trans enzymes that possess unusual activities (Figure 6).

2.4.1. Epoxidation, Cyclization, and Further Modifications of Polyethers.—In the monensin, nanchangmycin, nigericin, and salinomycin assembly lines, epoxidation and cyclization reactions are thought to occur while the polyketide is tethered to ACP of the final module, which also contains KS⁰ and AT⁰.¹⁸³ This ACP is not covalently attached to the assembly line and thus has the potential to diffuse away and be more accessible to trans enzymes. In the nanchangmycin pathway, the flavin-dependent epoxidase NanO (homologues MonC1, NigC1, SalC, as well as Lsd18 from the lasalocid pathway) stereospecifically operates on the trans double bonds of the acyl chain, and the epoxide hydrolase NanI (homologues MonBI/MonBII, NigBI/NigBII, SalBI/SalBII, as well as Lsd19 from the lasalocid pathway) conducts the cyclization reactions. The cyclization reactions are thought to resemble those conducted by Lsd19 in the lasalocid pathway even though this tailoring enzyme is not thought to operate on ACP-bound substrates. A crystal structure of this enzyme bound at each of its active sites by a substrate/product mimic (PDB 3RGA)²¹³ has provided clues as to how it facilitates anti-Baldwin ring closure. Whether other enzymes that modify the polyether substrate, such as OMTs, P450 hydroxylases, glycosyltransferases, and pyran-forming enzymes, also operate on the acyl-ACP intermediate before release from the ACP by TEII (see section 4.2.3) is unknown; however, the salinomycin pyran-forming enzyme, SalBIII, as well as the salinomycin hydroxylases, SlnE and SlnF, are each hypothesized to operate on an acyl-ACP intermediate.^{116,211}

2.4.2. Oxidation/Cyclization in Naphthalenic Ansamycins.—In the assembly-line biosynthesis of naphthalenic ansamycins (chaxamycin, hygrocine, naphthomycin, rifamycin, and rubradirin), an online oxidation of the 3-amino-5-hydroxybenzoic acid (AHBA) ring takes place.²⁰² Cyclization between the γ -carbon of the β,δ -diketoacyl chain and the ring follows this oxidation, although this may not be enzyme-catalyzed. In both naphthomycin and rifamycin pathways, deletion of the flavin-dependent monooxygenase responsible for AHBA oxidation (Nat2, Rif-Orf19, Cxm19, Hgc2, RubP1) resulted in an accumulation of tetraketide shunt products, suggesting that the hydroxylase functions while the intermediate is bound to ACP of the third module in each naphthalenic ansamycin assembly line and that acceptance of the polyketide by KS of the fourth module may be dependent on this modification.^{22,52} In each naphthalenic assembly line, the ACP thought to be bound to the acyl chain substrate during oxidation/cyclization is the most C-terminal domain of the first assembly-line polypeptide, which houses the loading machinery and the first three modules.

The related hydroxylase DivE operates in the assembly-line biosynthesis of divergolides, a collection of benzenic and naphthalenic ansamycins. Oxidation does not proceed to the

quinone stage for benzenic divergolides but does for naphthalenic divergolides. Perhaps the gatekeeping activity of the downstream KS in divergolide module 4 is not robust. While oxidation occurs at the same 2-position of the AHBA ring in the benzenic ansamycins (geldanamycin, herbimycin, macbecin), this is thought to occur after the polyketide has been released from these assembly lines.^{84,202}

2.4.3. Trans-Acyltransferase, Methyltransferase, Ketoreductase, and Dehydratase.—While trans-AT, MT, KR, DH, and ER enzymes commonly cooperate with trans-AT assembly lines, they rarely participate in biosyntheses orchestrated by cis-AT assembly lines.^{5,6}

2.4.3.1. Trans-Acyltransferase.: Bengamide module 3, conglobatin module 3, DKxanthene module 1, gephyronic acid modules 2 and 4, and piericidin module 4 contain AT⁰ domains where ATs are anticipated, and the fostriecin assembly line contains an AT⁰ where an AT_L is anticipated. Acyltransferases⁰ do not possess the catalytic serine necessary for transferring extender units to ACPs (see section 2.1.5). Since the neighboring ACPs still receive extender units, trans-ATs likely supply them. These ACPs usually receive malonyl units, and the overwhelming majority of trans-ATs are malonyl-specific; whether a trans-AT supplies conglobatin module 3 with a methylmalonyl unit or bengamide module 3 with a hydroxymalonyl unit is not known. Zwittermicin module 4 adds an aminomalonyl group to the growing acyl chain; however, it possesses an AT⁰ and its ACP actually receives the aminomalonyl unit from the trans-AT ZmaF. As discussed in section 2.3.1.3, some ATs predicted to be methylmalonyl-specific reside in modules that incorporate malonyl units into the growing polyketide. Perhaps ACPs in these modules are also receiving extender units from trans-ATs. Trans-ATs are also required for the trans-AT modules present within primarily cis-AT assembly lines (ambruticin module 8, conglobatin module 1, neocarzinil modules 2 and 4, and pyoluteorin module 2; see section 2.1.4).

2.4.3.2. Trans-Methyltransferase.: Ambruticin module 4 and jerangolid module 4 contain malonyl-specific ATs and rely on trans-CMTs AmbC and JerC to deliver an α -methyl group^{18,209} Fascinatingly, these CMTs methylate β/γ -unsaturated intermediates formed by a double-bond shift (see section 2.3.4).²¹⁴ Althiomycin module 5 generates a β -ketoacyl intermediate that is methylated in the cis-enol form by trans-OMT AlmC.¹⁷ The cis geometry may enable subsequent TE-mediated cyclization of an amide nitrogen to C1.

2.4.3.3. Trans-Ketoreductase and Trans-Dehydratase.: AntM is the trans-KR that reduces the β -ketoacyl chains bound to ACP of antimycin module 3 (the homologue NatF aids in neoantimycin biosynthesis).²¹⁵ AntM is highly homologous to SiaM, the trans-KR that reduces β -keto groups of acyl chains bound to the SIA7248 trans-AT assembly line and whose structure is known (PDB 3WOH).^{216,217}

A trans-KR may also be operating in thuggacin (*Chondromyces crocatus*) module 2.¹³⁰ Thuggacin (*C. crocatus*) module 2 is thought to generate an α/β -trans double bond, like thuggacin (*S. cellulorum*) module 2. However, in contrast to thuggacin (*S. cellulorum*) module 2, thuggacin (*C. crocatus*) module 2 lacks a KR domain. The most likely possibilities are that a KR from a neighboring module or a trans-KR is active. As there are

no precedents for KRs acting from a neighboring module and there are precedents for trans-KR activity, a trans-KR is probably operating at thuggacin (*C. crocatus*) module 2 prior to the apparently functional DH embedded in that module.

Neither chalcomycin module 6 nor DKxanthene module 1 contains the processing enzymes necessary to yield an α/β -trans double bond. Trans-KRs and trans-DHs may provide the missing activities.^{35,53}

2.4.4. Hydroxylation.—The macrolactone ester of stambomycin is formed when the C50-hydroxyl group connects with C1 in a TE-mediated cyclization.²¹⁸ The hydroxyl group on C50 is not generated by embedded enzymes, since C50 corresponds to the α -carbon of the methylmalonyl unit selected by AT_L and decarboxylated by KS_L. Thus, one of the two P450 hydroxylases encoded in the stambomycin gene cluster (SamR0478 or SamR0479) is thought to hydroxylate the acyl chain while bound to an ACP (the other could also operate online to generate the C28 hydroxyl group). At what point in the assembly line this occurs is unknown. The hydroxyl group that becomes part of the macrolactone ester of nocardopsin must also be installed during assembly-line biosynthesis.¹⁰⁰ A double bond generated by nocardopsin module 1 is epoxidized by the P450 epoxidase NsnF and the epoxide hydrolase NsnG to yield a vicinal diol. To which ACP the intermediate is bound during these reactions has not yet been determined. The P450 hydroxylase SlnF is also proposed to operate on the acyl chain bound to the final ACP in the salinomycin assembly line before its TEII-mediated hydrolysis from that ACP.¹¹⁶

A shunt product of the cyclizidine assembly line reveals that the terminal carbon of the acyl chain is hydroxylated (and then acetylated) before it is transferred to cyclizidine module 5. CycN, homologous to the ribonucleotide reductase β -subunit, which contains a diferric iron center, is thought to conduct the hydroxylation.⁴⁸

2.4.5. Deacylation.—In several pathways that result in the formation of a macrolactam (BE-14106, cremimycin, hitachimycin, incednine, lobosamide, ML-449, salinilactam, and vicenistatin), the amino group is acylated (by an amino acid) so that it does not perform any off-pathway chemistry as the growing chain travels through the assembly line. At the final module, a deacylase is hypothesized to remove the amino acid protecting group. In the cremimycin pathway the deacylase is CmiM6 (homologues BecP, HitF, IdnL5, LobG, MlaP, Strop_2777, Vinj), an α/β -hydrolase that shows high similarity to amino acid amidases.⁴⁴

2.4.6. β -Amination of α/β -Trans-Unsaturated Intermediates.—A few characterized macrolactam-producing assembly lines (BE-449, cremimycin, and ML-449) generate a β -amine from an α/β -unsaturated ACP-bound intermediate. The α/β -trans-unsaturated intermediate bound to the ACP of cremimycin module 3 is a substrate for CmiS1, a type III thioesterase (TEIII),²¹⁹ which both performs the Michael addition of glycine and hydrolyzes the thioester linkage to the assembly line.⁴⁴ The flavin-dependent enzyme CmiS2 then oxidizes the intermediate to generate the β -amine moiety and glyoxylate as a byproduct.⁴⁴ Several downstream enzymes help acylate the amine and load the chain onto a subsequent assembly line that extends and eventually lactamizes it through the introduced amine, which is deprotected prior to reaching the TE (see section 2.4.5).

Similar chemistry occurs in the BE-14106 and ML-449 pathways; however, in these pathways the action of the trans enzymes shifts the remaining double bond(s) (of which BE-14106 and ML-449 products contain one and two, respectively) away from the β -amine.^{30,91}

2.4.7. β -Branching.— β -Branching is quite common in trans-AT assembly lines; however, it also occurs in a few cis-AT assembly lines.^{5,6} The process involves several enzymes. Those interfacing directly with the β -ketoacyl-ACP substrate are (1) an enzyme homologous to a hydroxymethylglutaryl-CoA synthase (HCS) that usually adds an acetyl group at the β -position (the structure of the curacin HCS, CurD, is known; PDB 5KP8),²²⁰ (2) an enzyme homologous to an enoyl-CoA hydratase (ECH1) that forms an α/β double bond through elimination of the β -hydroxyl group, and (3) another enzyme homologous to enoyl-CoA hydratase (ECH2) that can decarboxylate what is usually an acetyl group to leave a carbon where the β -keto oxygen had been. Either an α/β or a β/γ double bond is formed through these reactions. The structure of the ECH2 embedded in the curacin assembly line (PDB 2Q2X)²²¹ helps elucidate how ECH2s catalyze decarboxylation.

β -Branching proceeds similarly in curacin module 1 and jamaicamide module 1. After the formation of β -ketoacyl-ACP intermediate by KS, the chain is acted on by a nonembedded HCS (CurD, JamH), an embedded halogenase (HAL), a nonembedded ECH1 (CurE, JamI), and an embedded ECH2.²²² The curacin intermediate is then a substrate for ER⁰-mediated cyclopropanation (see section 2.3.5.2).

β -Branching also occurs in the crocacin assembly line, although the details remain mysterious.⁴⁵ The KS and AT encoded on CroD likely generate a β -ketoacyl substrate for β -branching (see section 2.2.2); however, the ACP that presents this substrate to the β -branching enzymes (HCS, ECH1, and ECH2 correspond to CroE, CroF, and CroG) seems not to be located in the crocacin gene cluster (no ACP cognate for the KS and AT of CroD is evident). This ACP is anticipated to possess a strong sequence fingerprint.¹⁸⁶

2.4.8. Favorskii Rearrangement.—A cyclopropane ring is thought to be formed through a cascade reaction that involves a Favorskii rearrangement on the acyl chain attached to one of the two ACPs in ambruticin module 8, possibly mediated by the nonembedded flavin monooxygenase AmbI.¹⁸ Immediately C-terminal to ambruticin module 8 are two mysterious domains, the first with homology to a TE and the second with homology to a PLP-dependent AMT, that could help resolve the product of the Favorskii rearrangement. The A + ACP didomain encoded by AmbG may also participate. A Favorskii rearrangement is performed on a polyketide chain by the flavin monooxygenase EncM during the biosynthesis of enterocin.²²³

2.5. Remarkable Transformations That May Occur Online

Several remarkable transformations in the assembly lines discussed here merit further investigation. Two examples are the cyclizations that occur in the synthesis of lorneic acid and anatoxin. In lorneic acid biosynthesis, a P450 enzyme is thought to epoxidize a trans double bond to enable the conversion of a neighboring triene moiety into a benzyl ring.¹⁶ In anatoxin biosynthesis, cyclization to the oxidized prolyl primer unit to yield a cyclooctene

ring is hypothesized to occur online.²⁰ While the timing of the cyclization is not clear, the putative cyclase AnaJ may be involved. Interestingly, AnaJ is homologous to the embedded Cyc domain that mediates cyclization-coupled offloading from the stigmatellin assembly line (see section 4.3.4).¹²²

3. LOADING

Primer units are loaded onto the first module of an assembly line through a variety of loading machineries (Figure 7). Here, the loading machineries that supply primer units to cis-AT modules are discussed (129 of the 135 assembly lines covered in this review; the althiomycin, antimycin, leupyrin, tubulysin, yersiniabactin, and zwittermicin loading machineries supply primer units to NRPS modules). Three different types of loading machinery have been identified: AT_L-mediated (72%), detailed in section 3.1; A_L-mediated (26%), covered in section 3.2; and GNAT_L-mediated (2%), described in section 3.3.

3.1. Loading Acyltransferase

In these loading machineries, an AT_L transfers an acyl unit to ACP_L that will eventually transfer a primer unit to KS of the first module. The majority of AT_Ls select extender units versus primer units (60% vs 40%), and the majority of AT_Ls are embedded within the assembly line (89%).

3.1.1. Loading Acyltransferases That Select an Extender Unit for Loading Ketosynthase-Mediated Decarboxylation into Primer Unit.—

Loading acyltransferases that select extender units are apparently structurally and functionally equivalent to ATs that select extender units within modules. Loading ketosynthase, KS_L (also referred to as KS^Q or KS^S), primarily differs from KS within a module in that it possesses a residue such as a glutamine or serine in place of the active-site cysteine and behaves like an acylated KS, decarboxylating ACP_L-bound extender units.¹⁵³ The KS_L + AT_L + ACP_L architecture is equivalent to an α -module and is often referred to as a “loading module”.

The malonyl group is the most common extender unit selected by AT_L (70% malonyl, 29% methylmalonyl, 1% methoxymalonyl) and is decarboxylated by an accompanying KS_L into an acetyl primer unit (as in the abyssomicin, akaeolide, amphotericin, chalcomycin, chlorothricin, coelimycin, concanamycin, E-492, E-837, ebelactone, elaiophyllin, epothilone, FD-891, filipin, fostriecin, kijanimicin, lasalocid, lobophorin, lorneic acid, maklamicin, meridamycin, monensin, nanchangmycin, niddamycin, nigericin, nystatin, oligomycin, pimaricin, quartromycin, salinomycin, streptazone, streptolydigin, tautomycetin, tetrocarcin, tetronasin, tetronomycin, tirandamycin, versipelostatin, and both mycolactone assembly lines). Methylmalonyl extender units are also commonly selected by AT_L and are decarboxylated by the accompanying KS_L into propionyl primer units (as in the ambruticin, BE-14106, cremimycin, cyclizidine, halstoctacosanolide, herboxidiene, jerangolide, ML-449, mycinamycin, nocardiopepsin, pikromycin, pladienolide, reveromycin, spinosyn, stambomycin, and tylosin assembly lines). A methoxymalonyl group can also be selected by an AT_L and decarboxylated by its accompanying KS_L (as in the apoptolidin assembly line).

After the extender unit bound to ACP_L has been converted to a primer unit, it is transferred by ACP_L to the first module of the assembly line.

The pimarin assembly line may possess two loading machineries.¹⁰⁶ Its KS_L + AT_L + ACP_L is preceded by an A_L and another ACP_L. Whether this loading machinery can also supply primer units to pimarin module 1 is not known. The epothilone assembly line possesses the architecture KS_L + AT_L + ER⁰ + ACP_L; the role of ER⁰ is unknown.⁵⁸

3.1.2. Loading Acyltransferases That Directly Select Primer Units.—Loading acyltransferases selective for primer units differ from those selective for extender units. The most obvious difference is the lack of the arginine 25 residues C-terminal to the catalytic serine that, in an extender unit-accepting AT, forms a salt bridge with the extender unit α -carboxyl group. A crystal structure (PDB 4RL1)¹⁴⁶ reveals how a tryptophan replaces this arginine in the avermectin AT_L. Other residues in the active site form complementary interactions with the acyl group they help select.

3.1.2.1. Loading Acyltransferase + Loading Acyl Carrier Protein.—The combination of a primer unit-selective AT_L and an ACP_L N-terminal to the first module is often referred to as a “loading didomain”. These AT_Ls are selective for a range of acyl groups: the meilingmycin and piericidin AT_Ls select acetyl-CoA; the animycin, erythromycin, and tiacumicin AT_Ls select propionyl-CoA; the bafilomycin, lipomycin, and tautomycin AT_Ls select isobutyryl-CoA; the avermectin and lankamycin AT_Ls select *s*-pentanoyl-CoA; the phoslactomycin AT_L selects cyclohexanoyl-CoA; and the borrelidin AT_L selects *trans*-cyclopentane-(1*R*,2*R*)-dicarboxylic acid-CoA

In the related melithiazol and cystothiazole assembly lines, a mysterious KS⁰ is present within the loading machinery. In both assembly lines, AT_L and ACP_L seem to be operating as if they were the only domains N-terminal to the first module, with isobutyryl groups transferring from isobutyryl-CoA to AT_L to ACP_L to the KS of the first module.^{50,87}

3.1.2.2. Loading Acyl Carrier Protein + Ketosynthase + Loading Acyltransferase + Acyltransferase + Processing Enzymes + Acyl Carrier Protein. As described in section 2.1.1.2, AT_L is sometimes inserted between KS and the flanking subdomain of the first module, while its cognate ACP_L is located N-terminal to the first module. These AT_Ls are selective for a range of acyl groups: the ajudazol, aurafuron, chondramide, disciformycin, gulmirecin, haprolid, spirangien, stigmatellin, and thuggacin AT_Ls select acetyl-CoA; the pellasoren AT_L selects propionyl-CoA; the myxalamide AT_L selects isobutyryl-CoA; the bengamide AT_L selects both isobutyryl- and 2-methylbutyryl-CoA; the aurafuron and myxothiazol AT_Ls select isopentyl-CoA; and the soraphen and nannocystin AT_Ls select benzoyl-CoA. Each example of the ACP_L + KS + AT_L + AT + processing enzymes + ACP architecture is from a myxobacterial assembly line.

3.1.3. Nonembedded Loading Acyltransferase Transfers to Embedded Loading Acyl Carrier Protein.— β -Aminoacyl-ACPs are generated in the construction of polyene macrolactams either at the beginning of their biosynthesis (as in the hitachimycin, incednine, lobosamide, salinilactam, and vicienistatin pathways) or in the middle of their

biosynthesis (as in the BE-14106, cremimycin, and ML-449 pathways). The chain becomes acylated on its β -amino group and is then transferred by a nonembedded AT_L to an ACP_L embedded immediately N-terminal to a module. These ~320-residue AT_L s (HitC, IdnL2, LobN, Strop_2773, BecK, CmiS5, MlaK, and VinK) are highly similar to one another. These ATs are not accompanied by the flanking subdomain. The structure of VinK crosslinked to its ACP substrate (PDB 5CZD)²²⁴ may reveal the general interface between ATs and ACPs.

Nonembedded AT_L s may also accept acyl groups from acyl-CoAs. The nonembedded ECO-02301 AT_L , Orf18, is thought to transfer a γ -4-guanidinobutyryl primer unit from 4-guanidinobutyryl-CoA to the embedded ECO-02301 ACP_L .²²⁵ This AT_L is highly similar to the AT_L s of the polyene macrolactam pathways. The nonembedded sangliferin AT_L , SfaL, is hypothesized to transfer a (2*R*)-2-ethyl-malonamyl group from (2*R*)-2-ethylmalonamyl-CoA to the embedded sangliferin ACP_L .²²⁶

3.2. Loading Adenylation Domain

Adenylate-forming enzymes often participate in loading the primer unit onto the first module of an assembly line. These enzymes are known by many names (CoA-ligase, adenylate forming domain, acyl-ACP synthetase, AMP-dependent ligase, etc.) but are equivalent to NRPS adenylation (A) domains, and that is how they are referred to here. More A_L s are embedded than not (68% vs 32%).

3.2.1. Embedded Loading Adenylation Domain.—The loading machineries of the ansamitocin, barbamide, candicidin, chaxamycin, chondrochloren, conglobatin, crocacin, cryptophycin, cylindrospermopsin, divergolide, FK520, geldanamycin, guadinomine, herbimycin, hygrocine, kendomycin, macbecin, microcystin, microsclerodermin, naphthomycin, phenylannolone, rapamycin, rifamycin, and rubradirin assembly lines contain an embedded A_L and ACP_L . The A_L adenylates a carboxylic acid, enabling the phosphopantetheinyl arm of ACP_L to acquire it. The acyl chain is next passed to the KS of the first module, except in the FK520 and rapamycin assembly lines, in which it is first reduced by an ER_L .

The ansamitocin, chaxamycin, divergolide, geldanamycin, herbimycin, hygrocine, kendomycin, macbecin, microsclerodermin, naphthomycin, and rubradirin loading machineries possess a KR^0 between A_L and ACP_L . The sequence conservation of these KR^0 domains is surprisingly high. The rapamycin and FK520 assembly lines also possess KR^0 domains with a functional ER_L embedded between KR_S and KR_C (see section 2.1.7).

Most embedded A_L s transfer a ring-containing acyl group. 3-Amino-5-hydroxybenzoic acid (AHBA) is selected by the ansamitocin, chaxamycin, divergolide, geldanamycin, herbimycin, hygrocine, macbecin, naphthomycin, rifamycin, and rubradirin A_L s. *trans*-Cinnamic acid is selected by the crocacin, cryptophycin, and phenylannolone A_L s. 4-Aminobenzoic acid is selected by the candicidin A_L . 2,3,5,6-Tetrahydroxy-4-methylbenzoic acid, generated by an accompanying type III polyketide synthase, is selected by kendomycin A_L .⁷⁶ Phenyl-propanoid, not phenylacetate, is selected by microcystin and microsclerodermin A_L s.²²⁷ 4,5-Dihydroxycyclohex-1-enecarboxylic acid is selected by the FK520 and rapamycin A_L s (their ER_L s can reduce the α/β -double bond of an ACP_L -loaded

4,5-dihydroxycyclohex-1-enyl moiety).²²⁸ Ringless carboxylic acids are also selected by embedded A_Ls: guanidinoacetate is selected by the cylindrospermopsin and guadinomine A_Ls,^{49,208} butyric acid is selected by the chondrochloren A_L,⁴⁰ and a trichlorinated L-leucine is probably selected by the barbamide A_L.²⁹

3.2.2. Nonembedded Loading Adenylation Domain Transfers to

Nonembedded Loading Acyl Carrier Protein.—Nonembedded A_Ls operate very similarly to embedded A_Ls; however, in the majority of the pathways reviewed here, they transfer acyl groups to nonembedded ACP_Ls (A_L/ACP_L pairs: anatoxin, AnaC/AnaD; jamaicamide, JamA/JamC; chlorizidine, Clz14/Clz18; calcimycin, CalN2/CalN3; indanomycin, IdmJ/IdmK; puwainaphycin, PuwC/PuwD; and pyoluteorin, PltF/PltL). The acyl-ACP_L then docks in trans to the KS of the first module to provide the primer unit to the assembly line.

A few pathways do not fit this mold. The DKxanthene A_L, DkxA, transfers to an ACP_L that is part of a larger polypeptide; a KS⁰ without an obvious role is fused to its N-terminus (see section 2.1.6). The hectochlorin A_L, HctA, also transfers to an ACP_L that is part of a larger polypeptide; a HAL is fused to its N-terminus. The ACP_Ls to which the aureothin and neo-aureothin A_Ls, AurE and NorE, transfer acyl groups have not been identified; mysteriously, in both cases, a seemingly nonfunctional ACP (no DSL serine) is present N-terminal to the first module.

L-Proline is frequently selected by nonembedded A_Ls, as in the anatoxin, calcimycin, chlorizidine, DKxanthene, indanomycin, and pyoluteorin pathways. 4-Nitrobenzoic acid is selected by aureothin and neo-aureothin A_Ls.⁹⁷ 5-Hexynoic acid is selected by the jamaicamide A_L JamA;⁷⁵ hexanoic acid is selected by the hectochlorin A_L, HctA;⁶⁹ and fatty acids of variable lengths are selected by the puwainaphycin A_L, PuwC.¹⁰⁸

3.3. Loading GCN5 N-Acetyltransferase-like Enzyme

In two of the 135 assembly lines analyzed in this review (the curacin and gephyronic acid assembly lines), a GNAT_L domain mediates loading. The curacin GNAT_L is proposed to select malonyl-CoA, catalyzes its decarboxylation to acetyl-CoA, and then mediate transfer of the acetyl group from acetyl-CoA to ACP_L.⁴⁷ While the crystal structure of the curacin GNAT_L bound to acetyl-CoA (PDB 2REF) elucidates many details about its function, the purpose of an ~180-residue adapter (AR) domain N-terminal to the curacin GNAT_L domain is less clear.²²⁹

The gephyronic acid assembly line also possesses a GNAT_L.⁶⁵ N-Terminal to the GNAT_L, a ~490-residue region is predicted by the Fold and Function Assignment System (FFAS)²³⁰ to be similar to a dimeric SAM-dependent CMT (PDB 4KIB).²³¹ The N-terminal ~120 residues of this CMT mediate dimerization. Interestingly, the curacin AR domain aligns with the first ~180 residues of the gephyronic acid synthase CMT. Thus, the curacin AR may help mediate dimerization. A scheme was proposed that is in keeping with the function of GNAT_L in the curacin pathway and the function of CMT_L + GNAT_L + ACP_L hypothesized for the saxitoxin pathway (in which acetyl-ACP_L is converted to propionyl-ACP_L through methylation):¹⁹⁶ the gephyronic acid GNAT_L decarboxylates malonyl-CoA and transfers the

acetyl group to ACP_L, after which CMT_L converts acetyl-ACP_L to isovaleryl-ACP_L through two methylations. However, it is also possible that dimethylation of malonyl-ACP precedes GNAT_L-mediated decarboxylation. The CMT_L + GNAT_L + ACP_L machineries from the gephyronic acid and saxitoxin pathways align extremely well, as do the curacin AR and the first ~180 residues of the saxitoxin CMT_L.

4. OFFLOADING

The embedded and nonembedded domains that help break the covalent connection between the grown acyl chain and the assembly line are referred to as offloading machinery (Figures 8 and 9). Although offloading machinery collaborating with a cis-AT module is discussed here, general strategies for PKS and NRPS assembly line offloading have been recently reviewed.²³² Usually, the offloading machinery is embedded (71% embedded versus 29% nonembedded, not considering the mysterious offloading machineries discussed in section 4.3).

4.1. Embedded Enzymes

A thioesterase (TE) or reductase (R) domain located at the C-terminus of an assembly line can catalyze the removal of an acyl chain. Of the 72 discussed assembly lines with embedded offloading machinery, 94% are TEs and 6% are R domains.

4.1.1. Thioesterase.—Thioesterases are α/β -hydrolases and harbor a catalytic serine in a G(H/Y)SxG motif. The serine acquires the acyl chain from the final ACP of the assembly line through transesterification. Various nucleophiles can attack C1 of the acyl-TE intermediate. In 60% of the embedded TEs discussed here, the nucleophile is a hydroxyl group; in 25% it is water, and in 15% it is an amine. A nucleophilic competition seems to play out on each acyl-TE.²³³ If an amine is present, it usually wins. In the competition between secondary hydroxyl groups, the more distant or sterically accessible hydroxyl group usually wins. If hydroxyl groups are slow to approach C1, then water can win.

Several different types of TE can be embedded in the most C-terminal position of an assembly line.²³³ Of the 68 embedded TEs in the assembly lines analyzed here, 76% can be classified as PKS TEI, 6% as PKS TEII, 12% as NRPS TEI, and 4% as NRPS TEII. The curacin TE does not fit into any of these categories (see section 4.1.1.5). Polyketide synthase TEIs are represented by the C-terminal TE of the erythromycin synthase, whose crystal structure (PDB 1KEZ)²³⁴ was the first of an enzyme embedded in a cis-AT assembly line. Polyketide synthase TEIIs are most commonly trans enzymes, performing the editing role of releasing nonproductive acyl intermediates from assembly-line ACPs. However, in the ajudazol, gulfmirecin, jerangolide, and pellasoren assembly lines, embedded PKS TEIIs operate as offloading enzymes. The structure of an embedded PKS TEII is likely very similar to that of RifR, the nonembedded, editing TE of the rifamycin pathway (PDB 3FLA).²³⁵ The TEs that commonly collaborate with NRPS modules can also be divided into NRPS TEIs and NRPS TEIIs. A sequence alignment of TEs analyzed here indicates that cis-AT modules can collaborate with both NRPS TEIs (as in the althiomycin, antimycin, cryptophycin, gephyronic acid, hectochlorin, jamaicamide, microsclerodermin, and tubulysin assembly lines) and NRPS TEIIs (as in the ambruticin, epothilone, and

phenylannolone assembly lines). NRPS TEIs are represented by the TE embedded in the surfactin assembly line (PDB 1JMK),²³⁶ and NRPS TEIs are represented by the editing TE of the same assembly line (PDB 2RON).²³⁷ Only PKS TEI possesses the N-terminal, 30-residue dimerization motif of the TE from the erythromycin assembly line, although the curacin TE is also dimeric.²³⁸

4.1.1.1. Lactonization.: A lactone ring is formed when a hydroxyl substituent of the acyl chain nucleophilically attacks the acyl-TE intermediate, as in the assembly lines that synthesize ajudazol, akaeolide, amphotericin, antimycin, apoptolidin, avermectin, bafilomycin, borrelidin, candicidin, chalcomycin, concanamycin, conglobatin, disciformycin, E-492, E-837, elaiophylin, epothilone, erythromycin, FD-891, filipin, fostriecin, gulmirecin, halstoctacosanolide, jerangolide, lankamycin, meilingmycin, mycinamycin, mycolactone, niddamycin, nystatin, oligomycin, pellasoren, phenylannolone, phoslactomycin, pikromycin, pimarin, pladienolide, soraphen, spinosyn, stambomycin, thuggacin, tiacumicin, and tylosin.

4.1.1.1.1. Large Rings.: When a lactone ring contains 12 or more atoms, it is referred to as a macrolactone. From the analyzed assembly lines, the largest macrolactone produced is stambomycin (51-membered),²¹⁸ while the next largest macrolactones are amphotericin,¹⁹ candicidin,³⁴ and nystatin¹⁰¹ (38-membered).

4.1.1.1.2. Small Rings.: From the analyzed assembly lines, the smallest sized ring generated by a TE is six-membered, as produced in the akaeolide, aureothin, E-492, E-837, fostriecin, jerangolid, neo-aureothin, pellasoren, phenylannolone, and phoslactomycin assembly lines. Formation of the phenylannolone pyrone is aided by an α/β -cis double bond formed by the $KR_A + DH$ in the final module (an α/β -trans double bond would have prevented the formation of a six-membered ring).¹⁰³ The embedded TEs of the aureothin, the E-492, E-837, and neo-aureothin assembly lines also generate pyrones (E-492 and E-837 pyrones are subsequently tailored into furanones; see section 4.3.2).^{54,97}

4.1.1.1.3. Macrodiolide Formation.: Two of the discussed cis-AT assembly lines selectively produce macrodiolides.^{43,57} In both the conglobatin and elaiophylin assembly lines, the final ACP passes an acyl chain to TE to form an acyl-TE, acquires another acyl chain that nucleophilically attacks C1 of the first chain with its C7-hydroxyl group, and then passes the acyl chain to TE such that the C7-hydroxyl group of the first chain can nucleophilically attack C1 of the second chain.²³⁹ This requires the first acyl-TE be stable to cyclization and hydrolysis such that it awaits the arrival of the second chain. The α/β -trans double bond of conglobatin and the α/β -trans- γ/δ -trans-diene of elaiophylin ensure eight-membered macrolides (unfavored even for a saturated chain) are not formed. The conglobatin and elaiophylin TEs do not differ substantially on a sequence level from other PKS TEIs.

4.1.1.2. Hydrolysis.: A carboxylic acid is formed when a water molecule nucleophilically attacks the acyl-TE intermediate (as in the althiomycin, ambruticin, animycin, curacin, ECO-02301, gephyronic acid, herboxidiene, jamaicamide, kendomycin, lasalocid, lorneic acid, ptericidin, reveromycin, spirangien, tautomycetin, tautomycin, and tubulysin assembly

lines), although in the curacin, kendomycin, tautomycetin, and tautomycin assembly lines this carboxylate is lost as CO₂.^{76,125,238,240} The piericidin assembly line is interesting because one might expect a β,δ -diketo intermediate attached to TE to cyclize into a pyrone; however, the e/ζ -trans double bond may sterically inhibit pyrone formation.¹⁰⁵ No sequence fingerprint distinguishes hydrolyzing TEs from macrocyclizing TEs.

4.1.1.3. C2 Activation in Conjunction with Cyclization or Hydrolysis.: In the final stages of assembly-line synthesis of ajudazol, avermectin, kendomycin, lasalocid, and meilingmycin, unexplained cyclizations occur. In ajudazol, avermectin, lasalocid, and meilingmycin, these are between C2 and C7, where C2 is nucleophilic and C7 electrophilic.^{15,27,79,86} In kendomycin cyclization, C2 attacks a quinone carbonyl at the other end of the acyl chain.⁷⁶ The only enzyme common to each of these cyclizations is TE; thus these cyclizations may occur in the TE active site prior to the cyclization or hydrolysis it normally mediates, potentially after the acyl chain has been transesterified onto TE. On a sequence level, these TEs do not diverge from related TEs.

4.1.1.4. Lactamization.: A lactam is formed when an amino substituent on the acyl chain nucleophilically attacks the acyl-TE intermediate (as in the hitachimycin, incednine, lobosamide, microsclerodermin, salinilactam, and vicienistatin assembly lines, as well as in the second part of the BE-14016, cremimycin, and ML-449 assembly lines). The 30-residue N-terminal dimerization motif of these lactam-forming PKS TEIs is distinguishable from other PKS TEIs on a sequence level, but this may merely indicate an evolutionary relationship.

4.1.1.5. Terminal Alkene Formation.: The two most C-terminal enzymes embedded in the curacin assembly line are sulfotransferase (ST) and TE.⁴⁷ Sulfotransferase sulfonates the β -hydroxyl substituent generated by the final module of the curacin assembly line, thereby creating a leaving group for the decarboxylative elimination that follows thioester hydrolysis.²⁴¹ Thought to both hydrolyze and decarboxylate the intermediate to generate a terminal alkene product, the embedded curacin TE diverges significantly in both sequence and structure from other embedded TEs discussed here.²³⁸

4.1.2. Reductase.—An ~400-residue, NADPH-dependent thioester reductase (R) that belongs to the SDR family is located at the C-terminal end of the coelimycin, cyclizidine, and streptazone assembly lines, where it liberates thioester substrates from the assembly line as aldehydes (interestingly, in each of these pathways, the aldehyde products are subsequently converted to primary amines and cyclized to connect the newly installed nitrogen to C5).^{41,48,123} The R domain can also operate C-terminal to a final NRPS module, as in the myxalamid assembly line.²⁴² The myxalamid R domain continues past the aldehyde level to generate a primary alcohol (a side product of the cyclizidine assembly line is also produced when its R domain reduces twice).⁴⁸ A structure of the NADPH-bound myxalamid R domain is available (PDB 4U7W).²⁴³ No significant sequence differences are apparent between R domains that collaborate with cis-AT modules versus those that collaborate with NRPS modules.

A reductase-incompetent R domain (R*) in the chlorizidine assembly line conducts a cyclization reaction to yield a heterocyclic tricycle.³⁷ Other R* domains are known, such as in the cyclopiazonic acid synthase, where it catalyzes a Dieckmann cyclization.²⁴⁴ R* domains diverge significantly on a sequence level from reductase-competent R domains.

4.2. Nonembedded Enzymes

Nonembedded offloading enzymes can specifically interact with the final ACP of assembly line to release acyl chains. In the discussed pathways, tetronates, lactams, macrolides, carboxylic acids, and esters are generated.

4.2.1. Tetronate Condensing Enzyme.—The tetronate moiety of several tetronate-containing natural products (those produced by the abyssomicin, chlorothricin, kijanimicin, lobophorin, maklamicin, quartromicin, tetrocarcin, tetronasin, tetronomycin, and versipelostatin assembly lines) relies on a type III KS, known as a tetronate condensing enzyme (CE; respectively AbyA1, ChlM, KijB, LobC4, MakB1, QmnD5, TcaD4, Tsn13, Tmn15, and VstC5). This enzyme is hypothesized to acquire the β -ketoacyl chain bound to the final ACP of the assembly line through transthioesterification to its catalytic cysteine.⁸⁵ The condensing enzyme is then proposed to activate the α -hydroxyl group of a glyceryl-ACP (AbyA3, ChlD2, KijD, LobC2, MakB3, QmnD2, TcaD2, Tmn7a, or VstC2) or potentially glycolyl-ACP in the tetronasin pathway¹²⁸ for attack at C1 of the acyl chain and even catalyze a Dieckmann cyclization to generate the tetronic acid (although this cyclization may not need to be catalyzed). The structure of the abyssomicin CE, AbyA1, has been determined (PDB 5BY7).

Similar enzymology by NRPS machinery yields tetramates (as in the pyrroindomycin, streptolydigin, and tirandamycin assembly lines).^{124,132,206} Briefly, a C domain adds an α -aminoacyl-bound ACP through its α -amino group to a β -ketoacyl intermediate, and a spontaneous Dieckmann cyclization to C1 releases the tetramate.

4.2.2. Amide Synthase.—An ~260-residue amide synthase (AS) is usually encoded in each of the discussed benzenic and naphthalenic ansamycin gene clusters immediately downstream of the gene encoding the final module of the assembly line (as in the ansamitocin, chaxamycin, divergolide, geldanamycin, hygrocine, macbecin, naphthomycin, and rifamycin pathways; the rubraridin AS is encoded 13 kb upstream, whereas a paralogue may substitute for the AS expected for the herbimycin pathway⁶⁴). Amide synthases possess the same fold as arylamine N-acetyltransferases (e.g., PDB 2PQT), and each (Asm9, CxmF, DivN, GdmF, HgcF, MbcF, NatF, RifF, or RubF) possesses analogous catalytic Cys, His, and Asp residues.²⁴⁵ The arylamine N-acetyltransferase mechanism involves transfer of an acetyl group from acetyl-CoA onto a catalytic cysteine and subsequent attack of the thioester by the amino group of an arylamine substrate. By analogy, the catalytic cysteine of ASs may be trans-thioesterified by the acyl chain bound to the final ACP of the assembly line prior to an attack of the thioester by the aryl amino group present at the end of the acyl chain. No significant sequence difference is apparent between ASs of benzenic and naphthalenic ansamycin assembly lines. Amide synthases may have been naturally selected by

ansamycin-producing assembly lines over lactam-forming TEs due to favorable interactions with the aromatic group to which the amino group is attached.

4.2.3. Type II Thioesterase.—The monensin, nanchangmycin, nigericin, and salinomycin assembly lines are thought to transfer the grown acyl chain to a final, nonembedded ACP.¹⁸³ Epoxidases, epoxide cyclases, P450 hydroxylases, OMTs, glycosyltransferases, and pyran-forming enzymes may operate on acyl-ACPs until release by TEII (MonCII, NanE, NigCII, or SalBI). Type II thioesterases may recognize the chains as fully processed prior to hydrolyzing them from the ACP. In contrast, the polyether lasalocid is thought to be hydrolyzed from the final ACP of the lasalocid assembly line by the embedded TE prior to operation of the epoxidase and epoxide cyclase.⁷⁹

The pikromycin TEII, PikAV, is thought to acquire and cyclize the acyl chains bound to the final two modules of pikromycin synthase.²⁴⁶ The anatoxin TEII, AnaA, is suspected to cleave the grown acyl chain from the final module of the anatoxin assembly line, with the resulting β -keto acid spontaneously decarboxylating.²⁰

4.2.4. Type III Thioesterase.—The first portion of the BE-14106, cremimycin, and ML-449 assembly lines generates an α,β -unsaturated acyl chain that undergoes both Michael addition of a glycine and hydrolysis through the action of a TEIII (BecU, CmiSI, or MlaU; stereoselectivities unknown), such that the resulting β -amino acid can be prepared for loading onto the second part of these assembly lines (see section 2.4.6).^{30,44,91}

4.3. Mysterious Chain Release

The mechanisms of several offloading strategies are still being elucidated.

4.3.1. Mycolactone Chain-Joining Enzymes.—Two assembly lines are at work in the mycolactone pathway, one producing a macrolactone and the other producing a side chain that is appended to a hydroxyl group of the macrolactone.⁹⁴ The embedded TEs of both assembly lines are nearly identical to one another but very different from other embedded TEs; catalytic serines appear in the sequence AHSIV instead of the expected G(H/Y)SxG motif, and the C-terminal half does not align well with other embedded TEs. The potential joinase Mup045, encoded in the mycolactone gene cluster, is a type III ketosynthase similar to other C–O bond-forming type III ketosynthases (e.g., NonJ and NonK),²⁴⁷ albeit with a catalytic serine rather than a cysteine.⁹⁴ The mycolactone side chain could be transferred to Mup045 from the final ACP (or after being released by TE), and Mup045 could catalyze a nucleophilic attack by the macrolactone on C1 of the side chain.

4.3.2. Furanones from Pyrones.—The route by which the β,δ -diketoacyl chains synthesized by the aurafuron, E-492, and E-837 assembly lines are converted to furanone rings is not yet clear. E-492 and E-837 are hypothesized to be generated via pyrones offloaded by their embedded TEs.⁵⁴ The aurafuron assembly line does not possess an embedded TE so it was hypothesized that enzymes operate on the acyl chain bound to the final ACP, releasing the polyketide through a chain-shortening reaction similar to those performed by the type I Baeyer–Villiger monooxygenases of the pederin and FR901464 pathways (PedG and Fr9H-Ox).^{25,248,249} However, a comparison of tailoring enzymes in

these pathways suggests the furanone ring of aurafuron is most likely also formed through a pyrone intermediate. The E-492, E-837, and aurafuron Baeyer–Villiger monooxygenases (E-492 Orf13, E-837 Orf8, and AufJ) are highly similar on a sequence level and are type O, like the structurally characterized mithramycin Baeyer–Villiger monooxygenase MtmOIV (PDB 4K5R),²⁵⁰ which inserts an oxygen atom into a six-membered ring adjacent to a carbonyl. In the tailoring of pyrone intermediates of E-492 and E-837, the Baeyer–Villiger reaction is thought to be preceded by pyrone C2-hydroxylation by P450 hydroxylases (E-492 Orf 8 or 15, E-837 Orf 1 or 2). The aurafuron pathway possesses similar P450 hydroxylases (AufA, AufB, and AufH).⁵⁴ Each of these P450 hydroxylases is related to others that hydroxylate six-membered rings (e.g., SgcD3 in the C-1027 pathway).²⁵¹ Thus, an unidentified offloading enzyme may cyclize the β,δ -diketoacyl chain generated by the aurafuron assembly line into a pyrone prior to its conversion to a furanone.

4.3.3. Indanomycin Cyc, Pyran Formation, and Release.—The terminal Cyc domain of the indanomycin assembly line is homologous to SalBIII, the nonembedded pyran-forming enzyme in the salinomycin pathway, and may function in the same way to convert a β,ζ -dihydroxy moiety of an acyl chain bound to the terminal ACP into a pyran^{74,211}—(see section 2.4.1). In the salinomycin and related polyether pathways, the terminal $\text{KS}^0 + \text{AT}^0$ is thought to transfer the polyketide to a separately encoded ACP. While the terminal $\text{KS}^0 + \text{AT}^0$ of the indanomycin assembly line could be performing a similar transfer to an unidentified ACP (in both the salinomycin and indanomycin assembly lines, it is unclear which gene encodes the separate ACP), this would likely occur after online, Cyc-mediated pyran formation. Thus, the Cyc domain would be operating on an acyl chain bound to the previous module (indanomycin module 10) of the assembly line. A TEII is thought to release a precursor of salinomycin from ACP (see section 4.2.3). A TEII could similarly be involved in offloading from the indanomycin assembly line, either from the ACP of indanomycin module 10 or from an unidentified, separately encoded ACP.

4.3.4. Stigmatellin Cy and Chromone Formation.—The terminal Cy domain in the stigmatellin assembly line is thought to generate a chromone ring from a $\beta,\delta,\zeta,\theta$ -tetraketoacyl chain.¹²² Isotope labeling studies reveal that after the first ring is formed between C1 and C6, the next cyclization can occur either between the resulting C1- or C5-hydroxyl group and C9. FFAS suggests the Cy domain is similar to eight-stranded β -barrel proteins (e.g., PDB 4GZV); it is also similar to the putative trans-cyclase of the anatoxin pathway (see section 2.5).²⁰

4.3.5. Ebelactone Four-Membered Ring.—The ebelactone assembly line possesses a KS^0 at its C-terminus.⁵⁵ Whether this KS^0 is involved in catalyzing the cyclization of a β -hydroxyacyl chain to yield the four-membered β -lactone of ebelactone is unknown. It diverges from KSs of cis-AT modules both at a sequence level (residues 45–60 and 120–150) and by not being accompanied by an AT.

4.3.6. Calcimycin Intramolecular Amide Formation.—The nonembedded TEII of the calcimycin pathway, CalG, has been hypothesized to acquire the grown acyl chain from the last ACP of the assembly line and activate the amino group of 3-hydroxyanthranilic acid

for attack on C1 of the acyl chain.³³ CalG appears similar to other TEIIs at a sequence level, so it would be surprising if it activated a small molecule for an attack of the acyl-enzyme intermediate. Another possibility is that CalG mediates hydrolysis and that an undiscovered ligase joins the amino group of 3-hydroxyanthranilic acid, similar to the amide synthetase of the annimycin pathway, Ann1, that joins the amino group of 2-amino-3-hydroxycyclopent-2-enone to the TE-hydrolyzed acyl chain from the annimycin assembly line.²¹

4.3.7. Pyoluteorin Dieckmann Cyclization.—The nonembedded pyoluteorin TE, PltG, is a PKS TEII thought to catalyze a Dieckmann cyclization of C6 (between the δ - and ζ -keto groups) and C1 to generate an aromatic ring in the offloaded product.¹⁰⁹ PltG does not diverge significantly in sequence from other PKS TEIIs.

4.3.8. Cyindrospermopsin Uracil Ring Formation.—In the offloading of the tricyclic species generated by the cyindrospermopsin assembly line, two enzymes from the amidohydrolase/urease/dihydroorotase family (CyrG and CyrH) are thought to add a guanidino group to the β -keto thioester moiety of the intermediate bound to ACP of cyindrospermopsin module 5.⁴⁹ This helps generate a uracil ring that is important for the bioactivity of cyindrospermopsin. The guanidino group could be supplied by the donor molecule L-arginine. Thus, another enzymatic activity may be required to cleave the guanidino group from the donor molecule; a mysterious ~280-residue domain homologous to carbon–nitrogen hydrolases is present following cyindrospermopsin module 5 at the C-terminus of the assembly line.

5. CONCLUSION

This analysis of 135 diverse, well-characterized cis-AT assembly lines, comprised of 1129 PKS modules, has enabled more complete descriptions of the unusual structural and functional features within them. In the process, relatively obscure phenomena have been placed in the spotlight so that uncommon structures, such as embedded MTs, split DHs, and dimeric ERs, as well as uncommon activities, such as double-bond shifting, dehydration of δ -hydroxyl groups, and C2-activation during TE-mediated catalysis, can be treated as established phenomena that merit further study.

To realize the collective goal of engineering cis-AT assembly lines to produce new molecules and medicines, we should study nature's successes in engineering the same. While the typical α -, β -, γ -, and δ -type modules must possess sophisticated structural and functional properties to be naturally employed as often as they are, other domain arrangements are possible. The precious few examples of these are disproportionately informative. Tubulysin cis-AT modules show how KR domains can be positioned N-terminal to DH domains, FD-891 module 8 reveals that ER can insert into the DH domain, and the first modules of the curacin and jamaicamide assembly lines demonstrate how HALs can be inserted within cis-AT modules.

Our understanding of cis-AT modules and our ability to engineer them is being accelerated through communal databases. MIBiG has not only facilitated navigation between assembly-line sequence information and the literature but also enabled researchers to more quickly

share and compare what they have learned. The continued development of online resources will further thrust assembly-line science forward. Solved domain structures could be incorporated into domain prediction. Biosynthetic models containing hypothesized intermediates could be presented. Phenomenological spreadsheets that list domains or modules displaying certain architectural or functional characteristics could also be developed. Online resources would benefit from insightful posts from trusted researchers in the biosynthetic community.

An understanding of the uncommon phenomena within a system lies on the path to true understanding of that system. However, organizing and discussing such phenomena is only a starting point. Experiments must be conducted to figure out the significance of uncommon architectures as well as the mechanisms of uncommon chemistry. Why are so many seemingly inactive domains (KS^0 , AT^0 , DH^0 , ER^0 , and KR^0 , often possessing a complete set of catalytic residues) evolutionary retained? Is there a higher-order architecture to cis-AT assembly lines, as with the bacillaene trans-AT assembly line?¹⁸⁴ How is gatekeeping accomplished to ensure only dehydrated, cyclized, or twice-extended intermediates are passed down the assembly line? What is the code that signals a set of trans enzymes to operate on a particular acyl-ACP? These exciting questions, raised through a close look at the uncommon enzymology of cis-AT modules, will yield even more exciting answers in the years to come.

ACKNOWLEDGMENTS

A.T.K.-C. is supported by grants from NIH (GM106112) and the Welch Foundation (F-1712).

Biography

Adrian T. Keatinge-Clay is an associate professor in the Department of Molecular Biosciences at the University of Texas at Austin. He obtained his B.S. in chemistry and biology from Stanford University in 1999. In collaboration with Chaitan Khosla, he structurally characterized type II polyketide synthase enzymes as a graduate student in the crystallography lab of Robert Stroud at the University of California San Francisco. During a postdoctoral fellowship in the same lab, he collaborated with Kosan Biosciences to determine the structures and activities of several isolated type I polyketide synthase enzymes. He continues this work today.

ABBREVIATIONS

A	adenylation domain (equivalent to nonribosomal peptide synthetase A domain)
A_L	loading adenylation domain
ACP	acyl carrier protein
ACP_L	loading acyl carrier protein
AHBA	3-amino-5-hydroxybenzoic acid

am	aminomalonyl-specific
AMP	adenosine monophosphate
AMT	aminotransferase
AR	adapter domain (e.g., in curacin loading machinery)
AS	amide synthase
AT	acyltransferase
AT⁰	homologous to acyltransferase but different function
AT_L	loading acyltransferase
ATP	adenosine triphosphate
C	condensation domain (nonribosomal peptide synthetase)
CDD	C-terminal docking domain
CE	tetronate condensing enzyme
CMT	carbon methyltransferase
CMT_L	loading carbon methyltransferase
CoA	coenzyme A
Cy	cyclase (stigmatellin
Cyc	cyclase (indanomycin)
DH	dehydratase
DH⁰	homologous to dehydratase but different function
ECH1	enoyl-CoA hydratase-like enzyme (dehydrating)
ECH2	enoyl-CoA hydratase-like enzyme (decarboxylating)
EI	enoylisomerase
ER	enoylreductase
ER⁰	homologous to enoylreductase but different function
ER_L	loading enoylreductase
FFAS	Fold and Function Assignment System
GNAT	GCN5 <i>N</i> -acetyltransferase-like enzyme
GNAT_L	loading GCN5 <i>N</i> -acetyltransferase-like enzyme
HAL	halogenase

HCS	hydroxymethylglutaryl-CoA synthetase-like enzyme
hm	hydroxymalonyl-specific
KR	ketoreductase
KR⁰	homologous to ketoreductase but different function
KR_A	A-type ketoreductase
KR_B	B-type ketoreductase
KR_S	ketoreductase structural subdomain
KR_c	ketoreductase catalytic subdomain
KS	ketosynthase
KS⁰	homologous to ketosynthase but different function
KS_L	loading ketosynthase (decarboxylating)
m	malonyl-specific
MDR	medium-chain dehydrogenase/reductase
MIBiG	Minimum Information about a Biosynthetic Gene Cluster
mm	methylmalonyl-specific
MT	methyltransferase
NADPH	nicotinamide adenine dinucleotide phosphate
NDD	N-terminal docking domain
NRPS	nonribosomal peptide synthetase
OMT	oxygen methyltransferase
OX	α -oxidase
PDB	Protein Data Bank
PKS	polyketide synthase
PLP	pyridoxal phosphate
PS	pyran synthase
R	reductase domain (offloading enzyme)
R*	reductase-incompetent R domain (can mediate cyclization)
SAM	<i>S</i> -adenosylmethionine
SDR	short-chain dehydrogenase/reductase

ST	sulfotransferase
TE	thioesterase
TEI	type I thioesterase
TEII	type II thioesterase
TEIII	type III thioesterase

REFERENCES

- (1). Smith S; Tsai SC The Type I Fatty Acid and Polyketide Synthases: A Tale of Two Megasyntases. *Nat. Prod. Rep.* 2007, 24, 1041–1072. [PubMed: 17898897]
- (2). Khosla C; Tang Y; Chen AY; Schnarr NA; Cane DE Structure and Mechanism of the 6-Deoxyerythronolide B Synthase. *Annu. Rev. Biochem* 2007, 76, 195–221. [PubMed: 17328673]
- (3). Hertweck C The Biosynthetic Logic of Polyketide Diversity. *Angew. Chem., Int. Ed* 2009, 48, 4688–4716.
- (4). Keatinge-Clay AT The Structures of Type I Polyketide Synthases. *Nat. Prod. Rep* 2012, 29, 1050–1073. [PubMed: 22858605]
- (5). Piel J Biosynthesis of Polyketides by *trans*-AT Polyketide Synthases. *Nat. Prod. Rep* 2010, 27, 996–1047. [PubMed: 20464003]
- (6). Helfrich EJ; Piel J Biosynthesis of Polyketides by *trans*-AT Polyketide Synthases. *Nat. Prod. Rep* 2016, 33, 231–316. [PubMed: 26689670]
- (7). Sundaram S; Hertweck C On-Line Enzymatic Tailoring of Polyketides and Peptides in Thiotemplate Systems. *Curr. Opin. Chem. Biol* 2016, 31, 82–94. [PubMed: 26859622]
- (8). Hill AM The Biosynthesis, Molecular Genetics and Enzymology of the Polyketide-Derived Metabolites. *Nat. Prod. Rep* 2006, 23, 256–320. [PubMed: 16572230]
- (9). Hur GH; Vickery CR; Burkart MD Explorations of Catalytic Domains in Non-Ribosomal Peptide Synthetase Enzymology. *Nat. Prod. Rep* 2012, 29, 1074–1098. [PubMed: 22802156]
- (10). Conductor HL; Bruner SD Structure and Noncanonical Chemistry of Nonribosomal Peptide Biosynthetic Machinery. *Nat. Prod. Rep.* 2012, 29, 1099–1110. [PubMed: 22729219]
- (11). Medema MH; Kottmann R; Yilmaz P; Cummings M; Biggins JB; Blin K; de Bruijn I; Chooi YH; Claesen J; Coates RC; Cruz-Morales P; Duddela S; Dusterhus S; Edwards DJ; Fewer DP; Garg N; Geiger C; Gomez-Escribano JP; Greule A; Hadjithomas M; Haines AS; Helfrich EJ; Hillwig ML; Ishida K; Jones AC; Jones CS; Jungmann K; Kegler C; Kim HU; Kotter P; Krug D; Masschelein J; Melnik AV; Mantovani SM; Monroe EA; Moore M; Moss N; Nuttmann HW; Pan G; Pati A; Petras D; Reen FJ; Rosconi F; Rui Z; Tian Z; Tobias NJ; Tsunematsu Y; Wiemann P; Wyckoff E; Yan X; Yim G; Yu F; Xie Y; Aigle B; Apel AK; Balibar CJ; Balskus EP; Barona-Gomez F; Bechthold A; Bode HB; Borriss R; Brady SF; Brakhage AA; Caffrey P; Cheng YQ; Clardy J; Cox RJ; De Mot R; Donadio S; Donia MS; van der Donk WA; Dorrestein PC; Doyle S; Driessen AJ; Ehling-Schulz M; Entian KD; Fischbach MA; Gerwick L; Gerwick WH; Gross H; Gust B; Hertweck C; Hofte M; Jensen SE; Ju J; Katz L; Kayser L; Klassen JL; Keller NP; Kormanec J; Kuipers OP; Kuzuyama T; Kyrpides NC; Kwon HJ; Lautru S; Lavigne R; Lee CY; Linquan B; Liu X; Liu W; Luzhetskyy A; Mahmud T; Mast Y; Mendez C; Metsa-Ketela M; Micklefield J; Mitchell DA; Moore BS; Moreira LM; Muller R; Neilan BA; Nett M; Nielsen J; O’Gara F; Oikawa H; Osbourn A; Osburne MS; Ostash B; Payne SM; Pernodet JL; Petricek M; Piel J; Ploux O; Raaijmakers JM; Salas JA; Schmitt EK; Scott B; Seipke RF; Shen B; Sherman DH; Sivonen K; Smanski MJ; Sosio M; Stegmann E; Sussmuth RD; Tahlan K; Thomas CM; Tang Y; Truman AW; Viaud M; Walton JD; Walsh CT; Weber T; van Wezel GP; Wilkinson B; Willey JM; Wohlleben W; Wright GD; Ziemert N; Zhang C; Zotchev SB; Breitling R; Takano E; Glockner FO Minimum Information About a Biosynthetic Gene Cluster. *Nat. Chem. Biol* 2015, 11, 625–631. [PubMed: 26284661]
- (12). Weber T; Blin K; Duddela S; Krug D; Kim HU; Brucoleri R; Lee SY; Fischbach MA; Muller R; Wohlleben W; Breitling R; Takano E; Medema MH Antismash 3.0—A Comprehensive Resource

for the Genome Mining of Biosynthetic Gene Clusters. *Nucleic Acids Res.* 2015, 43, W237–243. [PubMed: 25948579]

- (13). Staunton J; Wilkinson B Biosynthesis of Erythromycin and Rapamycin. *Chem. Rev* 1997, 97, 2611–2630. [PubMed: 11851474]
- (14). Gottardi EM; Krawczyk JM; von Suchodoletz H; Schadt S; Muhlenweg A; Uguru GC; Pelzer S; Fiedler HP; Bibb MJ; Stach JE; Sussmuth RD Abyssomicin Biosynthesis: Formation of an Unusual Polyketide, Antibiotic-Feeding Studies and Genetic Analysis. *ChemBioChem* 2011, 12, 1401–1410. [PubMed: 21656887]
- (15). Buntin K; Weissman KJ; Muller R An Unusual Thioesterase Promotes Isochromanone Ring Formation in Ajudazol Biosynthesis. *ChemBioChem* 2010, 11, 1137–1146. [PubMed: 20432424]
- (16). Zhou T; Komaki H; Ichikawa N; Hosoyama A; Sato S; Igarashi Y Biosynthesis of Akaeolide and Lorneic Acids and Annotation of Type I Polyketide Synthase Gene Clusters in the Genome of *Streptomyces* sp. NPS554. *Mar. Drugs* 2015, 13, 581–596. [PubMed: 25603349]
- (17). Cortina NS; Revermann O; Krug D; Muller R Identification and Characterization of the Althiomycin Biosynthetic Gene Cluster in *Myxococcus xanthus* DK897. *ChemBioChem* 2011, 12, 1411–1416. [PubMed: 21626639]
- (18). Julien B; Tian ZQ; Reid R; Reeves CD Analysis of the Ambruticin and Jerangolid Gene Clusters of *Sorangium cellulosum* Reveals Unusual Mechanisms of Polyketide Biosynthesis. *Chem. Biol.* 2006, 13, 1277–1286. [PubMed: 17185223]
- (19). Caffrey P; Lynch S; Flood E; Finnan S; Oliynyk M Amphotericin Biosynthesis in *Streptomyces nodosus*: Deductions from Analysis of Polyketide Synthase and Late Genes. *Chem. Biol.* 2001, 8, 713–723. [PubMed: 11451671]
- (20). Mejean A; Paci G; Gautier V; Ploux O Biosynthesis of Anatoxin-a and Analogues (Anatoxins) in Cyanobacteria. *Toxicon* 2014, 91, 15–22. [PubMed: 25108149]
- (21). Kalan L; Gessner A; Thaker MN; Waglechner N; Zhu X; Szawiola A; Bechthold A; Wright GD; Zechel DL A Cryptic Polyene Biosynthetic Gene Cluster in *Streptomyces calvus* Is Expressed Upon Complementation with a Functional *bldA* Gene. *Chem. Biol.* 2013, 20, 1214–1224. [PubMed: 24120331]
- (22). Wu Y; Kang Q; Shen Y; Su W; Bai L Cloning and Functional Analysis of the Naphthomycin Biosynthetic Gene Cluster in *Streptomyces* sp. CS. *Mol. BioSyst.* 2011, 7, 2459–2469. [PubMed: 21614382]
- (23). Yan Y; Zhang L; Ito T; Qu X; Asakawa Y; Awakawa T; Abe I; Liu W Biosynthetic Pathway for High Structural Diversity of a Common Dilactone Core in Antimycin Production. *Org. Lett.* 2012, 14, 4142–4145. [PubMed: 22861048]
- (24). Du Y; Derewacz DK; Deguire SM; Teske J; Ravel J; Sulikowski GA; Bachmann BO Biosynthesis of the Apoptolidins in *Nocardioopsis* sp. FU 40. *Tetrahedron* 2011, 67, 6568–6575. [PubMed: 21869849]
- (25). Frank B; Wenzel SC; Bode HB; Scharfe M; Blocker H; Muller R From Genetic Diversity to Metabolic Unity: Studies on the Biosynthesis of Aurafurones and Aurafuron-Like Structures in Myxobacteria and Streptomycetes. *J. Mol. Biol.* 2007, 374, 24–38. [PubMed: 17919655]
- (26). He J; Hertweck C Iteration as Programmed Event During Polyketide Assembly; Molecular Analysis of the Aureothin Biosynthesis Gene Cluster. *Chem. Biol.* 2003, 10, 1225–1232. [PubMed: 14700630]
- (27). Ikeda H; Nonomiya T; Usami M; Ohta T; Omura S Organization of the Biosynthetic Gene Cluster for the Polyketide Anthelmintic Macrolide Avermectin in *Streptomyces avermitilis*. *Proc. Natl. Acad. Sci. U. S. A.* 1999, 96, 9509–9514. [PubMed: 10449723]
- (28). Zhang W; Fortman JL; Carlson JC; Yan J; Liu Y; Bai F; Guan W; Jia J; Maitainaho T; Sherman DH; Li S Characterization of the Bafilomycin Biosynthetic Gene Cluster from *Streptomyces lohii*. *ChemBioChem* 2013, 14, 301–306. [PubMed: 23362147]
- (29). Chang Z; Flatt P; Gerwick WH; Nguyen VA; Willis CL; Sherman DH The Barbamide Biosynthetic Gene Cluster: A Novel Marine Cyanobacterial System of Mixed Polyketide Synthase (PKS)–Non-Ribosomal Peptide Synthetase (NRPS) Origin Involving an Unusual Trichloroleucyl Starter Unit. *Gene* 2002, 296, 235–247. [PubMed: 12383521]

- (30). Jorgensen H; Degnes KF; Sletta H; Fjaervik E; Dikiy A; Herfindal L; Bruheim P; Klinkenberg G; Bredholt H; Nygard G; Doskeland SO; Ellingsen TE; Zotchev SB Biosynthesis of Macrolactam BE-14106 Involves Two Distinct PKS Systems and Amino Acid Processing Enzymes for Generation of the Aminoacyl Starter Unit. *Chem. Biol.* 2009, 16, 1109–1121. [PubMed: 19875084]
- (31). Wenzel SC; Hoffmann H; Zhang J; Debussche L; Haag-Richter S; Kurz M; Nardi F; Lukat P; Kochems I; Tietgen H; Schummer D; Nicolas JP; Calvet L; Czepczor V; Vrignaud P; Muhlenweg A; Pelzer S; Muller R; Bronstrup M Production of the Bengamide Class of Marine Natural Products in Myxobacteria: Biosynthesis and Structure-Activity Relationships. *Angew. Chem., Int. Ed.* 2015, 54, 15560–15564.
- (32). Olano C; Moss SJ; Brana AF; Sheridan RM; Math V; Weston AJ; Mendez C; Leadlay PF; Wilkinson B; Salas JA Biosynthesis of the Angiogenesis Inhibitor Borrelidin by *Streptomyces parvulus* TU4055: Insights into Nitrile Formation. *Mol. Microbiol.* 2004, 52, 1745–1756. [PubMed: 15186422]
- (33). Wu Q; Liang J; Lin S; Zhou X; Bai L; Deng Z; Wang Z Characterization of the Biosynthesis Gene Cluster for the Pyrrole Polyether Antibiotic Calcimycin (A23187) in *Streptomyces chartreusis* NRRL 3882. *Antimicrob. Agents Chemother.* 2011, 55, 974–982. [PubMed: 21173184]
- (34). Chen S; Huang X; Zhou X; Bai L; He J; Jeong KJ; Lee SY; Deng Z Organizational and Mutational Analysis of a Complete Fr-008/Candicidin Gene Cluster Encoding a Structurally Related Polyene Complex. *Chem. Biol.* 2003, 10, 1065–1076. [PubMed: 14652074]
- (35). Ward SL; Hu Z; Schirmer A; Reid R; Revill WP; Reeves CD; Petrakovsky OV; Dong SD; Katz L Chalcomycin Biosynthesis Gene Cluster from *Streptomyces bikiniensis*: Novel Features of an Unusual Ketolide Produced through Expression of the *chm* Polyketide Synthase in *Streptomyces fradiae*. *Antimicrob. Agents Chemother.* 2004, 48, 4703–4712. [PubMed: 15561847]
- (36). Castro JF; Razmilic V; Gomez-Escribano JP; Andrews B; Asenjo JA; Bibb MJ Identification and Heterologous Expression of the Chaxamycin Biosynthesis Gene Cluster from *Streptomyces leeuwenhoekii*. *Appl. Environ. Microbiol.* 2015, 81, 5820–5831. [PubMed: 26092459]
- (37). Mantovani SM; Moore BS Flavin-Linked Oxidase Catalyzes Pyrrolizine Formation of Dichloropyrrole-Containing Polyketide Extender Unit in Chlorizidine A. *J. Am. Chem. Soc.* 2013, 135, 18032–18035. [PubMed: 24246014]
- (38). Jia XY; Tian ZH; Shao L; Qu XD; Zhao QF; Tang J; Tang GL; Liu W Genetic Characterization of the Chlorothricin Gene Cluster as a Model for Spirotetronate Antibiotic Biosynthesis. *Chem. Biol.* 2006, 13, 575–585. [PubMed: 16793515]
- (39). Rachid S; Krug D; Kunze B; Kochems I; Scharfe M; Zabriskie TM; Blocker H; Muller R Molecular and Biochemical Studies of Chondramide Formation—Highly Cytotoxic Natural Products from *Chondromyces crocatus* Cm c5. *Chem. Biol.* 2006, 13, 667–681. [PubMed: 16793524]
- (40). Rachid S; Scharfe M; Blocker H; Weissman KJ; Muller R Unusual Chemistry in the Biosynthesis of the Antibiotic Chondrochlorens. *Chem. Biol.* 2009, 16, 70–81. [PubMed: 19171307]
- (41). Gomez-Escribano JP; Song L; Fox DJ; Yeo V; Bibb MJ; Challis GL Structure and Biosynthesis of the Unusual Polyketide Alkaloid Coelimycin P1, a Metabolic Product of the cpk Gene Cluster of *Streptomyces coelicolor* M145. *Chem. Sci* 2012, 3, 2716–2720.
- (42). Haydock SF; Appleyard AN; Mironenko T; Lester J; Scott N; Leadlay PF Organization of the Biosynthetic Gene Cluster for the Macrolide Concanamycin A in *Streptomyces neyagawaensis* ATCC 27449. *Microbiology* 2005, 151, 3161–3169. [PubMed: 16207901]
- (43). Zhou Y; Murphy AC; Samborskyy M; Prediger P; Dias LC; Leadlay PF Iterative Mechanism of Macrodilide Formation in the Anticancer Compound Conglobatin. *Chem. Biol.* 2015, 22, 745–754. [PubMed: 26091168]
- (44). Amagai K; Takaku R; Kudo F; Eguchi T A Unique Amino Transfer Mechanism for Constructing the Beta-Amino Fatty Acid Starter Unit in the Biosynthesis of the Macrolactam Antibiotic Cremimycin. *ChemBioChem* 2013, 14, 1998–2006. [PubMed: 24014395]
- (45). Muller S; Rachid S; Hoffmann T; Surup F; Volz C; Zaburanyi N; Muller R Biosynthesis of Crocacin Involves an Unusual Hydrolytic Release Domain Showing Similarity to Condensation Domains. *Chem. Biol.* 2014, 21, 855–865. [PubMed: 24981773]

- (46). Magarvey NA; Beck ZQ; Golakoti T; Ding Y; Huber U; Hemscheidt TK; Abelson D; Moore RE; Sherman DH Biosynthetic Characterization and Chemoenzymatic Assembly of the Cryptophycins. Potent Anticancer Agents from Cyanobionts. ACS Chem. Biol. 2006, 1, 766–779. [PubMed: 17240975]
- (47). Chang Z; Sitachitta N; Rossi JV; Roberts MA; Flatt PM; Jia J; Sherman DH; Gerwick WH Biosynthetic Pathway and Gene Cluster Analysis of Curacin A, an Antitubulin Natural Product from the Tropical Marine Cyanobacterium *Lyngbya majuscula*. J. Nat. Prod. 2004, 67, 1356–1367. [PubMed: 15332855]
- (48). Huang W; Kim SJ; Liu J; Zhang W Identification of the Polyketide Biosynthetic Machinery for the Indolizidine Alkaloid Cyclizidine. Org. Lett. 2015, 17, 5344–5347. [PubMed: 26473429]
- (49). Mihali TK; Kellmann R; Muenchhoff J; Barrow KD; Neilan BA Characterization of the Gene Cluster Responsible for Cylindrospermopsin Biosynthesis. Appl. Environ. Microbiol 2008, 74, 716–722. [PubMed: 18065631]
- (50). Feng Z; Qi J; Tsuge T; Oba Y; Kobayashi T; Suzuki Y; Sakagami Y; Ojika M Construction of a Bacterial Artificial Chromosome Library for a Myxobacterium of the Genus *Cystobacter* and Characterization of an Antibiotic Biosynthetic Gene Cluster. Biosci., Biotechnol., Biochem. 2005, 69, 1372–1380. [PubMed: 16041144]
- (51). Surup F; Viehrig K; Mohr KI; Herrmann J; Jansen R; Muller R Disciformycins A and B: 12-Membered Macrolide Glycoside Antibiotics from the Myxobacterium *Pyxidicoccus fallax* Active against Multiresistant Staphylococci. Angew. Chem., Int. Ed. 2014, 53, 13588–13591.
- (52). Xu Z; Baunach M; Ding L; Peng H; Franke J; Hertweck C Biosynthetic Code for Divergolide Assembly in a Bacterial Mangrove Endophyte. ChemBioChem 2014, 15, 1274–1279. [PubMed: 24867126]
- (53). Meiser P; Weissman KJ; Bode HB; Krug D; Dickschat JS; Sandmann A; Muller R DKxanthene Biosynthesis—Understanding the Basis for Diversity-Oriented Synthesis in Myxobacterial Secondary Metabolism. Chem. Biol. 2008, 15, 771–781. [PubMed: 18721748]
- (54). Banskota AH; McAlpine JB; Sorensen D; Aouidate M; Pirae M; Alarco AM; Omura S; Shiomi K; Farnet CM; Zazopoulos E Isolation and Identification of Three New 5-Alkenyl-3,3(2H)-Furanones from Two Streptomyces Species Using a Genomic Screening Approach. J. Antibiot. 2006, 59, 168–176. [PubMed: 16724457]
- (55). Wyatt MA; Ahilan Y; Argyropoulos P; Boddy CN; Magarvey NA; Harrison PH Biosynthesis of Ebelactone A: Isotopic Tracer, Advanced Precursor and Genetic Studies Reveal a Thioesterase-Independent Cyclization to Give a Polyketide Beta-Lactone. J. Antibiot. 2013, 66, 421–430. [PubMed: 23801186]
- (56). McAlpine JB; Bachmann BO; Pirae M; Tremblay S; Alarco AM; Zazopoulos E; Farnet CM Microbial Genomics as a Guide to Drug Discovery and Structural Elucidation: ECO-02301, a Novel Antifungal Agent, as an Example. J. Nat. Prod. 2005, 68, 493–496. [PubMed: 15844935]
- (57). Sheng Y; Lam PW; Shahab S; Santosa DA; Proteau PJ; Zabriskie TM; Mahmud T Identification of Elaiophylin Skeletal Variants from the Indonesian *Streptomyces* sp. ICBB 9297. J. Nat. Prod. 2015, 78, 2768–2775. [PubMed: 26510047]
- (58). Molnar I; Schupp T; Ono M; Zirkle R; Milnamow M; Nowak-Thompson B; Engel N; Toupet C; Stratmann A; Cyr DD; Grolach J; Mayo JM; Hu A; Goff S; Schmid J; Ligon JM The Biosynthetic Gene Cluster for the Microtubule-Stabilizing Agents Epothilones A and B from *Sorangium cellulosum* So ce90. Chem. Biol. 2000, 7, 97–109. [PubMed: 10662695]
- (59). Cortes J; Haydock SF; Roberts GA; Bevitt DJ; Leadlay PF An Unusually Large Multifunctional Polypeptide in the Erythromycin-Producing Polyketide Synthase of *Saccharopolyspora erythraea*. Nature 1990, 348, 176–178. [PubMed: 2234082]
- (60). Kudo F; Motegi A; Mizoue K; Eguchi T Cloning and Characterization of the Biosynthetic Gene Cluster of 16-Membered Macrolide Antibiotic FD-891: Involvement of a Dual Functional Cytochrome P450 Monooxygenase Catalyzing Epoxidation and Hydroxylation. ChemBioChem 2010, 11, 1574–1582. [PubMed: 20589823]
- (61). Ikeda H; Shin-ya K; Omura S Genome Mining of the *Streptomyces avermitilis* Genome and Development of Genome-Minimized Hosts for Heterologous Expression of Biosynthetic Gene Clusters. J. Ind. Microbiol. Biotechnol. 2014, 41, 233–250. [PubMed: 23990133]

- (62). Wu K; Chung L; Revill WP; Katz L; Reeves CD The FK520 Gene Cluster of *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) Contains Genes for Biosynthesis of Unusual Polyketide Extender Units. *Gene* 2000, 251, 81–90. [PubMed: 10863099]
- (63). Kong R; Liu X; Su C; Ma C; Qiu R; Tang L Elucidation of the Biosynthetic Gene Cluster and the Post-PKS Modification Mechanism for Fostriecin in *Streptomyces pulveraceus*. *Chem. Biol.* 2013, 20, 45–54. [PubMed: 23352138]
- (64). Rascher A; Hu Z; Buchanan GO; Reid R; Hutchinson CR Insights into the Biosynthesis of the Benzoquinone Ansamycins Geldanamycin and Herbimycin, Obtained by Gene Sequencing and Disruption. *Appl. Environ. Microbiol.* 2005, 71, 4862–4871. [PubMed: 16085885]
- (65). Young J; Stevens DC; Carmichael R; Tan J; Rachid S; Boddy CN; Muller R; Taylor RE Elucidation of Gephyronic Acid Biosynthetic Pathway Revealed Unexpected SAM-Dependent Methylations. *J. Nat. Prod.* 2013, 76, 2269–2276. [PubMed: 24298873]
- (66). Schieferdecker S; Konig S; Weigel C; Dahse HM; Werz O; Nett M Structure and Biosynthetic Assembly of Gulumirecins, Macrolide Antibiotics from the Predatory Bacterium *Pyxidicoccus fallax*. *Chem. - Eur. J* 2014, 20, 15933–15940. [PubMed: 25287056]
- (67). Tohyama S; Kakinuma K; Eguchi T The Complete Biosynthetic Gene Cluster of the 28-Membered Polyketide Macrolactones, Halstoctacosanolides, from *Streptomyces halstedii* HC34. *J. Antibiot.* 2006, 59, 44–52. [PubMed: 16568718]
- (68). Steinmetz H; Li J; Fu C; Ziburanyi N; Kunze B; Harmrolfs K; Schmitt V; Herrmann J; Reichenbach H; Hofle G; Kalesse M; Muller R Isolation, Structure Elucidation, and (Bio)Synthesis of Haprolid, a Cell-Type-Specific Myxobacterial Cytotoxin. *Angew. Chem., Int. Ed.* 2016, 55, 10113–10117.
- (69). Ramaswamy AV; Sorrels CM; Gerwick WH Cloning and Biochemical Characterization of the Hectochlorin Biosynthetic Gene Cluster from the Marine Cyanobacterium *Lyngbya majuscula*. *J. Nat. Prod.* 2007, 70, 1977–1986. [PubMed: 18001088]
- (70). Shao L; Zi J; Zeng J; Zhan J Identification of the Herboxidiene Biosynthetic Gene Cluster in *Streptomyces chromofuscus* ATCC 49982. *Appl. Environ. Microbiol.* 2012, 78, 2034–2038. [PubMed: 22247174]
- (71). Kudo F; Kawamura K; Uchino A; Miyanaga A; Numakura M; Takayanagi R; Eguchi T Genome Mining of the Hitachimycin Biosynthetic Gene Cluster: Involvement of a Phenylalanine-2,3-Aminomutase in Biosynthesis. *ChemBioChem* 2015, 16, 909–914. [PubMed: 25786909]
- (72). Li S; Wang H; Li Y; Deng J; Lu C; Shen Y; Shen Y Biosynthesis of Hygrocins, Antitumor Naphthoquinone Ansamycins Produced by *Streptomyces* sp. LZ35. *ChemBioChem* 2014, 15, 94–102. [PubMed: 24501776]
- (73). Takaishi M; Kudo F; Eguchi T Identification of the Incednine Biosynthetic Gene Cluster: Characterization of Novel β -Glutamate- β -Decarboxylase IdnL3. *J. Antibiot.* 2013, 66, 691–699. [PubMed: 23921821]
- (74). Li C; Roege KE; Kelly WL Analysis of the Indanomycin Biosynthetic Gene Cluster from *Streptomyces antibioticus* NRRL 8167. *ChemBioChem* 2009, 10, 1064–1072. [PubMed: 19301315]
- (75). Edwards DJ; Marquez BL; Nogle LM; McPhail K; Goeger DE; Roberts MA; Gerwick WH Structure and Biosynthesis of the Jamaicamides, New Mixed Polyketide-Peptide Neurotoxins from the Marine Cyanobacterium *Lyngbya majuscula*. *Chem. Biol.* 2004, 11, 817–833. [PubMed: 15217615]
- (76). Wenzel SC; Bode HB; Kochems I; Muller R A Type I/ Type III Polyketide Synthase Hybrid Biosynthetic Pathway for the Structurally Unique Ansa Compound Kendomycin. *ChemBioChem* 2008, 9, 2711–2721. [PubMed: 18972512]
- (77). Zhang H; White-Phillip JA; Melancon CE 3rd; Kwon HJ; Yu WL; Liu HW Elucidation of the Kijanimicin Gene Cluster: Insights into the Biosynthesis of Spirotetronate Antibiotics and Nitrosugars. *J. Am. Chem. Soc.* 2007, 129, 14670–14683. [PubMed: 17985890]
- (78). Arakawa K; Kodama K; Tatsuno S; Ide S; Kinashi H Analysis of the Loading and Hydroxylation Steps in Lankamycin Biosynthesis in *Streptomyces rochei*. *Antimicrob. Agents Chemother.* 2006, 50, 1946–1952. [PubMed: 16723550]

- (79). Migita A; Watanabe M; Hirose Y; Watanabe K; Tokiwano T; Kinashi H; Oikawa H Identification of a Gene Cluster of Polyether Antibiotic Lasalocid from *Streptomyces lasaliensis*. Biosci., Biotechnol., Biochem. 2009, 73, 169–176. [PubMed: 19129623]
- (80). Kopp M; Irschik H; Gemperlein K; Buntin K; Meiser P; Weissman KJ; Bode HB; Muller R Insights into the Complex Biosynthesis of the Leupyrrins in *Sorangium cellulosum* So ce690. Mol. Biosyst 2011, 7, 1549–1563. [PubMed: 21365089]
- (81). Bihlmaier C; Welle E; Hofmann C; Welzel K; Vente A; Breitling E; Muller M; Glaser S; Bechthold A Biosynthetic Gene Cluster for the Polyenoiltetramic Acid Alpha-Lipomycin. Antimicrob. Agents Chemother. 2006, 50, 2113–2121. [PubMed: 16723573]
- (82). Niu S; Li S; Chen Y; Tian X; Zhang H; Zhang G; Zhang W; Yang X; Zhang S; Ju J; Zhang C Lobophorins E and F, New Spirotetronate Antibiotics from a South China Sea-Derived *Streptomyces* sp. SCSIO 01127. J. Antibiot. 2011, 64, 711–716. [PubMed: 21897400]
- (83). Schulze CJ; Donia MS; Siqueira-Neto JL; Ray D; Raskatov JA; Green RE; McKerrow JH; Fischbach MA; Linington RG Genome-Directed Lead Discovery: Biosynthesis, Structure Elucidation, and Biological Evaluation of Two Families of Polyene Macrolactams against *Trypanosoma brucei*. ACS Chem. Biol. 2015, 10, 2373–2381. [PubMed: 26270237]
- (84). Martin CJ; Gaisser S; Challis IR; Carletti I; Wilkinson B; Gregory M; Prodromou C; Roe SM; Pearl LH; Boyd SM; Zhang MQ Molecular Characterization of Macbecin as an Hsp90 Inhibitor. J. Med. Chem. 2008, 51, 2853–2857. [PubMed: 18357975]
- (85). Daduang R; Kitani S; Hashimoto J; Thamchaipenet A; Igarashi Y; Shin-ya K; Ikeda H; Nihira T Characterization of the Biosynthetic Gene Cluster for Maklamycin, a Spirotetronate-Class Antibiotic of the Endophytic *Micromonospora* sp. NBRC 110955. Microbiol. Res. 2015, 180, 30–39. [PubMed: 26505309]
- (86). He Y; Sun Y; Liu T; Zhou X; Bai L; Deng Z Cloning of Separate Meilingmycin Biosynthesis Gene Clusters by Use of Acyltransferase-Ketoreductase Didomain PCR Amplification. Appl. Environ. Microbiol. 2010, 76, 3283–3292. [PubMed: 20348291]
- (87). Weinig S; Hecht HJ; Mahmud T; Muller R Melithiazol Biosynthesis: Further Insights into Myxobacterial PKS/NRPS Systems and Evidence for a New Subclass of Methyl Transferases. Chem. Biol. 2003, 10, 939–952. [PubMed: 14583260]
- (88). He M; Haltli B; Summers M; Feng X; Hucul J Isolation and Characterization of Meridamycin Biosynthetic Gene Cluster from *Streptomyces* sp. NRRL 30748. Gene 2006, 377, 109–118. [PubMed: 16806745]
- (89). Pearson L; Mihali T; Moffitt M; Kellmann R; Neilan B On the Chemistry, Toxicology and Genetics of the Cyanobacterial Toxins, Microcystin, Nodularin, Saxitoxin and Cylindrospermopsin. Mar. Drugs 2010, 8, 1650–1680. [PubMed: 20559491]
- (90). Hoffmann T; Muller S; Nadmid S; Garcia R; Muller R Microsclerodermins from Terrestrial Myxobacteria: An Intriguing Biosynthesis Likely Connected to a Sponge Symbiont. J. Am. Chem. Soc. 2013, 135, 16904–16911. [PubMed: 24124771]
- (91). Jorgensen H; Degnes KF; Dikiy A; Fjaervik E; Klinkenberg G; Zotchev SB Insights into the Evolution of Macrolactam Biosynthesis through Cloning and Comparative Analysis of the Biosynthetic Gene Cluster for a Novel Macrocyclic Lactam, ML-449. Appl. Environ. Microbiol. 2010, 76, 283–293. [PubMed: 19854930]
- (92). Bhatt A; Stark CB; Harvey BM; Gallimore AR; Demydchuk YA; Spencer JB; Staunton J; Leadlay PF Accumulation of an *E,E,E*-Triene by the Monensin-Producing Polyketide Synthase When Oxidative Cyclization Is Blocked. Angew. Chem., Int. Ed. 2005, 44, 7075–7078.
- (93). Anzai Y; Saito N; Tanaka M; Kinoshita K; Koyama Y; Kato F Organization of the Biosynthetic Gene Cluster for the Polyketide Macrolide Mycinamicin in *Micromonospora griseorubida*. FEMS Microbiol. Lett. 2003, 218, 135–141. [PubMed: 12583909]
- (94). Stinear TP; Mve-Obiang A; Small PL; Frigui W; Pryor MJ; Brosch R; Jenkin GA; Johnson PD; Davies JK; Lee RB; Adusumilli S; Garnier T; Haydock SF; Leadlay PF; Cole ST Giant Plasmid-Encoded Polyketide Synthases Produce the Macrolide Toxin of *Mycobacterium ulcerans*. Proc. Natl. Acad. Sci. U. S. A 2004, 101, 1345–1349. [PubMed: 14736915]

- (95). Sun Y; Zhou X; Dong H; Tu G; Wang M; Wang B; Deng Z A Complete Gene Cluster from *Streptomyces nanchangensis* NS3226 Encoding Biosynthesis of the Polyether Ionophore Nanchangmycin. *Chem. Biol.* 2003, 10, 431–441. [PubMed: 12770825]
- (96). Krastel P; Roggo S; Schirle M; Ross NT; Perruccio F; Aspesi P Jr.; Aust T; Buntin K; Estoppey D; Liechty B; Mapa F; Memmert K; Miller H; Pan X; Riedl R; Thibaut C; Thomas J; Wagner T; Weber E; Xie X; Schmitt EK; Hoepfner D Nannocystin A: An Elongation Factor 1 Inhibitor from Myxobacteria with Differential Anti-Cancer Properties. *Angew. Chem., Int. Ed.* 2015, 54, 10149–10154.
- (97). Traitcheva N; Jenke-Kodama H; He J; Dittmann E; Hertweck C Non-Colinear Polyketide Biosynthesis in the Aureothin and Neoareothin Pathways: An Evolutionary Perspective. *ChemBioChem* 2007, 8, 1841–1849.
- (98). Kakavas SJ; Katz L; Stassi D Identification and Characterization of the Niddamycin Polyketide Synthase Genes from *Streptomyces caelestis*. *J. Bacteriol* 1997, 179, 7515–7522. [PubMed: 9393718]
- (99). Harvey BM; Mironenko T; Sun Y; Hong H; Deng Z; Leadlay PF; Weissman KJ; Haydock SF Insights into Polyether Biosynthesis from Analysis of the Nigericin Biosynthetic Gene Cluster in *Streptomyces* sp. DSM4137. *Chem. Biol.* 2007, 14, 703–714. [PubMed: 17584617]
- (100). Bis DM; Ban YH; James ED; Alqahtani N; Viswanathan R; Lane AL Characterization of the Nocardiopsis Biosynthetic Gene Cluster Reveals Similarities to and Differences from the Rapamycin and FK-506 Pathways. *ChemBioChem* 2015, 16, 990–997. [PubMed: 25755076]
- (101). Brautaset T; Sekurova ON; Sletta H; Ellingsen TE; Strom AR; Valla S; Zotchev SB Biosynthesis of the Polyene Antifungal Antibiotic Nystatin in *Streptomyces noursei* ATCC 11455: Analysis of the Gene Cluster and Deduction of the Biosynthetic Pathway. *Chem. Biol.* 2000, 7, 395–403. [PubMed: 10873841]
- (102). Jahns C; Hoffmann T; Muller S; Gerth K; Washausen P; Hofle G; Reichenbach H; Kalesse M; Muller R Pellasoren: Structure Elucidation, Biosynthesis, and Total Synthesis of a Cytotoxic Secondary Metabolite from *Sorangium cellulosum*. *Angew. Chem., Int. Ed.* 2012, 51, 5239–5243.
- (103). Bouhired SM; Crusemann M; Almeida C; Weber T; Piel J; Schaberle TF; Konig GM Biosynthesis of Phenylannolone A, a Multidrug Resistance Reversal Agent from the Halotolerant Myxobacterium *Nannocystis pusilla* B150. *ChemBioChem* 2014, 15, 757–765. [PubMed: 24677362]
- (104). Palaniappan N; Kim BS; Sekiyama Y; Osada H; Reynolds KA Enhancement and Selective Production of Phoslactomycin B, a Protein Phosphatase IIa Inhibitor, through Identification and Engineering of the Corresponding Biosynthetic Gene Cluster. *J. Biol. Chem.* 2003, 278, 35552–35557. [PubMed: 12819191]
- (105). Liu Q; Yao F; Chooi YH; Kang Q; Xu W; Li Y; Shao Y; Shi Y; Deng Z; Tang Y; You D Elucidation of Piericidin A1 Biosynthetic Locus Revealed a Thioesterase-Dependent Mechanism of Alpha-Pyridone Ring Formation. *Chem. Biol.* 2012, 19, 243–253. [PubMed: 22365607]
- (106). Aparicio JF; Colina AJ; Ceballos E; Martin JF The Biosynthetic Gene Cluster for the 26-Membered Ring Polyene Macrolide Pimaricin. A New Polyketide Synthase Organization Encoded by Two Subclusters Separated by Functionalization Genes. *J. Biol. Chem.* 1999, 274, 10133–10139. [PubMed: 10187796]
- (107). Machida K; Arisawa A; Takeda S; Tsuchida T; Aritoku Y; Yoshida M; Ikeda H Organization of the Biosynthetic Gene Cluster for the Polyketide Antitumor Macrolide, Pladienolide, in *Streptomyces platensis* Mer-11107. *Biosci., Biotechnol., Biochem.* 2008, 72, 2946–2952. [PubMed: 18997414]
- (108). Mares J; Hajek J; Urajova P; Kopecky J; Hrouzek P A Hybrid Non-Ribosomal Peptide/Polyketide Synthetase Containing Fatty-Acyl Ligase (FAAL) Synthesizes the β -Amino Fatty Acid Lipopeptides Puwainaphycins in the Cyanobacterium *Cylindrospermum alatosporum*. *PLoS One* 2014, 9, e111904. [PubMed: 25369527]
- (109). Gross H; Loper JE Genomics of Secondary Metabolite Production by *Pseudomonas* spp. *Nat. Prod. Rep.* 2009, 26, 1408–1446. [PubMed: 19844639]

- (110). He HY; Pan HX; Wu LF; Zhang BB; Chai HB; Liu W; Tang GL Quartromycin Biosynthesis: Two Alternative Polyketide Chains Produced by One Polyketide Synthase Assembly Line. *Chem. Biol.* 2012, 19, 1313–1323. [PubMed: 23102224]
- (111). Schwecke T; Aparicio JF; Molnar I; Konig A; Khaw LE; Haydock SF; Oliynyk M; Caffrey P; Cortes J; Lester JB; et al. The Biosynthetic Gene Cluster for the Polyketide Immunosuppressant Rapamycin. *Proc. Natl. Acad. Sci. U. S. A.* 1995, 92, 7839–7843. [PubMed: 7644502]
- (112). Takahashi S; Toyoda A; Sekiyama Y; Takagi H; Nogawa T; Uramoto M; Suzuki R; Koshino H; Kumano T; Panthee S; Dairi T; Ishikawa J; Ikeda H; Sakaki Y; Osada H Reveromycin A Biosynthesis Uses RevG and RevJ for Stereospecific Spiroacetal Formation. *Nat. Chem. Biol.* 2011, 7, 461–468. [PubMed: 21642985]
- (113). Floss HG; Yu TW Rifamycin—Mode of Action, Resistance, and Biosynthesis. *Chem. Rev.* 2005, 105, 621–632. [PubMed: 15700959]
- (114). Kim CG; Lamichhane J; Song KI; Nguyen VD; Kim DH; Jeong TS; Kang SH; Kim KW; Maharjan J; Hong YS; Kang JS; Yoo JC; Lee JJ; Oh TJ; Liou K; Sohng JK Biosynthesis of Rubradirin as an Ansamycin Antibiotic from *Streptomyces achromogenes* var. *rubradiris* NRRI3061. *Arch. Microbiol.* 2008, 189, 463–473. [PubMed: 18080113]
- (115). Uduary DW; Zeigler L; Asolkar RN; Singan V; Lapidus A; Fenical W; Jensen PR; Moore BS Genome Sequencing Reveals Complex Secondary Metabolome in the Marine Actinomycete *Salinispora tropica*. *Proc. Natl. Acad. Sci. U. S. A.* 2007, 104, 10376–10381. [PubMed: 17563368]
- (116). Jiang C; Wang H; Kang Q; Liu J; Bai L Cloning and Characterization of the Polyether Salinomycin Biosynthesis Gene Cluster of *Streptomyces albus* XM211. *Appl. Environ. Microbiol.* 2012, 78, 994–1003. [PubMed: 22156425]
- (117). Pulsawat N; Kitani S; Nihira T Characterization of Biosynthetic Gene Cluster for the Production of Virginiamycin M, a Streptogramin Type A Antibiotic, in *Streptomyces virginiae*. *Gene* 2007, 393, 31–42. [PubMed: 17350183]
- (118). Ligon J; Hill S; Beck J; Zirkle R; Molnar I; Zawodny J; Money S; Schupp T Characterization of the Biosynthetic Gene Cluster for the Antifungal Polyketide Soraphen A from *Sorangium cellulosum* So ce26. *Gene* 2002, 285, 257–267. [PubMed: 12039053]
- (119). Waldron C; Matsushima P; Rosteck PR Jr.; Broughton MC; Turner J; Madduri K; Crawford KP; Merlo DJ; Baltz RH Cloning and Analysis of the Spinosad Biosynthetic Gene Cluster of *Saccharopolyspora spinosa*. *Chem. Biol.* 2001, 8, 487–499. [PubMed: 11358695]
- (120). Frank B; Knauber J; Steinmetz H; Scharfe M; Blocker H; Beyer S; Muller R Spiroketal Polyketide Formation in *Sorangium*: Identification and Analysis of the Biosynthetic Gene Cluster for the Highly Cytotoxic Spirangienes. *Chem. Biol.* 2007, 14, 221–233. [PubMed: 17317575]
- (121). Laureti L; Song L; Huang S; Corre C; Leblond P; Challis GL; Aigle B Identification of a Bioactive 51-Membered Macrolide Complex by Activation of a Silent Polyketide Synthase in *Streptomyces ambofaciens*. *Proc. Natl. Acad. Sci. U. S. A.* 2011, 108, 6258–6263. [PubMed: 21444795]
- (122). Gaitatzis N; Silakowski B; Kunze B; Nordsiek G; Blocker H; Hofle G; Muller R The Biosynthesis of the Aromatic Myxobacterial Electron Transport Inhibitor Stigmatellin Is Directed by a Novel Type of Modular Polyketide Synthase. *J. Biol. Chem.* 2002, 277, 13082–13090. [PubMed: 11809757]
- (123). Ohno S; Katsuyama Y; Tajima Y; Izumikawa M; Takagi M; Fujie M; Satoh N; Shin-Ya K; Ohnishi Y Identification and Characterization of the Streptazone E Biosynthetic Gene Cluster in *Streptomyces* sp. MSC090213JE08. *ChemBioChem* 2015, 16, 2385–2391. [PubMed: 26403163]
- (124). Olano C; Gomez C; Perez M; Palomino M; Pineda-Lucena A; Carbajo RJ; Brana AF; Mendez C; Salas JA Deciphering Biosynthesis of the RNA Polymerase Inhibitor Streptolydigin and Generation of Glycosylated Derivatives. *Chem. Biol.* 2009, 16, 1031–1044. [PubMed: 19875077]
- (125). Choi SS; Hur YA; Sherman DH; Kim ES Isolation of the Biosynthetic Gene Cluster for Tautomycetin, a Linear Polyketide T Cell-Specific Immunomodulator from *Streptomyces* sp. CK4412. *Microbiology* 2007, 153, 1095–1102. [PubMed: 17379718]
- (126). Adler JT; Cook M; Luo Y; Pitt SC; Ju J; Li W; Shen B; Kunnimalaiyaan M; Chen H Tautomycetin and Tautomycin Suppress the Growth of Medullary Thyroid Cancer Cells Via

- Inhibition of Glycogen Synthase Kinase-3 β . *Mol. Cancer Ther.* 2009, 8, 914–920. [PubMed: 19372564]
- (127). Fang J; Zhang Y; Huang L; Jia X; Zhang Q; Zhang X; Tang G; Liu W Cloning and Characterization of the Tetrocarcin a Gene Cluster from *Micromonospora chalicea* NRRL 11289 Reveals a Highly Conserved Strategy for Tetronate Biosynthesis in Spirotetronate Antibiotics. *J. Bacteriol.* 2008, 190, 6014–6025. [PubMed: 18586939]
- (128). Cooper HN; Cortes J; Bevitt DJ; Leadlay PF; Staunton J Analysis of a Gene Cluster from *S. longisporoflavus* Potentially Involved in Tetronasin Biosynthesis. *Biochem. Soc. Trans.* 1993, 21, 31S. [PubMed: 8449312]
- (129). Demydchuk Y; Sun Y; Hong H; Staunton J; Spencer JB; Leadlay PF Analysis of the Tetronomycin Gene Cluster: Insights into the Biosynthesis of a Polyether Tetronate Antibiotic. *ChemBioChem* 2008, 9, 1136–1145. [PubMed: 18404760]
- (130). Buntin K; Irschik H; Weissman KJ; Luxemburger E; Blocker H; Muller R Biosynthesis of Thuggacins in Myxobacteria: Comparative Cluster Analysis Reveals Basis for Natural Product Structural Diversity. *Chem. Biol.* 2010, 17, 342–356. [PubMed: 20416506]
- (131). Xiao Y; Li S; Niu S; Ma L; Zhang G; Zhang H; Zhang G; Ju J; Zhang C Characterization of Tiacumicin B Biosynthetic Gene Cluster Affording Diversified Tiacumicin Analogues and Revealing a Tailoring Dihaloxygenase. *J. Am. Chem. Soc.* 2011, 133, 1092–1105. [PubMed: 21186805]
- (132). Carlson JC; Fortman JL; Anzai Y; Li S; Burr DA; Sherman DH Identification of the Tirandamycin Biosynthetic Gene Cluster from *Streptomyces* sp. 307-9. *ChemBioChem* 2010, 11, 564–572. [PubMed: 20127927]
- (133). Sandmann A; Sasse F; Muller R Identification and Analysis of the Core Biosynthetic Machinery of Tubulysin, a Potent Cytotoxin with Potential Anticancer Activity. *Chem. Biol.* 2004, 11, 1071–1079. [PubMed: 15324808]
- (134). Cundliffe E; Bate N; Butler A; Fish S; Gandeche A; Merson-Davies L The Tylosin-Biosynthetic Genes of *Streptomyces fradiae*. *Antonie van Leeuwenhoek* 2001, 79, 229–234. [PubMed: 11816964]
- (135). Hashimoto T; Hashimoto J; Teruya K; Hirano T; Shinya K; Ikeda H; Liu HW; Nishiyama M; Kuzuyama T Biosynthesis of Versipelostatin: Identification of an Enzyme-Catalyzed [4 + 2]-Cycloaddition Required for Macrocyclization of Spirotetronate-Containing Polyketides. *J. Am. Chem. Soc.* 2015, 137, 572–575. [PubMed: 25551461]
- (136). Ogasawara Y; Katayama K; Minami A; Otsuka M; Eguchi T; Kakinuma K Cloning, Sequencing, and Functional Analysis of the Biosynthetic Gene Cluster of Macrolactam Antibiotic Vicenistatin in *Streptomyces halstedii*. *Chem. Biol.* 2004, 11, 79–86. [PubMed: 15112997]
- (137). Ahmadi MK; Fawaz S; Jones CH; Zhang G; Pfeifer BA Total Biosynthesis and Diverse Applications of the Nonribosomal Peptide-Polyketide Siderophore Yersiniabactin. *Appl. Environ. Microbiol.* 2015, 81, 5290–5298. [PubMed: 26025901]
- (138). Kevany BM; Rasko DA; Thomas MG Characterization of the Complete Zwittermicin a Biosynthesis Gene Cluster from *Bacillus cereus*. *Appl. Environ. Microbiol.* 2009, 75, 1144–1155. [PubMed: 19098220]
- (139). Zhang L; Hashimoto T; Qin B; Hashimoto J; Kozono I; Kawahara T; Okada M; Awakawa T; Ito T; Asakawa Y; Ueki M; Takahashi S; Osada H; Wakimoto T; Ikeda H; Shin-Ya K; Abe I Characterization of Giant Modular PKSs Provides Insight into Genetic Mechanism for Structural Diversification of Aminopolyol Polyketides. *Angew. Chem., Int. Ed.* 2017, 56, 1740–1745.
- (140). Keatinge-Clay AT Polyketide Synthase Modules Redefined. *Angew. Chem., Int. Ed.* 2017, DOI: 10.1002/anie.201701281.
- (141). Keatinge-Clay AT; Stroud RM The Structure of a Ketoreductase Determines the Organization of the Beta-Carbon Processing Enzymes of Modular Polyketide Synthases. *Structure* 2006, 14, 737–748. [PubMed: 16564177]
- (142). Del Vecchio F; Petkovic H; Kendrew SG; Low L; Wilkinson B; Lill R; Cortes J; Rudd BA; Staunton J; Leadlay PF Active-Site Residue, Domain and Module Swaps in Modular Polyketide Synthases. *J. Ind. Microbiol. Biotechnol* 2003, 30, 489–494. [PubMed: 12811585]

- (143). Starcevic A; Zucko J; Simunkovic J; Long PF; Cullum J; Hranueli D Clustscan: An Integrated Program Package for the Semi-Automatic Annotation of Modular Biosynthetic Gene Clusters and in Silico Prediction of Novel Chemical Structures. *Nucleic Acids Res.* 2008, 36, 6882–6892. [PubMed: 18978015]
- (144). Tang Y; Kim CY; Mathews II; Cane DE; Khosla C The 2.7-Angstrom Crystal Structure of a 194-kDa Homodimeric Fragment of the 6-Deoxyerythronolide B Synthase. *Proc. Natl. Acad. Sci. U. S. A.* 2006, 103, 11124–11129. [PubMed: 16844787]
- (145). Gay DC; Gay G; Axelrod AJ; Jenner M; Kohlhaas C; Kampa A; Oldham NJ; Piel J; Keatinge-Clay AT A Close Look at a Ketosynthase from a Trans-Acyltransferase Modular Polyketide Synthase. *Structure* 2014, 22, 444–451. [PubMed: 24508341]
- (146). Wang F; Wang Y; Ji J; Zhou Z; Yu J; Zhu H; Su Z; Zhang L; Zheng J Structural and Functional Analysis of the Loading Acyltransferase from Avermectin Modular Polyketide Synthase. *ACS Chem. Biol.* 2015, 10, 1017–1025. [PubMed: 25581064]
- (147). Sedgwick B; Cornforth JW The Biosynthesis of Long-Chain Fatty Acids. Stereochemical Differentiation in the Enzymic Incorporation of Chiral Acetates. *Eur. J. Biochem.* 1977, 75, 465–479. [PubMed: 18345]
- (148). Keatinge-Clay AT Stereocontrol within Polyketide Assembly Lines. *Nat. Prod. Rep.* 2016, 33, 141–149. [PubMed: 26584443]
- (149). Wu J; Kinoshita K; Khosla C; Cane DE Biochemical Analysis of the Substrate Specificity of the Beta-Ketoacyl-Acyl Carrier Protein Synthase Domain of Module 2 of the Erythromycin Polyketide Synthase. *Biochemistry* 2004, 43, 16301–16310. [PubMed: 15610024]
- (150). Guo X; Liu T; Deng Z; Cane DE Essential Role of the Donor Acyl Carrier Protein in Stereoselective Chain Translocation to a Fully Reducing Module of the Nanchangmycin Polyketide Synthase. *Biochemistry* 2012, 51, 879–887. [PubMed: 22229794]
- (151). Jenner M; Frank S; Kampa A; Kohlhaas C; Poplau P; Briggs GS; Piel J; Oldham NJ Substrate Specificity in Ketosynthase Domains from Trans-AT Polyketide Synthases. *Angew. Chem., Int. Ed.* 2013, 52, 1143–1147.
- (152). Jenner M; Afonso JP; Bailey HR; Frank S; Kampa A; Piel J; Oldham NJ Acyl-Chain Elongation Drives Ketosynthase Substrate Selectivity in Trans-Acyltransferase Polyketide Synthases. *Angew. Chem., Int. Ed.* 2015, 54, 1817–1821.
- (153). Witkowski A; Joshi AK; Lindqvist Y; Smith S Conversion of a Beta-Ketoacyl Synthase to a Malonyl Decarboxylase by Replacement of the Active-Site Cysteine with Glutamine. *Biochemistry* 1999, 38, 11643–11650. [PubMed: 10512619]
- (154). Valenzano CR; Lawson RJ; Chen AY; Khosla C; Cane DE The Biochemical Basis for Stereochemical Control in Polyketide Biosynthesis. *J. Am. Chem. Soc.* 2009, 131, 18501–18511. [PubMed: 19928853]
- (155). Zheng JT; Keatinge-Clay AT The Status of Type I Polyketide Synthase Ketoreductases. *MedChemComm* 2013, 4, 34–40.
- (156). Caffrey P Conserved Amino Acid Residues Correlating with Ketoreductase Stereospecificity in Modular Polyketide Synthases. *ChemBioChem* 2003, 4, 654–657. [PubMed: 12851937]
- (157). Reid R; Piagentini M; Rodriguez E; Ashley G; Viswanathan N; Carney J; Santi DV; Hutchinson CR; McDaniel R A Model of Structure and Catalysis for Ketoreductase Domains in Modular Polyketide Synthases. *Biochemistry* 2003, 42, 72–79. [PubMed: 12515540]
- (158). Keatinge-Clay AT A Tylosin Ketoreductase Reveals How Chirality Is Determined in Polyketides. *Chem. Biol.* 2007, 14, 898–908. [PubMed: 17719489]
- (159). Zheng JT; Keatinge-Clay AT Structural and Functional Analysis of C2-Type Ketoreductases from Modular Polyketide Synthases. *J. Mol. Biol.* 2011, 410, 105–117. [PubMed: 21570406]
- (160). Valenzano CR; You YO; Garg A; Keatinge-Clay A; Khosla C; Cane DE Stereospecificity of the Dehydratase Domain of the Erythromycin Polyketide Synthase. *J. Am. Chem. Soc.* 2010, 132, 14697–14699. [PubMed: 20925342]
- (161). Bonnett SA; Whicher JR; Papireddy K; Florova G; Smith JL; Reynolds KA Structural and Stereochemical Analysis of a Modular Polyketide Synthase Ketoreductase Domain Required for the Generation of a Cis-Alkene. *Chem. Biol.* 2013, 20, 772–783. [PubMed: 23790488]

- (162). Keatinge-Clay A Crystal Structure of the Erythromycin Polyketide Synthase Dehydratase. *J. Mol. Biol.* 2008, 384, 941–953. [PubMed: 18952099]
- (163). Akey DL; Razelun JR; Tehranisa J; Sherman DH; Gerwick WH; Smith JL Crystal Structures of Dehydratase Domains from the Curacin Polyketide Biosynthetic Pathway. *Structure* 2010, 18, 94–105. [PubMed: 20152156]
- (164). Kwan DH; Leadlay PF Mutagenesis of a Modular Polyketide Synthase Enoylreductase Domain Reveals Insights into Catalysis and Stereospecificity. *ACS Chem. Biol.* 2010, 5, 829–838. [PubMed: 20666435]
- (165). Zheng JT; Gay DC; Demeler B; White MA; Keatinge-Clay AT Divergence of Multimodular Polyketide Synthases Revealed by a Didomain Structure. *Nat. Chem. Biol.* 2012, 8, 615–621. [PubMed: 22634636]
- (166). Keatinge-Clay AT The Structural Relationship between Iterative and Modular PKSs. *Cell Chem. Biol.* 2016, 23, 540–542. [PubMed: 27203374]
- (167). Crosby J; Crump MP The Structural Role of the Carrier Protein–Active Controller or Passive Carrier. *Nat. Prod. Rep.* 2012, 29, 1111–1137. [PubMed: 22930263]
- (168). Lai JR; Koglin A; Walsh CT Carrier Protein Structure and Recognition in Polyketide and Nonribosomal Peptide Biosynthesis. *Biochemistry* 2006, 45, 14869–14879. [PubMed: 17154525]
- (169). Beld J; Sonnenschein EC; Vickery CR; Noel JP; Burkart MD The Phosphopantetheinyl Transferases: Catalysis of a Post-Translational Modification Crucial for Life. *Nat. Prod. Rep.* 2014, 31, 61–108. [PubMed: 24292120]
- (170). Busche A; Gottstein D; Hein C; Ripin N; Pader I; Tufar P; Eisman EB; Gu L; Walsh CT; Sherman DH; Lohr F; Guntert P; Dotsch V Characterization of Molecular Interactions between Acp and Halogenase Domains in the Curacin A Polyketide Synthase. *ACS Chem. Biol.* 2012, 7, 378–386. [PubMed: 22103656]
- (171). Vance S; Tkachenko O; Thomas B; Bassuni M; Hong H; Nietlispach D; Broadhurst W Sticky Swinging Arm Dynamics: Studies of an Acyl Carrier Protein Domain from the Mycolactone Polyketide Synthase. *Biochem. J.* 2016, 473, 1097–1110. [PubMed: 26920023]
- (172). Skiba MA; Sikkema AP; Fiers WD; Gerwick WH; Sherman DH; Aldrich CC; Smith JL Domain Organization and Active Site Architecture of a Polyketide Synthase C-Methyltransferase. *ACS Chem. Biol.* 2016, 11, 3319–3327. [PubMed: 27723289]
- (173). Maier T; Leibundgut M; Ban N The Crystal Structure of a Mammalian Fatty Acid Synthase. *Science* 2008, 321, 1315–1322. [PubMed: 18772430]
- (174). Khare D; Wang B; Gu L; Razelun J; Sherman DH; Gerwick WH; Hakansson K; Smith JL Conformational Switch Triggered by Alpha-Ketoglutarate in a Halogenase of Curacin A Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 2010, 107, 14099–14104. [PubMed: 20660778]
- (175). Herbst DA; Jakob RP; Zahringer F; Maier T Mycocerosic Acid Synthase Exemplifies the Architecture of Reducing Polyketide Synthases. *Nature* 2016, 531, 533–537. [PubMed: 26976449]
- (176). Pedelacq JD; Nguyen HB; Cabantous S; Mark BL; Listwan P; Bell C; Friedland N; Lockard M; Faille A; Mourey L; Terwilliger TC; Waldo GS Experimental Mapping of Soluble Protein Domains Using a Hierarchical Approach. *Nucleic Acids Res.* 2011, 39, e125. [PubMed: 21771856]
- (177). Trivedi OA; Arora P; Vats A; Ansari MZ; Tickoo R; Sridharan V; Mohanty D; Gokhale RS Dissecting the Mechanism and Assembly of a Complex Virulence Mycobacterial Lipid. *Mol. Cell* 2005, 17, 631–643. [PubMed: 15749014]
- (178). Dorival J; Annaval T; Risser F; Collin S; Roblin P; Jacob C; Gruez A; Chagot B; Weissman KJ Characterization of Intersubunit Communication in the Virginiamycin Trans-Acyl Transferase Polyketide Synthase. *J. Am. Chem. Soc.* 2016, 138, 4155–4167. [PubMed: 26982529]
- (179). Zeng J; Wagner DT; Zhang Z; Moretto L; Addison JD; Keatinge-Clay AT Portability and Structure of the Four-Helix Bundle Docking Domains of Trans-Acyltransferase Modular Polyketide Synthases. *ACS Chem. Biol.* 2016, 11, 2466–2474. [PubMed: 27362945]
- (180). Weissman KJ The Structural Basis for Docking in Modular Polyketide Biosynthesis. *ChemBioChem* 2006, 7, 485–494. [PubMed: 16470766]

- (181). Zhao C; Coughlin JM; Ju J; Zhu D; Wendt-Pienkowski E; Zhou X; Wang Z; Shen B; Deng Z Oxazolomycin Biosynthesis in *Streptomyces albus* JA3453 Featuring an “Acyltransferase-Less” Type I Polyketide Synthase That Incorporates Two Distinct Extender Units. *J. Biol. Chem.* 2010, 285, 20097–20108. [PubMed: 20406823]
- (182). Otsuka M; Ichinose K; Fujii I; Ebizuka Y Cloning, Sequencing, and Functional Analysis of an Iterative Type I Polyketide Synthase Gene Cluster for Biosynthesis of the Antitumor Chlorinated Polyenone Neocarzinil in *Streptomyces carzinostaticus*. *Antimicrob. Agents Chemother.* 2004, 48, 3468–3476. [PubMed: 15328113]
- (183). Liu TG; Cane DE; Deng ZX The Enzymology of Polyether Biosynthesis. *Methods Enzymol* 2009, 459, 187–214. [PubMed: 19362641]
- (184). Gay DC; Wagner DT; Meinke JL; Zogzas CE; Gay GR; Keatinge-Clay AT The Links Motif Zippers Trans-Acyltransferase Polyketide Synthase Assembly Lines into a Biosynthetic Megacomplex. *J. Struct. Biol.* 2016, 193, 196–205. [PubMed: 26724270]
- (185). Gu L; Eisman EB; Dutta S; Franzmann TM; Walter S; Gerwick WH; Skiniotis G; Sherman DH Tandem Acyl Carrier Proteins in the Curacin Biosynthetic Pathway Promote Consecutive Multienzyme Reactions with a Synergistic Effect. *Angew. Chem., Int. Ed.* 2011, 50, 2795–2798.
- (186). Rahman AS; Hotherall J; Crosby J; Simpson TJ; Thomas CM Tandemly Duplicated Acyl Carrier Proteins, Which Increase Polyketide Antibiotic Production, Can Apparently Function Either in Parallel or in Series. *J. Biol. Chem.* 2005, 280, 6399–6408. [PubMed: 15583005]
- (187). Wenzel SC; Kunze B; Hofle G; Silakowski B; Scharfe M; Blocker H; Muller R Structure and Biosynthesis of Myxochromides S1–3 in *Stigmatella aurantiaca*: Evidence for an Iterative Bacterial Type I Polyketide Synthase and for Module Skipping in Nonribosomal Peptide Biosynthesis. *ChemBioChem* 2005, 6 375–385. [PubMed: 15651040]
- (188). Li Y; Chen H; Ding Y; Xie Y; Wang H; Cerny RL; Shen Y; Du L Iterative Assembly of Two Separate Polyketide Chains by the Same Single-Module Bacterial Polyketide Synthase in the Biosynthesis of HSAF. *Angew. Chem., Int. Ed.* 2014, 53, 7524–7530.
- (189). Wenzel SC; Meiser P; Binz TM; Mahmud T; Muller R Nonribosomal Peptide Biosynthesis: Point Mutations and Module Skipping Lead to Chemical Diversity. *Angew. Chem., Int. Ed.* 2006, 45, 2296–2301.
- (190). Chan YA; Podevels AM; Kevany BM; Thomas MG Biosynthesis of Polyketide Synthase Extender Units. *Nat. Prod. Rep.* 2009, 26, 90–114. [PubMed: 19374124]
- (191). Wilson MC; Moore BS Beyond Ethylmalonyl-CoA: The Functional Role of Crotonyl-CoA Carboxylase/Reductase Homologs in Expanding Polyketide Diversity. *Nat. Prod. Rep.* 2012, 29, 72–86. [PubMed: 22124767]
- (192). Liscombe DK; Louie GV; Noel JP Architectures, Mechanisms and Molecular Evolution of Natural Product Methyltransferases. *Nat. Prod. Rep.* 2012, 29, 1238–1250. [PubMed: 22850796]
- (193). Wagner DT; Stevens DC; Mehaffey MR; Manion HR; Taylor RE; Brodbelt JS; Keatinge-Clay AT Alpha-Methylation Follows Condensation in the Gephyronic Acid Modular Polyketide Synthase. *Chem. Commun. (Cambridge, U. K.)* 2016, 52, 8822–8825.
- (194). Julien B; Shah S; Ziermann R; Goldman R; Katz L; Khosla C Isolation and Characterization of the Epothilone Biosynthetic Gene Cluster from *Sorangium cellulosum*. *Gene* 2000, 249, 153–160. [PubMed: 10831849]
- (195). Poust S; Phelan RM; Deng K; Katz L; Petzold CJ; Keasling JD Divergent Mechanistic Routes for the Formation of Gem-Dimethyl Groups in the Biosynthesis of Complex Polyketides. *Angew. Chem., Int. Ed* 2015, 54, 2370–2373.
- (196). Kellmann R; Mihali TK; Jeon YJ; Pickford R; Pomati F; Neilan BA Biosynthetic Intermediate Analysis and Functional Homology Reveal a Saxitoxin Gene Cluster in Cyanobacteria. *Appl. Environ. Microbiol.* 2008, 74, 4044–4053. [PubMed: 18487408]
- (197). Feng Y; Cronan JE *Escherichia coli* Unsaturated Fatty Acid Synthesis: Complex Transcription of the *fabA* Gene and in Vivo Identification of the Essential Reaction Catalyzed by FabB. *J. Biol. Chem.* 2009, 284, 29526–29535. [PubMed: 19679654]
- (198). Berkhan G; Merten C; Holec C; Hahn F The Interplay between a Multifunctional Dehydratase Domain and a C-Methyltransferase Effects Olefin Shift in Ambruticin Biosynthesis. *Angew. Chem., Int. Ed.* 2016, 55, 13589–13592.

- (199). Moldenhauer J; Gotz DC; Albert CR; Bischof SK; Schneider K; Sussmuth RD; Engeser M; Gross H; Bringmann G; Piel J The Final Steps of Bacillaene Biosynthesis in *Bacillus amyloliquefaciens* FZB42: Direct Evidence for β,γ Dehydration by a Trans-Acyltransferase Polyketide Synthase. *Angew. Chem., Int. Ed.* 2010, 49, 1465–1467.
- (200). Gay DC; Spear PJ; Keatinge-Clay AT A Double-Hotdog with a New Trick: Structure and Mechanism of the Trans-Acyltransferase Polyketide Synthase Enoyl-Isomerase. *ACS Chem. Biol.* 2014, 9, 2374–2381. [PubMed: 25089587]
- (201). Taft F; Brunjes M; Knobloch T; Floss HG; Kirschning A Timing of the $10,12\rightarrow 11,13$ Double Bond Migration During Ansamitocin Biosynthesis in *Actinosynnema pretiosum*. *J. Am. Chem. Soc.* 2009, 131, 3812–3813. [PubMed: 19292483]
- (202). Kang Q; Shen Y; Bai L Biosynthesis of 3,5-AHBA-Derived Natural Products. *Nat. Prod. Rep.* 2012, 29, 243–263. [PubMed: 22193711]
- (203). Fiers WD; Dodge GJ; Sherman DH; Smith JL; Aldrich CC Vinylogous Dehydration by a Polyketide Dehydratase Domain in Curacin Biosynthesis. *J. Am. Chem. Soc.* 2016, 138, 16024–16036. [PubMed: 27960309]
- (204). Tang L; Ward S; Chung L; Carney JR; Li Y; Reid R; Katz L Elucidating the Mechanism of Cis Double Bond Formation in Epothilone Biosynthesis. *J. Am. Chem. Soc.* 2004, 126, 46–47. [PubMed: 14709052]
- (205). Schneider K; Chen XH; Vater J; Franke P; Nicholson G; Borriss R; Sussmuth RD Macrolactin Is the Polyketide Biosynthesis Product of the PKS2 Cluster of *Bacillus amyloliquefaciens* FZB42. *J. Nat. Prod.* 2007, 70, 1417–1423. [PubMed: 17844999]
- (206). Tian Z; Sun P; Yan Y; Wu Z; Zheng Q; Zhou S; Zhang H; Yu F; Jia X; Chen D; Mandi A; Kurtan T; Liu W An Enzymatic [4 + 2] Cyclization Cascade Creates the Pentacyclic Core of Pyrroindomycins. *Nat. Chem. Biol.* 2015, 11, 259–265. [PubMed: 25730548]
- (207). Khare D; Hale WA; Tripathi A; Gu L; Sherman DH; Gerwick WH; Hakansson K; Smith JL Structural Basis for Cyclopropanation by a Unique Enoyl-Acyl Carrier Protein Reductase. *Structure* 2015, 23, 2213–2223. [PubMed: 26526850]
- (208). Holmes TC; May AE; Zaleta-Rivera K; Ruby JG; Skewes-Cox P; Fischbach MA; DeRisi JL; Iwatsuki M; Omura S; Khosla C Molecular Insights into the Biosynthesis of Guadinomine: A Type III Secretion System Inhibitor. *J. Am. Chem. Soc.* 2012, 134, 17797–17806. [PubMed: 23030602]
- (209). Berkhan G; Hahn F A Dehydratase Domain in Ambruticin Biosynthesis Displays Additional Activity as a Pyran-Forming Cyclase. *Angew. Chem., Int. Ed.* 2014, 53, 14240–14244.
- (210). Poplau P; Frank S; Morinaka BI; Piel J An Enzymatic Domain for the Formation of Cyclic Ethers in Complex Polyketides. *Angew. Chem., Int. Ed.* 2013, 52, 13215–13218.
- (211). Luhavaya H; Dias MV; Williams SR; Hong H; de Oliveira LG; Leadlay PF Enzymology of Pyran Ring Formation in Salinomycin Biosynthesis. *Angew. Chem., Int. Ed.* 2015, 54, 13622–13625.
- (212). Kwan JC; Donia MS; Han AW; Hirose E; Haygood MG; Schmidt EW Genome Streamlining and Chemical Defense in a Coral Reef Symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* 2012, 109, 20655–20660. [PubMed: 23185008]
- (213). Hotta K; Chen X; Paton RS; Minami A; Li H; Swaminathan K; Mathews II; Watanabe K; Oikawa H; Houk KN; Kim CY Enzymatic Catalysis of Anti-Baldwin Ring Closure in Polyether Biosynthesis. *Nature* 2012, 483, 355–358. [PubMed: 22388816]
- (214). Berkhan G; Merten C; Holec C; Hahn F The Interplay between a Multifunctional Dehydratase Domain and a C-Methyltransferase Effects Olefin Shift in Ambruticin Biosynthesis. *Angew. Chem., Int. Ed.* 2016, 55, 13589–13592.
- (215). Liu J; Zhu X; Kim SJ; Zhang W Antimycin-Type Depsipeptides: Discovery, Biosynthesis, Chemical Synthesis, and Bioactivities. *Nat. Prod. Rep.* 2016, 33, 1146. [PubMed: 27307039]
- (216). Zou Y; Yin H; Kong D; Deng Z; Lin S A Trans-Acting Ketoreductase in Biosynthesis of a Symmetric Polyketide Dimer SIA7248. *ChemBioChem* 2013, 14, 679–683. [PubMed: 23505159]

- (217). Wang H; Zhang H; Zou Y; Mi Y; Lin S; Xie Z; Yan Y; Zhang H Structural Insight into the Tetramerization of an Iterative Ketoreductase SiaM through Aromatic Residues in the Interfaces. *PLoS One* 2014, 9, e97996. [PubMed: 24901639]
- (218). Song L; Laureti L; Corre C; Leblond P; Aigle B; Challis GL Cytochrome P450-Mediated Hydroxylation Is Required for Polyketide Macrolactonization in Stambomycin Biosynthesis. *J. Antibiot.* 2014, 67, 71–76. [PubMed: 24220109]
- (219). Wang F; Langley R; Gulten G; Wang L; Sacchettini JC Identification of a Type III Thioesterase Reveals the Function of an Operon Crucial for *Mtb* Virulence. *Chem. Biol.* 2007, 14, 543–551. [PubMed: 17524985]
- (220). Maloney FP; Gerwick L; Gerwick WH; Sherman DH; Smith JL Anatomy of the Beta-Branching Enzyme of Polyketide Biosynthesis and Its Interaction with an Acyl-ACP Substrate. *Proc. Natl. Acad. Sci. U. S. A.* 2016, 113, 10316–10321. [PubMed: 27573844]
- (221). Geders TW; Gu L; Mowers JC; Liu H; Gerwick WH; Hakansson K; Sherman DH; Smith JL Crystal Structure of the ECH2 Catalytic Domain of CurF from *Lyngbya majuscula*. Insights into a Decarboxylase Involved in Polyketide Chain Beta-Branching. *J. Biol. Chem.* 2007, 282, 35954–35963. [PubMed: 17928301]
- (222). Gu L; Wang B; Kulkarni A; Geders TW; Grindberg RV; Gerwick L; Hakansson K; Wipf P; Smith JL; Gerwick WH; Sherman DH Metamorphic Enzyme Assembly in Polyketide Diversification. *Nature* 2009, 459, 731–735. [PubMed: 19494914]
- (223). Teufel R; Miyanaga A; Michaudel Q; Stull F; Louie G; Noel JP; Baran PS; Palfey B; Moore BS Flavín-Mediated Dual Oxidation Controls an Enzymatic Favorskii-Type Rearrangement. *Nature* 2013, 503, 552–556. [PubMed: 24162851]
- (224). Miyanaga A; Iwasawa S; Shinohara Y; Kudo F; Eguchi T Structure-Based Analysis of the Molecular Interactions between Acyltransferase and Acyl Carrier Protein in Vicenistatin Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 2016, 113, 1802–1807. [PubMed: 26831085]
- (225). Hong H; Fill T; Leadlay PF A Common Origin for Guanidinobutanoate Starter Units in Antifungal Natural Products. *Angew. Chem., Int. Ed.* 2013, 52, 13096–13099.
- (226). Qu X; Jiang N; Xu F; Shao L; Tang G; Wilkinson B; Liu W Cloning, Sequencing and Characterization of the Biosynthetic Gene Cluster of Sanglifehrin A, a Potent Cyclophilin Inhibitor. *Mol. Biosyst* 2011, 7, 852–861. [PubMed: 21416665]
- (227). Hicks LM; Moffitt MC; Beer LL; Moore BS; Kelleher NL Structural Characterization of in Vitro and in Vivo Intermediates on the Loading Module of Microcystin Synthetase. *ACS Chem. Biol.* 2006, 1, 93–102. [PubMed: 17163649]
- (228). Gregory MA; Petkovic H; Lill RE; Moss SJ; Wilkinson B; Gaisser S; Leadlay PF; Sheridan RM Mutasynthesis of Rapamycin Analogues through the Manipulation of a Gene Governing Starter Unit Biosynthesis. *Angew. Chem., Int. Ed.* 2005, 44, 4757–4760.
- (229). Gu L; Geders TW; Wang B; Gerwick WH; Hakansson K; Smith JL; Sherman DH GNAT-Like Strategy for Polyketide Chain Initiation. *Science* 2007, 318, 970–974. [PubMed: 17991863]
- (230). Jaroszewski L; Li Z; Cai XH; Weber C; Godzik A FFAS Server: Novel Features and Applications. *Nucleic Acids Res.* 2011, 39, W38–44. [PubMed: 21715387]
- (231). Zou XW; Liu YC; Hsu NS; Huang CJ; Lyu SY; Chan HC; Chang CY; Yeh HW; Lin KH; Wu CJ; Tsai MD; Li TL Structure and Mechanism of a Nonhaem-Iron SAM-Dependent C-Methyltransferase and Its Engineering to a Hydratase and an O-Methyltransferase. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2014, 70, 1549–1560. [PubMed: 24914966]
- (232). Du L; Lou L PKS and NRPS Release Mechanisms. *Nat. Prod. Rep.* 2010, 27, 255–278. [PubMed: 20111804]
- (233). Horsman ME; Hari TP; Boddy CN Polyketide Synthase and Non-Ribosomal Peptide Synthetase Thioesterase Selectivity: Logic Gate or a Victim of Fate? *Nat. Prod. Rep.* 2016, 33, 183–202. [PubMed: 25642666]
- (234). Tsai SC; Miercke LJ; Krucinski J; Gokhale R; Chen JC; Foster PG; Cane DE; Khosla C; Stroud RM Crystal Structure of the Macrocyclic-Forming Thioesterase Domain of the Erythromycin Polyketide Synthase: Versatility from a Unique Substrate Channel. *Proc. Natl. Acad. Sci. U. S. A.* 2001, 98, 14808–14813. [PubMed: 11752428]

- (235). Claxton HB; Akey DL; Silver MK; Admiraal SJ; Smith JL Structure and Functional Analysis of RifR, the Type II Thioesterase from the Rifamycin Biosynthetic Pathway. *J. Biol. Chem.* 2009, 284, 5021–5029. [PubMed: 19103602]
- (236). Bruner SD; Weber T; Kohli RM; Schwarzer D; Marahiel MA; Walsh CT; Stubbs MT Structural Basis for the Cyclization of the Lipopeptide Antibiotic Surfactin by the Thioesterase Domain SrfTE. *Structure* 2002, 10, 301–310. [PubMed: 12005429]
- (237). Koglin A; Lohr F; Bernhard F; Rogov VV; Frueh DP; Strieter ER; Mofid MR; Guntert P; Wagner G; Walsh CT; Marahiel MA; Dotsch V Structural Basis for the Selectivity of the External Thioesterase of the Surfactin Synthetase. *Nature* 2008, 454, 907–911. [PubMed: 18704089]
- (238). Gehret JJ; Gu L; Gerwick WH; Wipf P; Sherman DH; Smith JL Terminal Alkene Formation by the Thioesterase of Curacin A Biosynthesis: Structure of a Decarboxylating Thioesterase. *J. Biol. Chem.* 2011, 286, 14445–14454. [PubMed: 21357626]
- (239). Zhou Y; Prediger P; Dias LC; Murphy AC; Leadlay PF Macrolide Formation by the Thioesterase of a Modular Polyketide Synthase. *Angew. Chem., Int. Ed.* 2015, 54, 5232–5235.
- (240). Ju J; Li W; Yuan Q; Peters NR; Hoffmann FM; Rajski SR; Osada H; Shen B Functional Characterization of *ttmM* Unveils New Tautomycin Analogs and Insight into Tautomycin Biosynthesis and Activity. *Org. Lett.* 2009, 11, 1639–1642. [PubMed: 19281218]
- (241). McCarthy JG; Eisman EB; Kulkarni S; Gerwick L; Gerwick WH; Wipf P; Sherman DH; Smith JL Structural Basis of Functional Group Activation by Sulfotransferases in Complex Metabolic Pathways. *ACS Chem. Biol.* 2012, 7, 1994–2003. [PubMed: 22991895]
- (242). Beyer S; Kunze B; Silakowski B; Muller R Metabolic Diversity in Myxobacteria: Identification of the Myxalamid and the Stigmatellin Biosynthetic Gene Cluster of *Stigmatella aurantiaca* Sg a15 and a Combined Polyketide-(Poly)Peptide Gene Cluster from the Epothilone Producing Strain *Sorangium cellulosum* So ce90. *Biochim. Biophys. Acta, Gene Struct. Expression* 1999, 1445, 185–195.
- (243). Barajas JF; Phelan RM; Schaub AJ; Kliewer JT; Kelly PJ; Jackson DR; Luo R; Keasling JD; Tsai SC Comprehensive Structural and Biochemical Analysis of the Terminal Myxalamid Reductase Domain for the Engineered Production of Primary Alcohols. *Chem. Biol.* 2015, 22, 1018–1029. [PubMed: 26235055]
- (244). Liu X; Walsh CT Cyclopiazonic Acid Biosynthesis in *Aspergillus* Sp.: Characterization of a Reductase-Like R* Domain in Cyclopiazonate Synthetase That Forms and Releases Cyclo-Acetoacetyl-L-Tryptophan. *Biochemistry* 2009, 48, 8746–8757. [PubMed: 19663400]
- (245). Zhou X; Ma Z; Dong D; Wu B Arylamine N-Acetyltransferases: A Structural Perspective. *Br. J. Pharmacol.* 2013, 169, 748–760. [PubMed: 23517104]
- (246). Xue Y; Zhao L; Liu HW; Sherman DH A Gene Cluster for Macrolide Antibiotic Biosynthesis in *Streptomyces venezuelae*: Architecture of Metabolic Diversity. *Proc. Natl. Acad. Sci. U. S. A.* 1998, 95, 12111–12116. [PubMed: 9770448]
- (247). Kwon HJ; Smith WC; Scharon AJ; Hwang SH; Kurth MJ; Shen B C–O Bond Formation by Polyketide Synthases. *Science* 2002, 297, 1327–1330. [PubMed: 12193782]
- (248). Tang MC; He HY; Zhang F; Tang GL Baeyer–Villiger Oxidation of Acyl Carrier Protein-Tethered Thioester to Acyl Carrier Protein-Linked Thiocarbonate Catalyzed by a Monooxygenase Domain in FR901464 Biosynthesis. *ACS Catal* 2013, 3, 444–447.
- (249). Piel J A Polyketide Synthase-Peptide Synthetase Gene Cluster from an Uncultured Bacterial Symbiont of *Paederus* Beetles. *Proc. Natl. Acad. Sci. U. S. A.* 2002, 99, 14002–14007. [PubMed: 12381784]
- (250). Bosserman MA; Downey T; Noinaj N; Buchanan SK; Rohr J Molecular Insight into Substrate Recognition and Catalysis of Baeyer–Villiger Monooxygenase MtmOIV, the Key Frame-Modifying Enzyme in the Biosynthesis of Anticancer Agent Mithramycin. *ACS Chem. Biol* 2013, 8, 2466–2477. [PubMed: 23992662]
- (251). Liu W; Christenson SD; Standage S; Shen B Biosynthesis of the Enediyne Antitumor Antibiotic C-1027. *Science* 2002, 297, 1170–1173. [PubMed: 12183628]
- (252). Jeon B.-s.; Wang S-A; Rusczycki MW; Liu H.-w. Natural [4 + 2]-Cyclases. *Chem. Rev.* 2017, DOI: 10.1021/acs.chemrev.6b00578.

Figure 1.

Model of an assembly line composed of α -, β -, γ -, and δ -type cis-AT modules. The stereodimeric assembly line shows how domains may be oriented relative to one another along the 2-fold axis of the homodimeric assembly line [structures utilized: KS (PDB 2QO3), AT (PDB 2QO3), DH (PDB 3EL6), KR (PDB 3SLK and 4IMP), ER (PDB 3SLK), ACP (PDB 2LIW), and TE (PDB 1MN6)]. The corresponding biosynthetic model is shown below. Domains that insert into others (AT and ER) branch off. Acyl carrier proteins are represented by solid circles.

2.1.1. Unusual domain organizations

2.1.2. Dimeric ER

2.1.3. Split module

2.1.4. Incorporation of *trans*-AT module 2.1.5. AT⁰ 2.1.6. KS⁰ 2.1.7. KR⁰ in loading machinery

2.1.8. Tandem ACPs

Figure 2.

Uncommon architecture. Domain insertions, rare domain orderings, split modules, and other peculiar features provide insight into the organization of *cis*-AT modules.

2.2.1. Use of domain outside module boundaries

2.2.2. Iteration

2.2.3. Module-skipping

Figure 3.

Uncommon module behavior. Modules can behave non-collinearly by accessing neighboring processing enzymes, by performing multiple catalytic cycles, or by not modifying the acyl chain at all.

2.3.1. Unusual AT function

Figure 4.

Uncommon activities not requiring trans enzymes. Acyltransferases can select unusual extender units. Embedded MTs can methylate α -carbon as well as α -hydroxyl, β -hydroxyl, and β -keto groups. Dehydratases can help generate shifted β/γ double bonds.

Figure 5.

More uncommon activities not requiring trans enzymes. Certain DHs can dehydrate δ -hydroxy intermediates. Carbocyclizations can occur while the chain is module-bound. PLP-dependent AMTs stereoselectively convert β -keto groups into β -amines. Pyran formation, halogenation, and α -oxidation can also be mediated by embedded enzymes.

Figure 6.

Uncommon activities requiring *trans* enzymes. *Trans* enzymes can be recruited to a particular ACP to perform chemistry on its bound acyl chain. This chemistry is often conducted on intermediates bound to ACP of the final module (e.g., polyether reactions, deacylation) or to ACP at the C-terminal end of a polypeptide (e.g., oxidation/cyclization of naphthalenic ansamycins, β -amination of α/β -unsaturated intermediates, β -branching). A Favorskii rearrangement in the synthesis of ambruticin (perhaps the most mysterious reaction discussed) yields a cyclopropyl moiety.

Figure 7.

Loading. Acyl groups can be loaded onto a cis-AT module through AT_L , A_L , or $GNAT_L$. Modification of an acquired acyl group is possible for each type of loading machinery before transfer of the primer unit to the first module.

4.1.1. TE

4.1.1.1. Lactonization

4.1.1.2. Hydrolysis

4.1.1.3. C2 activation in conjunction with cyclization or hydrolysis

4.1.1.4. Lactamization

4.1.1.5. Terminal alkene formation

4.1.2. Reductase (R)

Figure 8.

Offloading by embedded enzymes. Embedded TEs mediate lactonization, hydrolysis, lactamization, or macrodiolide formation, depending on the nucleophile that most quickly attacks at C1 of the acyl-TE intermediate. The curacin TE collaborates with an embedded sulfotransferase (ST) to generate a terminal alkene product. Acyl chains are offloaded as aldehydes by the reductase (R) domains of the coelimycin, cyclizidine, and streptazone assembly lines.

Figure 9.

Offloading by nonembedded enzymes and mysterious chain release. Tetronate condensing enzymes (CEs), amide synthases (ASs), TEIIs, and TEIIIs each release acyl chains from assembly lines in a characteristic fashion. More mysterious chain release events are represented by the Cy domain of the stigmatellin assembly line that helps generate a chromone ring system, the KS⁰ from the ebelactone assembly line that may participate in β -lactone formation, and a mysterious C-terminal enzymatic domain that, together with two aminohydrolases, mediates uracil ring formation during offloading of the acyl chain from the cylindrospormycin assembly line.