Chapter 2

The Assembly Line Enzymology of Polyketide Biosynthesis

Marisa Till and Paul R. Race

Abstract

Polyketides are a structurally and functionally diverse family of bioactive natural products that have found widespread application as pharmaceuticals, agrochemicals, and veterinary medicines. In bacteria complex polyketides are biosynthesized by giant multifunctional megaenzymes, termed modular polyketide synthases (PKSs), which construct their products in a highly coordinated assembly line-like fashion from a pool of simple precursor substrates. Not only is the multifaceted enzymology of PKSs a fascinating target for study, but it also presents considerable opportunities for the reengineering of these systems affording access to functionally optimized unnatural natural products. Here we provide an introductory primer to modular polyketide synthase structure and function, and highlight recent advances in the characterization and exploitation of these systems.

Key words Polyketide synthase, Biosynthesis, Structural enzymology, Natural products, Synthetic biology

1 Introduction

Polyketide natural products are a proven source of high-value bioactive small molecules [1]. These include compounds that have become mainstays for the treatment of human and animal diseases, and others that have found application as agrochemicals, flavors, fragrances, and nutraceuticals [2, 3]. Illustrative examples of clinically relevant polyketides include the statin family of cholesterol lowering agents, erythromycin A and related macrolide antibiotics, and the antiparasitic avermectins, which are used extensively in veterinary medicine (Fig. 1) [3–5]. The wealth of structural and functional diversity found within the polyketides is almost unsurpassed amongst natural products [6]. This, coupled to their exploitable bioactivities, has made both the compounds themselves, and the cellular machineries responsible for their biosynthesis, a source of intrigue and fascination for researchers worldwide.

Polyketides of bacterial origin represent a significant proportion of the natural product pool [1]. These compounds function to confer evolutionary fitness to the producing host, acting as defense

Bradley S. Evans (ed.), *Nonribosomal Peptide and Polyketide Biosynthesis: Methods and Protocols*, Methods in Molecular Biology, vol. 1401, DOI 10.1007/978-1-4939-3375-4_2,

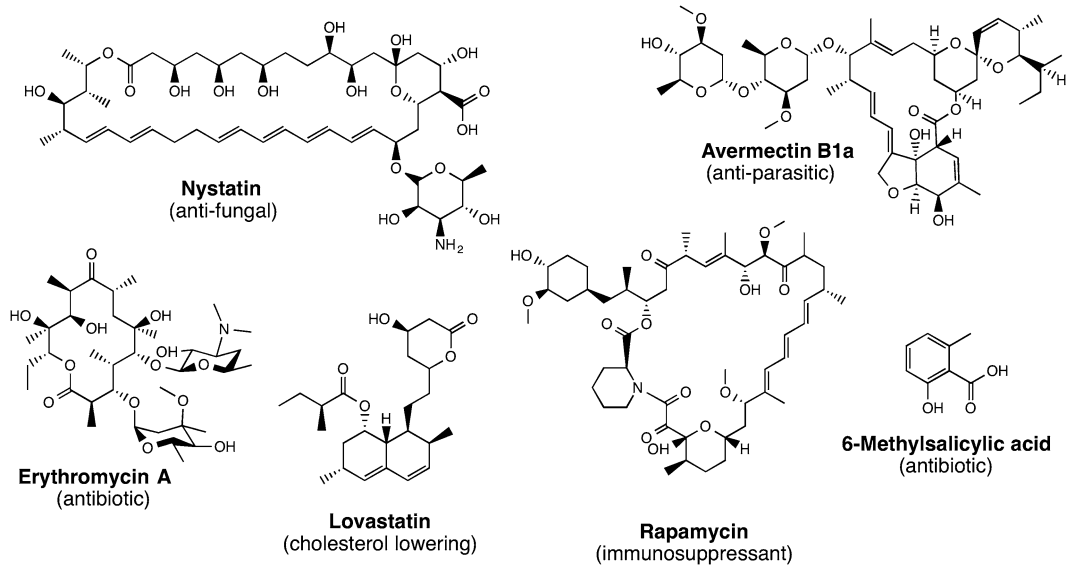

Fig. 1 Illustrative examples of polyketide natural products and their respective bioactivities

agents or signaling molecules [7, 8]. Commonly, they are biosynthesized through the action of giant multifunctional megaenzymes, termed modular polyketide synthases (PKSs), which utilize sequential condensation chemistry to construct elaborate product scaffolds from simple carboxylic acid substrates [9, 10]. This biosynthetic logic permits the assembly of a suite of molecules that are in principle of almost unlimited chemical diversity. In addition, the highly modular architecture of these systems raises intriguing possibilities for their modification and manipulation through the addition, removal, or substitution of synthase modules or domains therein [11, 12]. This approach has been exploited to generate an even larger portfolio of useful molecules, the so-called unnatural natural products [13]. With the advent of synthetic biology and the increasing adoption of the concepts and practices of rational design, it is inevitable that synthase reengineering will become more reliable and expedient in the future [14, 15]. Here we outline the basic principles of modular PKS enzymology, highlighting recent developments in the characterization of these systems, and discuss future prospects for harnessing their biosynthetic potential.

2 Modular PKSs and the Biosynthetic Process

Modular PKSs are amongst the largest and most sophisticated enzymes known. In extreme cases their proteinogenic mass can exceed 5 MDa. The distinguishing structural feature of these systems is their assembly line-like architecture, which comprises a series of linearly arranged, multi-domain extension modules,

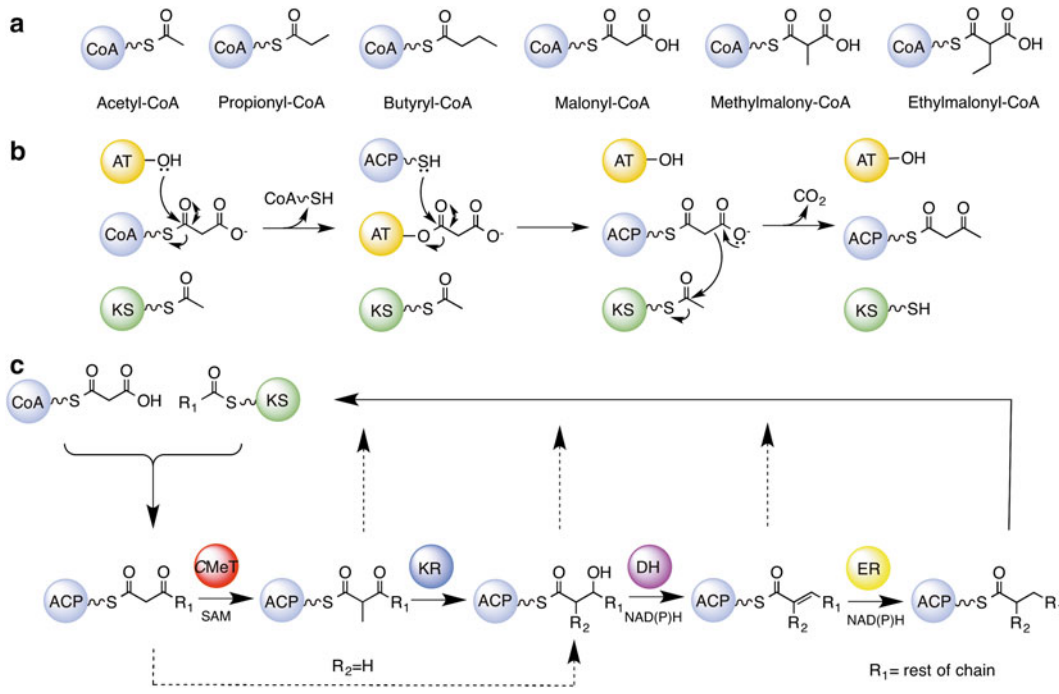

Fig. 2 (**a**) Chemical structures of example PKS starter and extender units. (**b**) Individual steps involved in extender unit selection and chain incorporation. (**c**) Illustrative chain modifications catalyzed by module embedded PKS domains

housed in sequence within giant polypeptide chains [1, 16]. Biosynthesis upon these systems proceeds in step-wise fashion, through the processive transfer and extension of the nascent polyketide chain, as it progresses from one module to the next. Each module within the synthase complex incorporates a single carboxylic acid substrate, derived from a pool of simple precursors, into the pathway product (Fig. 2) [1, 6, 17, 18]. The first module of the PKS initiates biosynthesis through the selection and loading of a starter unit. Each subsequent module within the PKS catalyzes the thiotemplated addition of a defined extender unit to this initial priming substrate. In some cases, however, phenomena including module 'skipping' or iterative 'stuttering' have been reported [19, 20]. The structures and properties of polyketides biosynthesized via this route are dictated by the number, identity, and incorporation order of the starter and extender units from which they are assembled. These factors are defined by the sequence of the modules that constitute the PKS assembly line. For this reason modular PKSs can be considered to possess an inherent biosynthetic programming, where synthase nucleotide sequence is colinearly related to product chemistry [21, 22]. This relationship provides a powerful tool for the prediction and analysis of synthase function, requiring knowledge only of PKS gene sequence. In addition,

given that the open reading frames (ORFs) that encode PKS assembly lines and their associated tailoring enzymes are commonly clustered within the genomes of biosynthetically competent microorganisms, PKS nucleotide sequences provide a useful analytical tool by which synthase components can be readily identified and functionally assigned [23, 24].

To perform the required set of chemical reactions necessary to extend a polyketide chain by a single acyl unit each PKS extension module must minimally house three essential domains. These are; an acyltransferase (AT) domain, responsible for the selection and loading of carboxylic acid extender units derived from their coenzyme-A (CoA) thioesters; an acyl carrier protein (ACP), possessing a phosphopantetheine arm, which acts as a site of covalent tethering for both the growing product chain and selected extender units; and a ketosynthase (KS) domain, which catalyzes the carbon–carbon bond forming Claisen condensation of the downstream product chain with the upstream ACP tethered extender unit (Fig. 2). This minimal KS-AT-ACP module architecture may be elaborated to include additional auxiliary domains that further modify the incorporated acyl unit. Such modifications may take place at the reactive α-position or the β-carbonyl. An example of the former is the introduction of an α-methyl substituent via a reaction catalyzed by a *C*-methyltransferase (*C*MeT), which exploits *S*-adenosyl methionine as a methyl source (Fig. 2). Alternatively the β-carbonyl of the incorporated acyl unit may be reduced in a reaction catalyzed by a ketoreductase domain (KR) to form a β-hydroxy substituent. This may in turn undergo dehydration, catalyzed by a dehydratase domain (DH), yielding an α/β-unsaturated thioester. Finally an enoyl reductase (ER) may act to reduce this species yet further yielding a product fully reduced at the β-position (Fig. 2). Both the substrate selectivity of the module embedded AT domain and the module's compliment of auxiliary domains, offer complementary mechanisms for achieving diverse product chemistry [1, 6, 17, 18].

Upon reaching the final module of the PKS assembly line the immature polyketide product is liberated from the synthase via cleavage of its covalent tether. In the majority of modular PKSs this process is catalyzed by a thioesterase that forms either the *C*-terminal domain of the synthase, or functions as a stand-alone enzyme [25]. Frequently, following release from the PKS, the biosynthesized product is subjected to additional site-specific tailoring modifications catalyzed by free-standing enzymes. Tailoring modifications are in many cases implicitly required to establish the bioactivity of the synthase product. Examples include glycosylation, halogenation, methylation, and hydroxylation among others [1, 10, 26].

Modular PKS enzymology is readily distinguishable from that of other polyketide synthases. Iteratively acting type I PKSs (iPKSs) found in fungi are smaller but inherently more sophisticated

systems, which use a single set of functional domains to accomplish the biosynthesis of their products, in a manner analogous to that of the type I fatty acid synthase [27–30]. Uniquely, the domains that comprise iPKSs often only catalyze their specific transformations during a subset of chain extension cycles, establishing a defined biosynthetic program that violates the colinearity of synthase sequence and product chemistry. Type II PKSs are transiently assembled iteratively acting complexes composed of discrete proteins [31, 32]. Consequently type II systems are significantly smaller and less intricate than both modular and iteratively acting type I PKSs. Biosynthesis upon type II PKSs proceeds in two discernable phases, initiation and chain extension, which are catalyzed by the minimal type II PKS assembly, comprising two KS like condensing enzymes and an ACP. The minimal KS-ACP machinery may be further elaborated through the incorporation of a selection of additional domains including ketoreductases, cyclases, and aromatases. Type III PKSs, which are widely distributed amongst bacteria, fungi, and plants, are in essence free standing KS domains that act to catalyze the sequential condensation of acetate units to a CoA derived starter unit [33–35]. Chemical diversity in the products of type III systems arises due to the choice of starter unit, the number of chain extension steps catalyzed, the mechanism of cyclization, and product modification through the action of allied tailoring enzymes.

3 The DEBS Paradigm

Early attempts to delineate the enzymology of modular PKSs focused on the deoxyerythronolide B synthase (DEBS). This system is responsible for the biosynthesis of 6-deoxyerythronolide (6-dEB), the macrolide nucleus of the clinically relevant antibiotic erythromycin A [4]. Studies of DEBS resulted in the establishment of many of the central tenets of modular PKS enzymology and as such this system serves as a useful test case for illustrating the basic principles of PKS function [36].

DEBS is encoded for by a series of ORFs clustered on the genome of the soil-dwelling actinomycete *Saccharopolyspora erythraea* [37]. This cluster comprises the genes *ery*AI-III that encode the core biosynthetic machinery of DEBS (DEBS 1-3), along with a number of associated tailoring and regulatory proteins (Fig. 3) [36, 38]. DEBS was the first modularPKS gene cluster to be sequenced, thus providing initial evidence for the distinctive modular assembly line-like architecture of bacterial PKSs [21, 37]. This analysis also proved instrumental in establishing the colinear relationship between synthase nucleotide sequence and product chemistry.

Each module within DEBS consists of a discrete set of domains that are responsible for the incorporation and reductive tailoring of

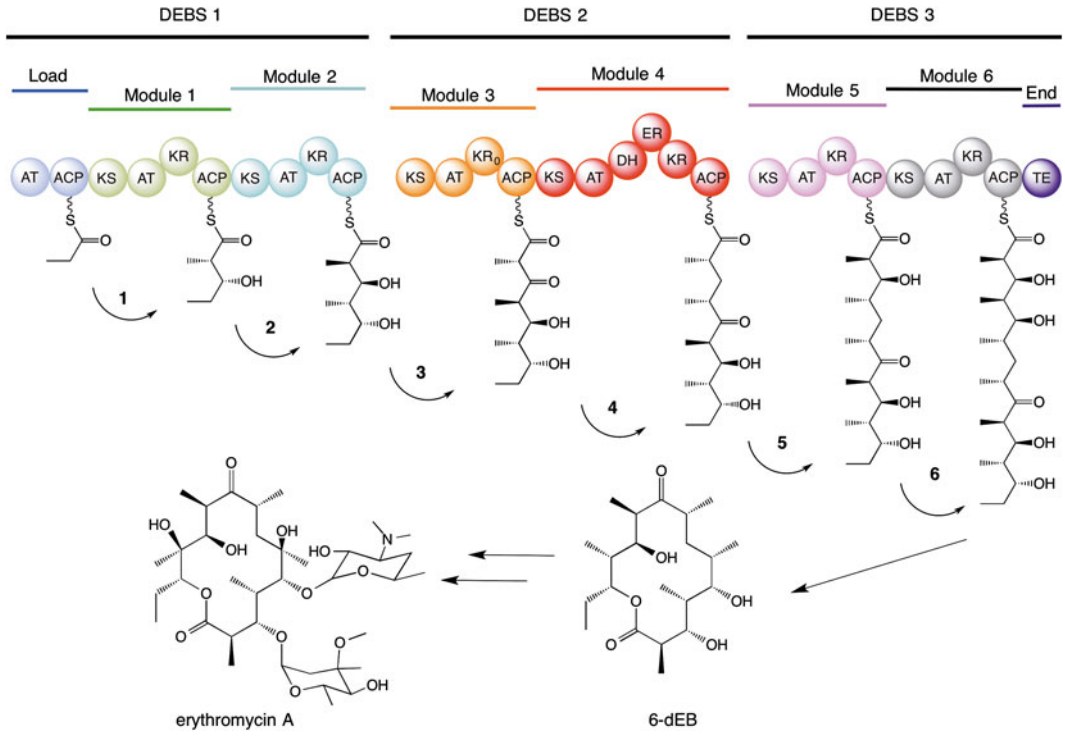

Fig. 3 Domain organization of the 6-deoxyerythronolide B synthase (DEBS) and the biosynthetic route to erythromycin A. Domains are colored based on their host module. *Numbered arrows* indicate the direction and order of product chain extension and transfer

a designated extender unit at a precise location within the growing product chain. Each domain is functionally specific and contributes to the incorporation of a single extender unit within its host module. Biosynthesis upon DEBS is initiated by a loading module, located at the *N*-terminus of DEBS 1, which consists of an AT, nominally selective for a propionate starter unit, and an ACP onto which this unit becomes tethered. It has been reported that this initial loading step is more promiscuous than was initially presumed and that in addition to propionate, non-native extender units can also be accepted by the starter module AT [39–41].

The remainder of DEBS comprises six extension modules distributed across DEBS 1-3 and a *C*-terminal TE domain that catalyzes the release and cyclization of the immature PKS product, the 14-membered 6-dEB macrolactone [42]. The modules within DEBS are numbered in the order in which they act during biosynthesis and each houses a single AT, ACP, and KS domain. Each of the extender module embedded AT domains of DEBS is selective for (2*S*)-methylmalonyl-CoA. In addition to their compliment of core domains, modules 1, 2, 5, and 6 also house active KRs. Consequently, extender units incorporated at these locations undergo reductive modification yielding a hydroxyl substituent at

their β-position. The stereochemistries of the hydroxyl groups are dictated at the level of protein structure. The KR domain can attack the β-ketone from either side of the acyl chain, resulting in either a d- or l-configuration. The structure of the KR domain dictates the direction of substrate entry into the enzyme's active site, which in turn determines the face of the keto-group that is presented for hydride transfer from NADPH [43]. Module 4 of DEBS has a full complement of reductive domains (KR, DH, and ER). Extender units incorporated at this location undergo complete β-keto reduction, yielding an α/β alkane. By contrast, DEBS module 3 houses no functional reductive domains resulting in retention of the β-keto group. Notably, however, this module does contain an oft unreported 'broken' KR that controls epimerization of the incorporated methylmalonyl extender unit, a function that was previously ascribed to this module's KS domain [44, 45].

To generate the fully functional DEBS PKS there is a requirement to assemble DEBS 1-3 into a single complex. This is facilitated by protein–protein interaction interfaces at the *N*- and *C*-termini of the individual DEBS polypeptides, which serve to link these three proteins together to form the intact megaenzyme. DEBS docking domains comprise complimentary *N*- and *C*-terminal helical bundles of 30–50 and 80–130 residues respectively [46]. Complex formation between complimentary docking domains is implicitly required to permit transfer of the growing product chain between neighboring modules that are housed on different polypeptides.

Upon release from DEBS the cyclized 6-dEB macrolactone is subjected to post-PKS tailoring, catalyzed by a complement of stand-alone enzymes encoded within the *ery* gene cluster. 6-dEB is hydroxylated and glycosylated to form erythromycin D, which is then converted to erythromycin A through the action of the *O*-methyl transferase EryG and the hydroxylase EryK [47]. Interestingly these two transformations occur in a sequence independent fashion, forming the isolatable intermediates erythromycin B and erythromycin C respectively.

4 The Structural Enzymology of Modular PKSs

Significant insights into the enzymology of PKS chain extension and processing have been provided by structural studies of isolated synthase domains. Crystal structures have been reported for AT, ACP, KS, KR, DH, and ER domains amongst others, and NMR spectroscopy has been used extensively for the structural characterization of ACPs [38, 48–51]. These data have proven informative in establishing the roles and contributions of each domain during biosynthesis. Here we provide structural descriptions of the three core PKS domains (AT, ACP, and KS), but direct readers to

references [49, 50] for more comprehensive descriptions of PKS domain structure.

Acyltransferases have been shown to possess a distinct two sub-domain architecture comprising a larger hydrolase like sub-domain fused to a smaller ferredoxin like sub-domain [52–55]. The AT active site sits at the interface of the two sub-domains, at the base of a solvent exposed channel, and houses an invariant His-Ser catalytic dyad. ATs are proposed to employ a ping-pong bi-bi mechanism that proceeds via an acyl-enzyme intermediate. The intermediate is stabilized during catalysis by an oxyanion hole formed by backbone amides from neighboring amino acids within the enzyme active site. Resolution of the acyl-enzyme intermediate occurs only in the presence of thiol nucleophiles rendering this intermediate sufficiently stable to permit its isolation and characterization in vitro [56]. The topology of the AT active site is in part dictated by amino acids that form defined substrate-selectivity motifs. Residues that occupy these positions play a role in dictating the specific acyl extender unit that the AT selects [57, 58]. Substitution of these motifs has been used to alter AT selectivity giving rise to new natural products, though the universality of this approach remains questionable [59–61]. Recently it has been proposed that extender unit selectivity is dictated more generally by a combination of structural features distributed throughout the enzyme fold [50].

ACPs are small negatively charged helical bundles that provide a site of anchorage for acyl intermediates during biosynthesis [51, 62]. The covalent attachment of intermediates to the ACP occurs via a post-translationally modified serine residue bearing a phosphopantetheine arm, which forms part of a conserved Asp-Ser-Leu motif. This motif is located at the *N*-terminus of helix 2 of the protein, which is considered the key portion of the ACP for mediating interactions with each of its respective binding partners [63–65]. In many instances ACPs exhibit a high degree of specificity for their cognate intra-module domains [63, 66]. Undoubtedly this property is dictated by a combination of ACP sequence and the identity of the substrate or intermediate to which the ACP is tethered. Although originally considered as somewhat as a passive component of the biosynthetic machinery, it is becoming increasingly clear that ACPs play a more active role. This includes for example shielding of tethered acyl units to allow their presentation at appropriate time points or locations within their host module [67–69].

Ketosynthase domains are dimeric proteins with a conserved thiolase fold [52, 53, 70]. Structurally KSs are composed of two α-β-α-β-α protomers arranged in the form of a five-layered core, within which three layers of α-helices are separated by two layers of β-sheets. Although there is some structural divergence between KSs, they all exhibit a small number of universal features. These

include retention of the overall fold described above, extensive and highly hydrophobic dimer interfaces, and a conserved active site cysteine that acts as a site for the covalent attachment of substrates and intermediates. Variations in KS structure are largely confined to the enzyme's active site and associated regions, and have significant impact on substrate selectivity [52, 53, 71]. Examples of this include the identity and location of the key catalytic active site residues (excluding the universally conserved Cys) and the steric and electrostatic topology of the active site and its associated solvent exposed access channel. Such variations influence the ability of KSs to act upon intermediates of different chain length, saturation state and stereochemistry, and inadvertently provide a proofreading or gate-keeping function [66, 72, 73]. KS catalyzed Claisen condensation can be achieved using either a Cys-His-His or Cys-His-Asn catalytic triad, along with an intermediate stabilizing oxyanion hole.

Despite continued progress in the structural characterization of isolated PKS domains, it is only recently that structural techniques have been successfully applied to the study of intact PKS modules [74–77]. These analyses have afforded a step-change in the understanding of PKS enzymology and provided insight into the dynamic nature of chain extension and processing. Most significant amongst these studies has been the elucidation of the structure of module 5 from the pikromycin PKS (PikAIII) using cryo-electron microscopy (cryo-EM) [74, 75]. Uniquely, in this study, a number of structures were determined in a range of states that mimic all stages of the module's catalytic cycle. In addition, the resolution range within which these structures were obtained (7–11 Å) allowed the unambiguous placement of high resolution crystal structures and NMR models of domains homologous to those of PikAIII, yielding a series of pseudo-atomic models describing the process of chain extension in its entirety. In an elegant series of experiments, PikAIII, which possesses a KS-AT-KR-ACP domain architecture, is shown to adopt a distinctive symmetrical arch-like structure comprising a single PikAIII homodimer. Dimerization occurs via the KS domain, which sits at the top of the arch. The AT and KR domains from each monomer form descending struts, which together with the capping KS dimer form a single central reaction chamber accessible by the active sites of each of the KS, AT and KR domains that constitute the module (Fig. 4). The relative positions of each domain deviate significantly from those proposed in structural models of intact PKS modules. The location of the PikAIII module's two ACPs are also resolved and are shown, using targeted chemical modifications designed to mimic a range of reaction cycle intermediates, to occupy distinct locations within the chamber during catalysis (Fig. 4). These include; a *holo*-ACP form bearing a phosphopantetheine arm, where the ACPs are located adjacent to their respective AT or KR domains (state 1); a form in which the ACPs carry a pentaketide intermediate, locating

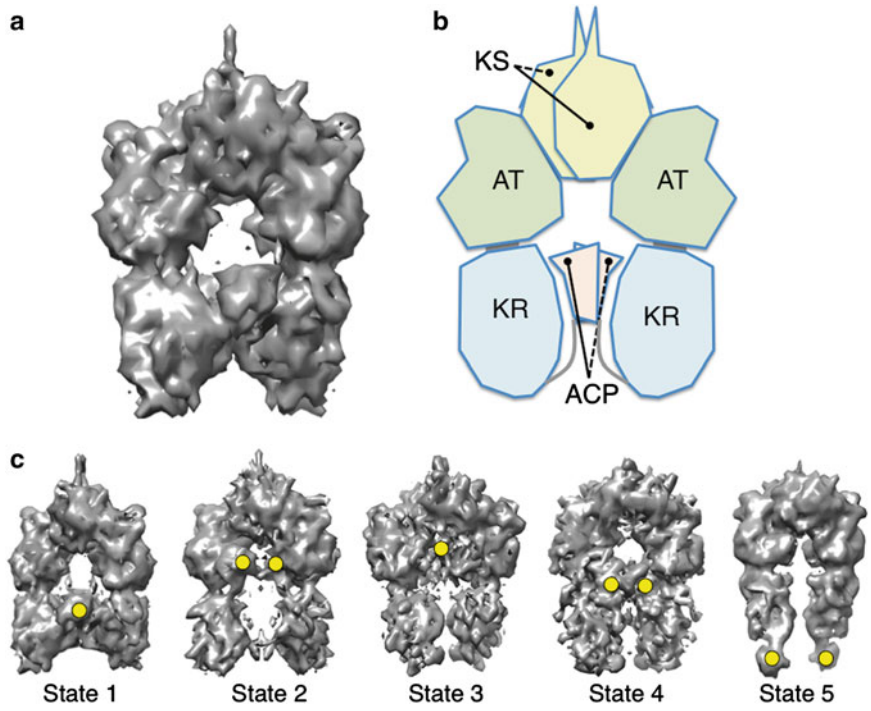

Fig. 4 (**a**) Solid rendering of the cryo-EM map of *holo* PikAIII (EMD-5647). (**b**) Cartoon representation of *holo* PikAIII showing the relative position of each domain within the module. (**c**) Conformational states of PikAIII observed by cryo-EM (state 1 EMD-5647, state 2 EMD-5663, state 3 EMD-5653, state 4 EMD-5664, state 5 EMD-5666). Each state equates to a defined step within the module's catalytic cycle. The location of each PikAIII ACP is indicated by a *yellow circle*. In states 1 and 3 the ACPs occupy overlapping positions at the front and rear of the module

them next to their respective AT's active sites (state 2); a form in which the ACPs are acylated with methylmalonate, and within which they occupy positions below their cognate KSs, primed for active site entry (state 3); a form within which the ACPs carries a β-ketohexaketide, mimicking the system following chain extension, where the ACPs locate next to their respective KRs ready for ketoreduction (state 4); and finally a form mimicking the culmination of the chain extension processes, where the ACPs, carrying β-hydroxyhexaketide groups, are expelled from the PKS reaction chamber, facilitating chain transfer to the down-stream module (state 5). In each of the structures reported the two ACPs occupy equivalent positions on either side of the module dimer, suggesting that they operate in synchronous fashion. This is a likely consequence of space constraints within the reaction chamber and steric hindrance imposed by the downstream KS dimer. Analysis of the reported PikAIII cryo-EM structures also hints at dynamic motions in other domains within the PKS module and is a powerful illustration of the value of analyzing the structures of individual proteins in the context of their interacting partners.

5 *Trans*-AT Synthases: A New Paradigm in Modular PKS Enzymology

A significant recent development in the study of PKSs was the identification of a second class of modular synthases whose domain and module architectures diverge significantly from those of the canonical DEBS like systems [78–80]. This new family of modular PKSs, the *trans*-AT synthases, are notable for their highly mosaic structures that incorporate disparate biosynthetic features within a single megaenzyme complex [81]. *Trans*-AT PKSs have been shown to have an evolutionary lineage distinct from that of the *cis*-AT DEBS like systems [46]. Although *trans*-AT PKSs make use of the same step-wise sequential condensation chemistry employed by *cis*-AT synthases, they exploit a much broader repertoire of functional domains, make use of *trans*-acting elements to modify synthase intermediates, and exhibit module architectures that diverge from the classical KS-AT-ACP paradigm [46, 81, 82]. Here we describe two examples of biosynthetic peculiarities common to *trans*-AT PKSs, but direct readers to more comprehensive reviews of this area [81, 82].

The defining feature of *trans*-AT synthases is the absence of module embedded AT domains throughout the PKS. Substrate loading in these systems is instead provided in *trans*, by free-standing *trans*-acting ATs encoded within the synthase gene cluster [83]. *Trans*-acting ATs are found either as stand-alone enzymes, or as di- or tri-domain fusions proteins partnered with decarboxylases (DCs), ERs, and/or proof-reading acyl-hydrolases (AH; Fig. 5) [78, 84–86]. The role of *trans*-ATs in substrate loading has been demonstrated both in vivo and in vitro, and initial structural and functional characterization of these enzymes suggests that they possess the same general structure and catalytic mechanism as module embedded ATs, but are able to furnish multiple ACPs throughout the PKS complex with extender units [87–90]. The mechanism of *trans*-AT recruitment to each extension module remains to be established, however, it has been suggested that there may be distinct acyltransferase docking domains located within *trans*-AT PKS modules that facilitate this process [91, 92].

Another distinctive biosynthetic feature common to *trans*-AT PKSs is their use of a multi-protein hydroxymethylglutaryl CoA synthase (HCS) enzyme cassette to catalyze the introduction of methyl groups at β-carbon positions within product incorporated acyl extender units [93]. To achieve this acetyl-ACP is condensed with the unreduced β-carbon of the nascent polyketide chain forming a β-hydroxy-β-carboxymethyl intermediate. This intermediate is then subjected to sequential dehydration and decarboxylation to form the β-methyl group (Fig. 5) [89, 94–97]. Structural studies of *trans*-AT PKS di-domain ACPs upon which β-branching takes place have identified a distinct amino acid signature that appears to

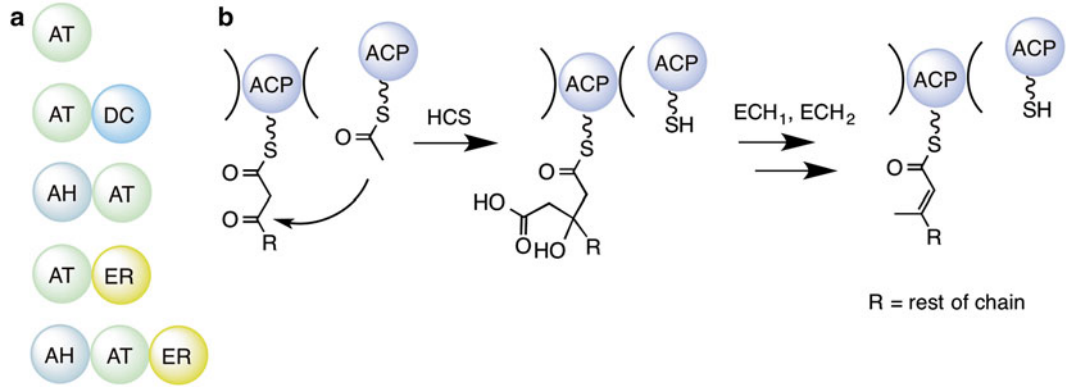

Fig 5 (**a**) Domain organization within polypeptides that house *trans*-acting acyltransferases. (**b**) Scheme for the introduction of β-methyl branches in to polyketides via the HCS cassette route. Only the ACP domains of the modules that support the product chain are shown. *ECH* enoyl CoA hydratase

target the branching machinery to these locations. This discovery raises the intriguing possibility of implementing β-branching chemistry at a range of locations throughout PKSs, by introducing appropriate amino acid signatures into ACPs that do not, in their native state, support branch formation.

6 Reengineering of PKS Assembly Lines

The highly modular architectures of PKSs make these systems attractive targets for reengineering as a mechanism for accessing unnatural natural products with novel or enhanced functionality. This approach gains credence from the notion that polyketide scaffolds have been evolutionarily selected for optimum performance within their producer's environmental niche, and as such they may be of limited utility in wider contexts. The targeted modification of polyketides through manipulation of the cellular machineries responsible for their biosynthesis may therefore yield derivatives of the parent compound with, for example, improved clinical efficacy. The tractability of PKS reengineering was initially demonstrated in DEBS [98]. These investigations resulted in the establishment of a set of rules for the purposeful manipulation of modular synthases [98]. Despite these initial successes, the reengineering of modular PKSs has, however, proved to be much more challenging than was initially anticipated. Though with recent advances in the delineation of synthase enzymology, the development of more robust experimental tools for targeted high-throughput genetic manipulation and DNA synthesis, and improvements in analytical techniques and computation, PKS reengineering appears poised for a renascence.

A number of strategies have been proposed for the successful manipulation of PKSs. These focus on targeted modifications at the module, domain, or amino acid level, or involve the refactoring of precursor biosynthesis or post-PKS tailoring. Approaches targeted at the module level have focused on the deletion, insertion or substitution of intact modules within PKS complexes. This approach has been successfully applied to DEBS, where substitution of the loading module of DEBS 1 with that from the tylosin PKS resulted in a hybrid system selective solely for a propionate starter unit [99]. Similarly, replacement of the DEBS 1 starter module with that from the oleandomycin PKS yielded a hybrid system within which acetate was exclusively incorporated as the starter unit [99]. Although pioneered in DEBS, loading module substitution has been successfully implemented in other systems. For example, replacement of the loading module initiating avermectin biosynthesis, which specifically incorporates isobutyryl-CoA, with that of the cyclohexanecarboxylic (CHC) phoslactomycin PKS loading module from *Streptomycesplatensis*, permitted the biosynthesis of the antiparasitic veterinary medicine doramectin [100]. To achieve this outcome it was also necessary to express the five proteins responsible for the biosynthesis of the CHC-CoA precursor in tandem with the hybrid PKS.

Targeted manipulation at the domain level requires less extensive interference with the PKS, however, the importance of intramodule protein–protein interactions should always be considered. The tolerance of DEBS to domain swapping has been probed extensively [101, 102]. Each extension module within this PKS contains an AT domain selective for (2*S*)-methylmalonyl-CoA. Substitution of these domains with ATs selective for malonyl, ethylmalonyl, or methoxymalonyl extender units has permitted the biosynthesis of regioselectively modified polyketides with the expected chemical composition [102–106]. Substitution of individual domains in DEBS modules 2, 5, and 6, with counterparts from the rapamycin PKS, possessing different substrate specificities and reductive capabilities, as well as the insertion of additional domains from the same system, has been used to produce an extensive range of 6-dEB analogs [102, 107, 108].

To minimize the deleterious effects of PKS reengineering, the targeted mutation of individual residues within modules or domains represents an attractive, less invasive approach. As the number of publically available genome sequences and protein structures increases exponentially, so increases the ability of researchers to make informed site-specific changes that confer or modulate protein function. This may include targeted changes that impact catalytic activity, substrate selectivity, cofactor binding, or stereoselectivity. Examples of the use of this approach in DEBS include deactivation of the enoyl reductase domain of

module 4 permitting the biosynthesis of $\Delta^{6,7}$-anhydroerythromycin C [109], and targeted alteration of the substrate selectivities of extender module AT domains to allow malonyl-CoA and fluoromalonyl-CoA to be accepted as substrates [110]. This approach has been further informed by computational methods. For example the use of quantum mechanics/molecular mechanics (QM/MM) methods was instrumental in the design of a mutagenesis strategy to reengineer DEBS AT6 to accept a non-natural 2-propargylmalonyl extender unit [111]. The polyketide biosynthesized by this modified PKS possesses a synthetically functionalizable handle that can be exploited to generate an even greater number of useful derivatives.

In addition to reengineering strategies that focus explicitly on the PKS itself, complimentary approaches that target the allied starter and extender unit biosynthetic pathways, or focus on post synthase tailoring reactions have been investigated. The former have included the targeted mutation of acyl-CoA synthetases [112, 113], or the replacement of endogenous acyl-CoA pathways with those that generate alternative precursors [114]. For the latter, many successful approaches have involved the repurposing of glycosyltransferases to generate polyketide products with altered glycosylation patterns and consequently more favorable toxicities, solubilities, and bioavailabilities [115, 116].

7 Conclusions

In recent years significant progress has been made in the genetic, chemical, biochemical and structural characterization of modular PKSs and their constituent parts. Despite this, many questions still remain, and without doubt the study of modular synthases will remain a fertile area of research for many years to come. Studies of PKS structure and function are providing unexpected insights into the assembly, operation, and dynamics of these systems, and the frequency with which new PKS gene clusters are being identified and annotated continues to rise exponentially.

Notably, one area of modular PKS research that has progressed more slowly than was hoped is that of synthase reengineering. Methods developed to enable the rational redesign of modular PKSs and their products have proven nontrivial to implement successfully across multiple, often closely related systems, and consequently many approaches that showed initial promise have failed to deliver. That said, new fundamental insights into the enzymology of these systems will undoubtedly expedite the development of the necessary tools and technologies that are required for robust, broadly implementable synthase reengineering.

References

1. Staunton J, Weissman KJ (2001) Polyketide biosynthesis: a millennium review. Nat Prod Rep 18:380–416
2. Li JW-H, Vederas JC (2009) Drug discovery and natural products: end of an era or an endless frontier? Science 325:161–166
3. Demain AL, Vaishnav P (2011) Natural products for cancer chemotherapy. Microb Biotechnol 4:687–699
4. Washington JA, Wilson WR (1985) Erythromycin: a microbial and clinical perspective after 30 years of clinical use (1). Mayo Clin Proc 60:189–203
5. Campbell WC (2012) History of avermectin and ivermectin, with notes on the history of other macrocyclic lactone antiparasitic agents. Curr Pharm Biotechnol 13:9–11
6. Weissman KJ (2009) Introduction to polyketide biosynthesis. Methods Enzymol 459:3–16
7. Challis GL, Hopwood DA (2003) Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. Proc Natl Acad Sci U S A 100(Suppl 2):14555–14561
8. Firn RD, Jones CG (2003) Natural products - a simple model to explain chemical diversity. Nat Prod Rep 20:382
9. Walsh CT (2004) Polyketide and nonribosomal peptide antibiotics: modularity and versatility. Science 199:1805–1811
10. Walsh CT (2008) The chemical versatility of natural-product assembly lines. Acc Chem Res 41:4–10
11. Williams G (2013) Engineering polyketide synthases and nonribosomal peptide synthetases. Curr Opin Struct Biol 23:603–612
12. Wong FT, Khosla C (2012) Combinatorial biosynthesis of polyketides - a perspective. Curr Opin Chem Biol 16:117–123
13. Hertweck C (2015) Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. Trends Biochem Sci 40:189–199
14. Cummings M, Breitling R, Takano E (2014) Steps towards the synthetic biology of polyketide biosynthesis. FEMS Microbiol Lett 351:116–125
15. Goss RJM, Shankar S, Fayad AA (2012) The generation of "unnatural" products: synthetic biology meets synthetic chemistry. Nat Prod Rep 29:870–889
16. Weissman KJ (2004) Polyketide biosynthesis: understanding and exploiting modularity. Philos Trans A Math Phys Eng Sci 362:2671–2690
17. Fischbach MA, Walsh CT (2006) Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. Chem Rev 106:3468–3496
18. Hertweck C (2009) The biosynthetic logic of polyketide diversity. Angew Chem Int Ed Engl 48:4688–4716
19. Xue Y, Sherman DH (2000) Alternative modular polyketide synthase expression controls macrolactone structure. Nature 403: 571–575
20. Wilkinson B et al (2000) Novel octaketide macrolides related to 6-deoxyerythronolide B provide evidence for iterative operation of the erythromycin polyketide synthase. Chem Biol 7:111–117
21. Donadio S, Staver MJ, Mcalpine JB et al (1991) Modular organization of genes required for complex polyketide biosynthesis. Science 252:675–679
22. Callahan B, Thattai M, Shraiman BI (2009) Emergent gene order in a model of modular polyketide synthases. Proc Natl Acad Sci U S A 106:19410–19415
23. Fedorov O, Niesen FH (2012) Kinase inhibitor selectivity profiling using differential scanning fluorimetry. Methods Mol Biol 795:109–118
24. Helfrich EJN, Reiter S, Piel J (2014) Recent advances in genome-based polyketide discovery. Curr Opin Biotechnol 29:107–115
25. Horsman ME, Hari TPA, Boddy CN (2015) Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: logic gate or a victim of fate? Nat Prod Rep (in Press)
26. Lin S, Huang T, Shen B (2012) Tailoring enzymes acting on carrier protein-tethered substrates in natural product biosynthesis. Methods Enzymol 516:321–343
27. Maier T, Leibundgut M, Ban N (2008) The crystal structure of a mammalian fatty acid synthase. Science 321:1315–1323
28. Townsend CA (2014) Aflatoxin and deconstruction of type I, iterative polyketide synthase function. Nat Prod Rep 31:1260–1265
29. Vederas JC (2014) Explorations of fungal biosynthesis of reduced polyketides - a personal viewpoint. Nat Prod Rep 31: 1253–1259
30. Simpson TJ (2014) Fungal polyketide biosynthesis - a personal perspective. Nat Prod Rep 31:1247–1252

31. Hertweck C, Luzhetskyy A, Rebets Y et al (2007) Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. Nat Prod Rep 24:162–190
32. Das A, Khosla C (2009) Biosynthesis of aromatic polyketides in bacteria. Acc Chem Res 42:631–639
33. Yu D, Xu F, Zeng J et al (2012) Type III polyketide synthases in natural product biosynthesis. IUBMB Life 64:285–295
34. Hashimoto M, Nonaka T, Fujii I (2014) Fungal type III polyketide synthases. Nat Prod Rep 31:1306–1317
35. Austin MB, Noel JP (2003) The chalcone synthase superfamily of type III polyketide synthases. Nat Prod Rep 20:79–110
36. Katz L (2009) The DEBS paradigm for type I modular polyketide synthases and beyond, 1st edn. Elsevier Inc., Amsterdam
37. Cortes J, Haydock SF, Roberts GA et al (1990) An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. Nature 348:176–178
38. Weissman KJ (2015) Uncovering the structures of modular polyketide synthases. Nat Prod Rep 32:436–453
39. Kao CM, Katz L, Khosia C (1994) Engineered biosynthesis of a complete macrolactone in a heterologous host. Science 265:509–512
40. Wiesmann KEH et al (1995) Polyketide synthesis in vitro on a modular polyketide synthase. Chem Biol 2:582–589
41. Rowe CJ, Gaisser S, Staunton J et al (1998) Construction of new vectors for high-level expression in actinomycetes. Gene 216: 215–223
42. Pinto A, Wang M, Horsman M et al (2012) 6-Deoxyerythronolide B synthase thioesterase-catalyzed macrocyclization is highly stereoselective. Org Lett 14:2278–2281
43. Siskos AP et al (2005) Molecular basis of Celmer's rules: stereochemistry of catalysis by isolated ketoreductase domains from modular polyketide synthases. Chem Biol 12:1145–1153
44. Weissman KJ et al (1997) The molecular basis of Celmer's rules: the stereochemistry of the condensation step in chain extension on the erythromycin polyketide synthase. Biochemistry 36:13849–13855
45. Keatinge-Clay AT, Stroud RM (2006) The structure of a ketoreductase determines the organization of the beta-carbon processing enzymes of modular polyketide synthases. Structure 14:737–748
46. Nguyen T et al (2008) Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. Nat Biotechnol 26: 225–233
47. Zhang Q et al (2011) Knocking out of tailoring genes eryK and eryG in an industrial erythromycin-producing strain of Saccharopolyspora erythraea leading to overproduction of erythromycin B, C and D at different conversion ratios. Lett Appl Microbiol 52:129–137
48. Khosla C (2009) Structures and mechanisms of polyketide synthases. J Org Chem 74: 6416–6420
49. Keatinge-Clay AT (2012) The structures of type I polyketide synthases. Nat Prod Rep 29:1050–1073
50. Tsai S-CS, Ames BD (2009) Structural enzymology of polyketide synthases, 1st edn. Elsevier Inc., Amsterdam
51. Crosby J, Crump MP (2012) The structural role of the carrier protein—active controller or passive carrier. Nat Prod Rep 29: 1111–1137
52. Tang Y, Kim C, Mathews II et al (2006) The 2.7-Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. Proc Natl Acad Sci U S A 103:11124–11129
53. Tang Y, Chen AY, Kim C-Y et al (2007) Structural and mechanistic analysis of protein interactions in module 3 of the 6-deoxyerythronolide B synthase. Chem Biol 14:931–943
54. Bergeret F et al (2012) Biochemical and structural study of the atypical acyltransferase domain from the mycobacterial polyketide synthase Pks13. J Biol Chem 287:33675–33690
55. Park H, Kevany BM, Dyer DH et al (2014) A polyketide synthase acyltransferase domain structure suggests a recognition mechanism for its hydroxymalonyl-acyl carrier protein substrate. PLoS One 9:e110965
56. Lau J, Cane DE, Khosla C (2000) Substrate specificity of the loading didomain of the erythromycin polyketide synthase. Biochemistry 39:10514–10520
57. Haydock SF et al (1995) Divergent sequence motifs correlated with the substrate specificity of (methyl) malonyl-CoA: acyl carrier protein transacylase domains in modular polyketide synthases. FEBS Lett 374:246–248
58. Yadav G, Gokhale RS, Mohanty D (2003) Computational approach for prediction of domain organization and substrate specificity

of modular polyketide synthases. J Mol Biol 328:335–363
59. Reeves CD et al (2001) Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. Biochemistry 40: 15464–15470
60. Lau J, Fu H, Cane DE et al (1999) Dissecting the role of acyltransferase domains of modular polyketide synthases in the choice and stereochemical fate of extender units. Biochemistry 38:1643–1651
61. Del Vecchio F et al (2003) Active-site residue, domain and module swaps in modular polyketide synthases. J Ind Microbiol Biotechnol 30:489–494
62. Lai JR, Koglin A, Walsh CT (2006) Carrier protein structure and recognition in polyketide and nonribosomal peptide biosynthesis. Biochemistry 45:14869–14879
63. Crump MP et al (1997) Solution structure of the actinorhodin polyketide synthase acyl carrier protein from *Streptomyces coelicolor* A3 (2). Biochemistry 36:6000–6008
64. Findlow SC, Winsor C, Simpson TJ et al (2003) Solution structure and dynamics of oxytetracycline polyketide synthase acyl carrier protein from *Streptomyces rimosus*. Biochemistry 42:8423–8433
65. Li Q, Khosla C, Puglisi JD, Liu CW (2003) Solution structure and backbone dynamics of the holo form of the frenolicin acyl carrier protein. Biochemistry 42:4648–4657
66. Chen AY, Schnarr NA, Kim C et al (2006) Extender unit and acyl carrier protein specificity of ketosynthase domains of the 6-Deoxyerythronolide B synthase. J Am Chem Soc 128:3067–3074
67. Płoskoń E et al (2010) Recognition of intermediate functionality by acyl carrier protein over a complete cycle of fatty acid biosynthesis. Chem Biol 17:776–785
68. Evans SE et al (2009) Probing the interactions of early polyketide intermediates with the actinorhodin ACP from *S. coelicolor* A3(2). J Mol Biol 389:511–528
69. Johnson MNR, Londergan CH, Charkoudian LK (2014) Probing the phosphopantetheine arm conformations of acyl carrier proteins using vibrational spectroscopy. J Am Chem Soc 136:11240–11243
70. Gay DC et al (2014) A close look at a ketosynthase from a trans-acyltransferase modular polyketide synthase. Structure 22:444–451
71. Khosla C, Gokhale RS, Jacobsen JR et al (1999) Tolerance and specificity of polyketide synthases. Annu Rev Biochem 68:219–253
72. Chen AY, Cane DE, Khosla C (2007) Structure-based dissociation of a type I polyketide synthase module. Chem Biol 14:784–792
73. Watanabe K, Wang CCC, Boddy CN et al (2003) Understanding substrate specificity of polyketide synthase modules by generating hybrid multimodular synthases. J Biol Chem 278:42020–42026
74. Dutta S et al (2014) Structure of a modular polyketide synthase. Nature 510:512–517
75. Whicher JR et al (2014) Structural rearrangements of a polyketide synthase module during its catalytic cycle. Nature 510:560–564
76. Edwards AL, Matsui T, Weiss TM et al (2014) Architectures of whole-module and bimodular proteins from the 6-deoxyerythronolide B synthase. J Mol Biol 426:2229–2245
77. Davison J et al (2014) Insights into the function of *trans*-acyl transferase polyketide synthases from the SAXS structure of a complete module. Chem Sci 5:3081
78. Cheng Y, Tang G, Shen B (2003) Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. Proc Natl Acad Sci U S A 100:3149–3154
79. El-sayed AK et al (2003) Characterization of the mupirocin biosynthesis gene cluster from *Pseudomonas fluorescens* NCIMB 10586. Chem Biol 10:419–430
80. Moldenhauer J, Chen X-H, Borriss R et al (2007) Biosynthesis of the antibiotic bacillaene, the product of a giant polyketide synthase complex of the *trans*-AT family. Angew Chem Int Ed Engl 46:8195–8197
81. Piel J (2010) Biosynthesis of polyketides by *trans*-AT polyketide synthases. Nat Prod Rep 27:996–1047
82. Till M, Race PR (2014) Progress challenges and opportunities for the re-engineering of *trans*-AT polyketide synthases. Biotechnol Lett 36:877–888
83. Cheng Y-Q, Coughlin JM, Lim S-K et al (2009) Type I polyketide synthases that require discrete acyltransferases. Methods Enzymol 459:165–186
84. Liu T, Huang Y, Shen B (2009) The bifunctional acyltransferase/decarboxylase LnmK as the missing link for-alkylation in polyketide biosynthesis. J Am Chem Soc 131:6900–6901
85. Thomas CM, Hothersall J, Willis CL et al (2010) Resistance to and synthesis of the antibiotic mupirocin. Nat Rev Microbiol 8:281–289
86. Jensen K et al (2012) Polyketide proofreading by an acyltransferase-like enzyme. Chem Biol 19:329–339

87. Wong FT, Jin X, Mathews II et al (2011) Structure and mechanism of the *trans*-acting acyltransferase from the disorazole synthase. Biochemistry 50:6539–6548
88. Musiol EM et al (2011) Supramolecular templating in kirromycin biosynthesis: the acyltransferase KirCII loads ethylmalonyl-CoA extender onto a specific ACP of the *trans*-AT PKS. Chem Biol 18:438–444
89. Calderone CT, Kowtoniuk WE, Kelleher NL et al (2006) Convergence of isoprene and polyketide biosynthetic machinery: isoprenyl-S-carrier proteins in the pksX pathway of Bacillus subtilis. Proc Natl Acad Sci U S A 103:8977–8982
90. Lopanik NB et al (2010) *In vivo* and *in vitro trans*-acylation by BryP, the putative bryostatin pathway acyltransferase derived from an uncultured marine symbiont. Chem Biol 15:1175–1186
91. Chan YA, Thomas MG (2010) Recognition of (2S)-aminomalonyl-acyl carrier protein (ACP) and (2R)-hydroxymalonyl-ACP by acyltransferases in zwittermicin A biosynthesis. Biochemistry 49:3667–3677
92. Musiol EM, Weber T (2012) Discrete acyltransferases involved in polyketide biosynthesis. Med Chem Commun 3:871
93. Calderone CT (2008) Isoprenoid-like alkylations in polyketide biosynthesis. Nat Prod Rep 25:845–853
94. Calderone CT, Iwig DF, Dorrestein PC et al (2007) Incorporation of nonmethyl branches by isoprenoid-like logic: multiple beta-alkylation events in the biosynthesis of myxovirescin A1. Chem Biol 14:835–846
95. Gu L et al (2006) Metabolic coupling of dehydration and decarboxylation in the curacin A pathway: functional identification of a mechanistically diverse enzyme pair. J Am Chem Soc 128:9014–9015
96. Simunovic V, Müller R (2007) 3-hydroxy-3-methylglutaryl-CoA-like synthases direct the formation of methyl and ethyl side groups in the biosynthesis of the antibiotic myxovirescin A. Chembiochem 8(5):497–500
97. Simunovic V, Müller R (2007) Mutational analysis of the myxovirescin biosynthetic gene cluster reveals novel insights into the functional elaboration of polyketide backbones. Chembiochem 8:1273–1280
98. McDaniel R, Ebert-Khosla S, Hopwood DA et al (1995) Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. Nature 375:549–554
99. Long PF et al (2002) Engineering specificity of starter unit selection by the erythromycin-producing polyketide synthase. Mol Microbiol 43:1215–1225
100. Wang J-B, Pan H-X, Tang G-L (2011) Production of doramectin by rational engineering of the avermectin biosynthetic pathway. Bioorg Med Chem Lett 21:3320–3323
101. Rodriguez E, McDaniel R (2001) Combinatorial biosynthesis of antimicrobials and other natural products. Curr Opin Microbiol 4:526–534
102. McDaniel R et al (1999) Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel 'unnatural' natural products. Proc Natl Acad Sci U S A 96:1846–1851
103. Oliynykl M, Brown MJB, Cort J et al (1996) A hybrid modular polyketide synthase obtained by domain swapping. Chem Biol 3:833–839
104. Ruan X et al (1997) Acyltransferase domain substitutions in erythromycin polyketide synthase yield novel erythromycin derivatives. J Bacteriol 179:6416–6425
105. Stassi DL et al (1998) Ethyl-substituted erythromycin derivatives produced by directed metabolic engineering. Proc Natl Acad Sci U S A 95:7305–7309
106. Kato Y et al (2002) Functional expression of genes involved in the biosynthesis of the novel polyketide chain extension unit, methoxymalonyl-acyl carrier protein, and engineered biosynthesis of 2-desmethyl-2-methoxy-6-deoxyerythronolide B. J Am Chem Soc 124:5268–5269
107. McDaniel R et al (1997) Gain-of-function mutagenesis of a modular polyketide synthase. J Am Chem Soc 119:4309–4310
108. Bedford D, Jacobsen JR, Luo G et al (1996) A functional chimeric modular polyketide synthase generated via domain replacement. Chem Biol 3:827–831
109. Donadio S, Mcalpine JB, Sheldont PJ et al (1993) An erythromycin analog produced by reprogramming of polyketide synthesis. Proc Natl Acad Sci U S A 90:7119–7123
110. Walker MC et al (2013) Expanding the fluorine chemistry of living systems using engineered polyketide synthase pathways. Science 341:1089–1094
111. Sundermann U et al (2013) Enzyme-directed mutasynthesis: a combined experimental and theoretical approach to substrate recognition of a polyketide synthase. ACS Chem Biol 8:443–450

112. Koryakina I, Williams GJ (2011) Mutant malonyl-CoA synthetases with altered specificity for polyketide synthase extender unit generation. Chembiochem 12:2289–2293
113. Koryakina I et al (2013) Poly specific *trans*-acyltransferase machinery revealed via engineered acyl-CoA synthetases. ACS Chem Biol 8:200–208
114. Lechner A et al (2013) Designed biosynthesis of 36-methyl-FK506 by polyketide precursor pathway engineering. ACS Synth Biol 2: 379–383
115. Rix U, Fischer C, Remsing LL et al (2002) Modification of post-PKS tailoring steps through combinatorial biosynthesis. Nat Prod Rep 19:542–580
116. Olano C, Méndez C, Salas JA (2010) Post-PKS tailoring steps in natural product-producing actinomycetes from the perspective of combinatorial biosynthesis. Nat Prod Rep 27:571