Published in final edited form as:
Nat Rev Microbiol. 2010 December ; 8(12): 879–889. doi:10.1038/nrmicro2465.

New insights into the formation of fungal aromatic polyketides

Jason M. Crawford[*] and **Craig A. Townsend**[†]

[*]Department of Biological Chemistry & Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA

[†]Department of Chemistry, The Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA

Abstract

Fungal aromatic polyketides constitute a large family of bioactive natural products and are synthesized by the non-reducing group of iterative polyketide synthases (NR-PKSs). Their diverse structures arise from selective enzymatic modifications of reactive enzyme-bound poly-β-keto intermediates. How iterative PKSs control starter unit selection, polyketide chain initiation and elongation, intermediate folding and cyclization, selective redox or modification reactions during assembly, and product release are central mechanistic questions underlying iterative catalysis. This review highlights recent insights into these questions, with a particular focus on the biosynthetic programming of fungal aromatic polyketides, and draws comparisons to the allied biosynthetic processes in bacteria.

Polyketide natural products possess great commercial value for their diverse biological activities, which result from their structural variation implemented during synthesis by polyketide synthases (PKSs). The fungal polyketide statins, such as lovastatin (**1**) for example, are among the most successful cholesterol-lowering agents on the market (Fig.1a). Bacterial polyketides such as the tetracycline (**4**) antibiotics or the daunomycin (**5**)-inspired doxorubicin anticancer agents have been prescribed for decades. On the other hand, fungal polyketide mycotoxins, including aflatoxin (**10**), fumonisin (**2**), zearalenone (**3**), and the 6-methylsalicylic acid (**6**)-derived patulin (**7**), profoundly and negatively affect agriculture. Like the bacterial-derived tetracyclines or daunomycins, the fungal features that reflect similarities in their biosynthetic origins (Fig. 1b). Understanding how polyketide structural variation is generated is key to identifying and predicting new products encoded in the vast number of emerging sequenced microbial genomes and to developing new bioactive polyketides through rational pathway or enzyme engineering.

Fungal and bacterial secondary metabolite biosynthetic genes are often clustered with regulatory and resistance genes, whose proximity accelerate the assignment of putative functions to individual enzymes in a given pathway. Fungal polyketide structural diversity is further enhanced by nearby auxiliary or post-PKS tailoring genes in the clusters. Genome sequence information has revealed and continues to reveal that microorganisms, especially filamentous fungi, house many more biosynthetic pathways than previously realized [1], although their tight regulatory control often restricts expression levels needed to isolate and characterize their encoded products [2]. While such regulation is not surprising, as it would conserve metabolic resources during times the pathway is not required, it begs the question: what unknown signals are required to activate expression of the unknown pathways? The

Correspondence to C.A.T. ctownsend@jhu.edu.

Competing interests statement The authors declare no competing financial interests.

regulation of some of these pathways is beginning to become clear and has been reviewed elsewhere [2-5]. The manipulation and up-regulation of unknown secondary metabolite pathways in the lab through environmental stimuli, nutritional optimization, heterologous expression, or genetics/epigenetics will be critical for future small molecule discovery efforts.

PKSs utilize chemistry common to fatty acid biosynthesis [6, 7] (Fig. 2). Short-chain fatty acids, activated as their CoA esters, are used as starter units (*e.g.* acetyl-CoA) and extender units (*e.g.* malonyl-CoA) and are homologated to form linear chains, such as the C_{16} fatty acid palmitate, through effective decarboxylative Claisen condensations (involving malonyl-ACP hydration and loss of bicarbonate [8]). In fatty acid synthases (FASs) and PKSs, the acyl transacylase (AT) loads in turn the starter and extender units onto an acyl-carrier protein (ACP), which carries the intermediates through the catalytic cycle. The ketosynthase (KS) accepts the acyl unit from the acyl-ACP via transthioesterification and subsequently condenses malonyl-CoA units that are shuttled by the ACP onto the starter unit to add a ketide unit one catalytic step at a time. In fatty acid biosynthesis, the resulting β-ketone is fully reduced by stepwise ketoreductase (KR), dehydrase (DH) and enoyl reductase (ER) reactions to give an inert hydrocarbon chain. In polyketide biosynthesis, full reduction does not always occur at the β-carbon of the extending chain and the PKSs selectively use these reductive/dehydrative processing steps to introduce one of the four possible chemical states at the β-carbon (ketone, hydroxyl, alkene, saturated methylene) during polyketide assembly. In some cases, PKS systems recruit additional auxiliary enzymes or domains to perform specialized reactions. All of these biosynthetic steps can be programmatically varied to produce a vast array of structural outcomes. Likewise, the absence of reduction steps leads to poly-β-keto chains that can be folded and cyclized in different ways to produce the many different scaffolds found in aromatic polyketides.

PKSs have been the subject of extensive review and are currently divided into three general classes (type I [6, 9], type II [10-12], and type III [13-15]) [16, 17], although the exchange and evolution of genetic information among organisms has led to mixed classes [18, 19]. Fungal PKSs are generally multidomain systems (type I) that elongate their polyketide products iteratively. Unlike the bacterial type I PKSs, which are often modular megasynthases that produce products such as erythromycin in a predictable assembly-line fashion, the iterative fungal type I systems reuse a single set of active sites through multiple catalytic cycles. Iterative type I PKSs are further subdivided based on the level of reductive processing [non-reducing (NR-), partially reducing (PR-) [20], or highly reducing (HR-) [21]] during chain assembly [22, 23]. Whereas fungal aromatic PKSs are of the non-reducing type I class (NR-PKSs), bacterial aromatic PKSs are discretely expressed free-standing proteins (type II) that associate to carry out the required chemistry. Regardless of whether the chemistry occurs in a single multidomain protein or in a multi-enzyme complex, aromatic PKSs must all contend with reactive poly-β-keto intermediates, controlling their chain length, folding and cyclization, as well as product release.

Over the past five years our understanding of fungal aromatic polyketide biosynthesis has advanced greatly. Here, we review the latest progress, including our increased understanding of starter unit selection, chain-length control and cyclization specificity. These advances should provide the foundation for many discoveries ahead as more complex PKSs begin to reveal their biosynthetic programming codes.

Enzyme deconstruction

Multidomain proteins are common in nature and are most often thought to be the result of evolutionary fusion events of smaller ancestral genes. Before the advent of PCR-based

molecular approaches, partial proteolysis of multidomain proteins was the primary method to obtain the individual catalytic domains for further study [24]. These classical approaches were of intense interest for the study of animal FASs [25] — close relatives of the fungal PKSs. Domains from animal FASs could be reacted with their free-standing bacterial FAS type II counterparts to assess enzymatic activity and, consequently, to further support their evolutionary relationships [26]. To that end, the Townsend lab developed a bioinformatics-based method, the UMA algorithm, which combines sequence similarity, predicted secondary structural conservation and local hydrophobicity to predict the location of the linker regions between catalytic domains in large multifunctional proteins such as PKSs [27]. Application to the NR-PKS group predicted six well-defined domains (Fig. 1b). Four were readily identified by sequence conservation to known proteins (the KS domain, the malonyl-CoA:ACP transacylase (MAT) domain, the ACP domain, and the terminal thioesterase/ Claisen cyclase (TE/CLC) domain), but the other two domains identified, a long *N*-terminal domain and a centrally located domain, lacked significant sequence conservation to any known enzyme. The central domain was termed the product template (PT) domain based on its location adjacent to the ACP and a hypothesized requirement for chemical stabilization of reactive poly-β-keto intermediates [27]. This *in silico* approach greatly accelerated efforts to clone and isolate individual domains or combinations of properly folded domains from various fungal polyketide systems for biochemical deconstruction experiments – domain dissection and *in vitro* reconstitution studies [28-34]. Others have used the sequence conservation of the type II bacterial enzymes or other structurally characterized proteins to successfully select sites at or near predicted domain boundaries for further biochemical interrogation [35-37].

Biochemical basis for starter-unit selection

Since the 1955 discovery of the incorporation of radiolabeled acetate into the now archetypal fungal polyketide 6-methylsalicylic acid in *Penicillium griseofuluvum* [38], many groups have employed isotopic labeling studies with the aim of understanding polyketide biosynthesis in both eukaryotes and bacteria. Beginning in the 1950s, ^{14}C-labeled acetate or malonate was administered to fungal or bacterial cultures, and the sites of radioisotope incorporation in the resulting polyketides were carefully tracked by chemical degradation [29]. Radiotracer methods were later replaced by NMR techniques utilizing various stable isotopic labels (^{13}C, ^{18}O, ^{15}N, ^{2}H) and, consequently, a large body of isotopic labeling data has accrued over the years. The central dogma arising from more than 60 years of these studies posits that fungal PKSs typically select acetyl-CoA as the starter unit to initiate biosynthesis rather than malonyl-CoA, which has come to be known as the acetyl "starter unit effect" [30,31]. Current efforts have focused on correlating these labeling trends with the underlying biosynthetic enzymes in bacteria and fungi.

PKSs can accept starter units that are more advanced than acetyl, particularly when they are downstream of other PKSs or associated with other biosynthetic systems, such as non-ribosomal peptide synthetases (NRPSs), FASs and terpene biosynthetic enzymes. For example, the biosynthesis of the fungal polyketide aflatoxin in select *Aspergillus* species [39-41] involves an NR-PKS, PksA, [42] in a FAS-PKS hybrid complex [43, 44]. The NR-PKS accepts a C_6-fatty acid starter [45-47] from a dedicated fungal FAS [31, 48, 49], further homologates seven malonyl-CoA extender units and controls specific folding and cyclization to produce the aflatoxin B_1 (**10**) precursor, norsolorinic acid anthrone (**8**). The anthrone is oxidized to the anthraquinone norsolorinic acid (**9**) by an anthrone oxidase [50]. The aflatoxins are potent environmental pro-carcinogens that commonly contaminate agricultural foods, and chronic ingestion and subsequent oxidation/activation in the liver and kidney can lead to hepatocellular carcinoma [51].

Functional interrogation of PksA led to the elucidation of a biochemical rationale for the starter unit effect in fungal NR-PKSs [28]. The long *N*-terminal domain in NR-PKSs showed minimal sequence conservation to enzymes of known function but shared predicted secondary structural similarity to MAT domains. Kinetics and mutation experiments revealed that this domain was responsible for accepting the hexanoyl starter unit from the associated FAS to initiate polyketide biosynthesis [28]. This unusual domain utilized thioester chemistry, as opposed to the oxyester chemistry found in typical AT domains, and so was termed the starter unit:ACP transacylase (SAT) domain. It was proposed that when associated proteins, such as the dedicated FAS proteins in aflatoxin biosynthesis, were absent, the SAT domain could mediate acetyl incorporation, thus providing a biochemical basis for the classical acetyl starter unit effect [28]. Indeed, the SAT domains involved in the biosynthesis of the naphthopyrone YWA1 (**12**), tetrahydroxynaphthalene (THN, **14**), cercosporin (**15**) and bikaverin (**16**) are selective for acetyl-CoA rather than malonyl-CoA or other longer-chain acyl-CoA substrates [30] (Fig. 1b).

The *Gibberella fujikuroi* bikaverin NR-PKS [52] was produced in *Escherichia coli*, and the purified protein was capable of utilizing a malonyl-CoA, rather than an acetyl-CoA starter unit, *in vitro* to produce the bikaverin precursor (**20**) [53]. The catalytic Cys residue of the SAT domain of this NR-PKS was not present in the annotated sequence. However, re-sequencing of the SAT region led to the discovery of a frame-shift and an alternatively spliced form that translated through the active-site Cys [30]. The encoded domain was indeed selective for acetyl-CoA [30]. Recent *in vivo/in vitro* comparisons with the lovastatin PKS show that, like other fungal PKSs, acetyl-CoA is preferentially accepted by this PKS *in vivo* while malonyl-CoA can be accepted *in vitro* [54]. The conservation of acetyl-CoA starter unit specificity in fungal PKSs is probably due to the metabolic benefits of sampling this readily available TCA cycle precursor as opposed to converting it to malonyl-CoA.

In addition to short- and medium-chain fatty acids, SAT domains have since been implicated in accepting other substrates, such as the advanced resorcylic acid lactone intermediates [29, 55-58], which was recently demonstrated biochemically [59]. Given that NR-PKS genes are ubiquitous in filamentous fungi and multiple homologues are often encoded by a single genome, substrate selectivity at the SAT domain provides a means for providing alternative starter units to begin the structural variation found in aromatic polyketides [29].

Control of reactive poly-β-keto intermediates

One approach to deconstruct PKS function has been to dissect the full-length enzyme into individual catalytic domains or small groups of adjacent domains for *in vitro* reconstitution experiments. Initial experiments aimed at understanding how NR-PKSs control poly-β-keto intermediate formation and subsequent processing were also carried out using PksA. By analyzing both the intermediates directly tethered to the ACP domain and the products released by PksA, the global division of labour among the individual domains in the iterative catalytic cycle was defined [32]. Domain combinations lacking the ACP did not assemble detectable released products, affirming the critical role of this domain in carrying intermediates during polyketide extension and processing. When the ACP domain was present but lacked the necessary tailoring domains, a stalled linear poly-β-keto intermediate (**17**) was detected and subsequent MS/MS fragmentation analysis confirmed the proposed structure, providing the first direct experimental evidence for these long-proposed enzyme-bound intermediates (Fig. 3).

Because the SAT [28] and MAT [35] domains load the starter and extender units and the ACP domain carries the elongated intermediates through the catalytic cycle, it was proposed that the KS domain is mostly, if not entirely, responsible for controlling the chain length of the

poly-β-keto intermediate [32]. Supporting this notion, the minimal NR-PKS components (KS, MAT, and ACP) from the bikaverin (**16**) PKS system (Pks4) primarily produce the correct chain length intermediate but lack the ability to control cyclization chemistry [37]. Moreover, non-cognate ACP domains [35] and the MAT proteins from FASs [12] can often qualitatively complement activity of aromatic polyketide biosynthesis. The solution structure of the PksA ACP domain exhibited minimal interaction with a tethered non-polar acyl-thioester substrate [60], further narrowing chain-length control to the KS domain. Non-natural starter units with different lengths alter the number of ketide extensions to largely reflect the natural chain length, which is likely dictated by the available KS cavity volume [53]. In the bacterial type II PKSs chain length is controlled by KS (KS_{α}) and chain-length factor (CLF, also known as KS_{β}) [61] (type I NR-PKS systems lack CLF).

When an active PT domain was added to the loading, carrier and KS domains, there was a decrease in the linear substrate (**17**) attached to the ACP domain, a reduction in unknown cyclization products, and a marked increase in the correct cyclization of rings A and B (**18**) (Fig. 3), where norpyrone (**11**) was detected as the major product [32]. The PksA PT domain is responsible for accepting the fully extended linear polyketide intermediate and catalyzing its stepwise cyclization and aromatization produces the penultimate polyketide intermediate (**18**). Lastly, when the full complement of all six domains was present, the correct final cyclization outcome was observed. The TE/CLC domain catalyzed C-C bond closure, as opposed to spontaneous O-C bond derailment, leading to the release of norsolorinic acid anthrone (**8**), which is subsequently oxidized to norsolorinic acid (**9**) on the path to aflatoxins [32].

It was evident in these experiments that domain fusion enhances catalytic efficiency, which undoubtedly constitutes an evolutionary basis for the formation of iterative type I systems in eukaryotes [32]. Heterologous co-expression of individual fungal domains has not led to readily observable chemistry, mandating direct covalent linkage or engineered linker peptides for efficient production *in vivo* [37, 62]. Nevertheless, dissected domains that are reconstituted *in vitro* maintain the ability to faithfully interact *in trans* to enable polyketide formation, like the iterative type II bacterial systems. Tang and co-workers expressed the NR-PKS gene responsible for biosynthesis of bikaverin (**16**) in *E. coli* for *in vitro* biochemical analysis [53] and showed that non-cognate bacterial proteins (a KR and an aromatase/cyclase (Aro/Cyc), which is known to catalyze cyclization in bacterial aromatic polyketide biosynthesis) and the dissected fungal TE/CLC domain can interact *in trans* with the fungal PKS [63]. Most intriguingly, they showed that a free-standing bacterial type II KR that normally interacts with its cognate type II PKS could access the fungal PKS-generated poly-β-keto intermediate and carry out regioselective reduction [63]. This suggests that despite evolutionary separation, substrate recognition and proper protein-protein docking can still take place.

The relative ease of using non-cognate enzymes to carry out alternative reactions not only opens the door for engineering new product scaffolds, but also illustrates the ease of gene cluster diversification in the evolution of bioactive fungal aromatic metabolites. Unlike the exquisite specificity of enzymes involved in primary metabolism that is emblematic of the ‘one substrate, one enzyme, one product’ rule for metabolic fidelity, many of the enzymes involved in secondary metabolism show greater tolerance where one enzyme can often accept multiple substrates to produce multiple products. Or, a common intermediate can diverge to multiple products, such as a series of antibiotics, that give the host a competitive advantage [64]. Alternatively, such enzymes can be only a few mutations away from promiscuity or alternative functions. Directed evolution is now a powerful approach to engineer entirely new enzymatic specificities. In keeping with the promiscuity of individual biosynthetic enzymes, entire secondary metabolite pathways are, in fact, diversity-

oriented [65]. The ability of alternative proteins or domains to act *in trans* clearly illuminates the evolutionary malleability of iterative fungal metabolite gene clusters.

Cyclization specificity of aromatic PKSs

The stable isotope labeling patterns collected for polyketide metabolites not only supported the fungal starter unit effect, but also defined polyketide folding patterns – that is, how the PKSs regioselectively control polyketide geometry prior to cyclization and aromatization [66, 67]. In aromatic polyketides, multiple folding modes can lead to the same product. For example, the polyketide chrysophanol, which is produced by both fungi and bacteria, is derived from a common poly-β-keto intermediate but it is cyclized differently in fungi (F-mode) than in bacteria (S-mode) [68]. Recent structural and biochemical studies have revealed the underlying enzymatic bases for folding and cyclization.

The iterative PKSs in fungi are architecturally related to mammalian FASs. Indeed, the crystal structure of a mammalian FAS illustrates that this multidomain enzyme harbours an inactive methyl transferase domain that is closely related to domains found in reducing PKSs, indicating that this structural relic has been maintained from its ancestral origins infungi [69]. Recent high-resolution crystal structures of the dimeric PT monodomain from PksA with linear or bicyclic substrate mimics bound showed that, despite its low sequence conservation to known enzymes, the domain adopted a modified double hot dog fold (Fig. 4a) [33]. The double hot dog fold [70] has been observed in bacterial dehydratases/ isomerases [71, 72] and mammalian hydratases [73], in addition to the pseudodimeric dehydratase domains in mammalian FASs [69]. Interestingly, in the PT domains of NR-PKSs this fold controls one of the key programmed steps in aromatic polyketide biosynthesis, cyclization specificity [32, 33]. Native PAGE mobility experiments on individual domains or combinations of domains from PksA further showed that PT- or KS-domain-containing constructs exist as dimers, whereas the loading and tailoring domains appear as monomers. Taken together, the observed associative interactions, the high-resolution structural comparisons, and the phylogenetic [74] and evolutionary relationships between fungal PKSs and animal FASs indicate that these enzymes are intriguingly similar, yet they catalyze quite different synthetic outcomes.

The modified double hot dog fold of the PT domain of PksA dictates the cyclization outcome [33]. The substrate binding regions of this domain can be divided into three sections: the phosphopantetheine localization channel, which binds the extended phosphopantetheine arm of the ACP domain that delivers the linear intermediate into the interior binding pocket; the central cyclization chamber, which promotes cyclization of the delivered substrate and accommodates the cyclization products; and the hexyl-binding region, which lies deep in the protein core and binds the fatty acid starter unit (the PT hexyl-binding channel is believed to be closed in acetyl-initiated systems). The linear arrangement of binding sites from the periphery to the protein core is clearly aligned to hold the poly-β-keto intermediate in a near linear fashion, centrally kinked into the cyclization chamber to initiate F-mode C4-C9 and C2-C11 cyclizations (Fig. 4b). Intriguingly, the catalytic chemistry of the PT domain is influenced by distinct “wet” and “dry” sides of the cyclization chamber. Because the keto form, as opposed to the enol form, is favoured in water [75, 76], the wet side of the cyclization chamber ensures that the substrate is in the electrophilic keto form for cyclization, and the resulting cyclized product is displaced into the hydrophobic, or dry, reaction chamber by a thermodynamic driving force afforded by C–C bond formation, loss of water and aromatization [33]. The counterpart bacterial Aro/Cyc protein from the antibiotic tetracenomycin pathway adopts a topologically related helix-grip fold that is thought to bind its poly-β-keto substrate in a distinct hairpin conformation and drive S-mode C9-C14 and C7-C16 cyclizations [77]. Aro/Cyc proteins are thought to have evolved from ancient ligand-

binding proteins [78], whereas PT domains are believed to have evolved from DH domains in ancient reducing PKSs [33]. The different evolutionary paths and structural topologies observed between eukaryotic PT domains and bacterial Aro/Cyc proteins have resulted in the altered chemistry seen across the kingdoms, and it is likely that the wet/dry active site arrangement in PT domains is common in enzymatic carbonyl-based chemistry.

It has been proposed that chain length control might be facilitated by the PT domain based on phylogenetic analysis [79], although the apparent relationship between PT clade structure and chain length is more likely to be correlated to cavity volume [22] and cyclization mode to accommodate the mature linear intermediates [32]. Comparison of active site residues between the known C4-C9 and C2-C7 first ring cyclizing PT domains showed a delineation in sequence homology between the two cyclization modes [33]. Subsequent phylogenetic studies illustrated that PT domains fall into five groups based on cyclization specificity and the number of rings in the final product [62]. The putative group of C6-C11 cyclizing PT domains was verified by incorporating the asperthecin [80] or viridicatumotoxin [81] PT domains (C6-C11) into the minimal PKS (KS, MAT, and ACP domains) from the bikaverin biosynthetic pathway (which utilizes a C2-C7 PT domain) [62]. In these studies, chain length again was set prior to cyclization and the regiospecificity was altered by the introduced PT domain. The clade structure of the individual PT domains [62], in addition to common active site residues [33] in known cyclization modes (C4-C9 [32], C2-C7 [37], or C6-C11[62] first ring cyclizing PT domains), provides a first approach to predicting the cyclization specificity of unknown NR-PKSs.

Claisen condensation governs product release

PKSs have acquired a number of different strategies to terminate the biosynthetic program and release their final products [82]. The canonical means for PKSs and FASs is TE mediated product release. TEs are capable of catalyzing not only simple hydrolysis, macrolactonization, and macrolactamization, but also Claisen/Dieckman C-C bond-forming condensation reactions in NR-PKSs. The first report of a thioesterase/Claisen Cyclase (TE/CLC) domain was obtained from the *Aspergillus nidulans* NR-PKS responsible for biosynthesis of the heptaketide naphthopyrone spore pigment YWA1 (**11**). Asexual spores, or conidia, produce pigments that contribute to the strength, rigidity, and impermeability of their cell walls. By using heterologous expression of the YWA1 NR-PKS in *Aspergillus oryzae* and subsequent product analysis, Ebizuka and co-workers identified WA as the protein responsible for biosynthesis of YWA1 (**12**) (Fig. 3) [83]. WA was originally thought to produce citreoisocoumarins (O-C cyclization product as opposed to correct TE/CLC C-C cyclization product) based on an unintentional deletion of the C-terminal TE domain [84]. This fortuitous deletion, later coupled with site-directed mutagenesis in the TE domain, led to the identification of a new class of TE/CLC domains that catalyze C-C bond cyclization as opposed to the simple hydrolysis or macrocyclizations previously reported in TE domains [85]. TE/CLC activity has since been identified in other NR-PKSs responsible for the synthesis of tetrahydroxynaphthalene (**13**) [86], bikaverin (**15**) [63], and norsolorinic acid anthrone (**7**) [32, 34] (Fig. 3). The YWA1 pigment (**11**) can serve as a substrate for melanin biosynthesis through a hydrolytic chain-shortening reaction to tetrahydroxynaphthalene (**13**) [87]. The protein Ayg1p catalyzes this retro-Claisen reaction.

Since the early 1960s, fungi have been known to produce melanins. In addition to contributing to cell wall integrity [88], melanins are known to act as virulence factors in human and plant pathogens [89, 90]. They have been reported to have antioxidant and antiphagocytic properties, contribute to cAMP-dependent signalling, and even block antimicrobial entry. Even though they have such important physiological roles, the mechanistic and biochemical bases for the biosynthesis of polyketide monomer units of

melanin have only emerged in the past few years. Identification of the NR-PKS responsible for the production of the monomer unit in *Colletotrichum lagenarium* was confirmed by heterologous expression of the *C. lagenarium* PKS1 in *Aspergillus oryzae* and product analysis [91]. In this expression system, tetrahydroxynaphthalene (THN, **14**), a precursor to the monomer unit found in dihydroxynaphthalene (DHN)-based melanin, was the main product observed (50%) [86]. Other minor products were also present, including acetyl-THN (**13**, 15%). Partial purification of PKS1 and feeding studies with radiolabeled acetyl- and malonyl-CoA precursors showed that radioactivity from malonyl-CoA, but not acetyl-CoA, remained in the symmetrical THN product, and the authors proposed that acetyl-CoA was not used as a starter unit [92].

Simpson and co-workers discovered similar labeling trends for scytalone biosynthesis in the early 1980s. Scytalone is derived from THN, but the authors proposed an alternative biosynthesis based on the general acetyl starter unit effect trends that were known at the time [93]. They proposed that either they were not able to detect the acetyl starter under the conditions of their experiment or the product could be derived from a hexaketide precursor (**19**) that loses its acetyl starter in the formation of the pentaketide THN product. Mutagenesis of the C-terminal TE/CLC led to the production of the hexaketide isocoumarin (>95% of the produced products were O-C cyclization derailments) as opposed to the pentaketide THN (C-C cyclization product) [86], indicating that the NR-PKS was likely a hexaketide synthase; the authors proposed that the TE/CLC domain controls chain length. Because doubts remained over its biosynthetic origins, Townsend and co-workers analyzed the substrate preferences for the SAT domain in PKS1 and found that it was selective for acetyl-CoA but not malonyl-CoA [30]. Based on the new study and previous results, a unified biosynthetic path was proposed, in which acetyl-CoA is incorporated into the hexaketide acetyl-THN (**13**), and an Ayg1p-mediated hydrolytic chain shortening reaction could occur, as had been found for YWA1 (Fig 3). Indeed, a knockout of the Ayg1p homologue in *Wangiella dermatitidis* later confirmed that a PKS1 homologue produced acetyl-THN [94], and that the enzyme was capable of catalyzing the retro-Claisen reaction *in vitro* [95]. The same phenomenon has yet to be confirmed experimentally for the *C. lagenarium* PKS1 system. It is also important to note that because malonyl-ACP can be decarboxylated, it is possible that malonyl-CoA can partially contribute to starter unit initiation in fungal polyketides. The effects are likely to vary based on growth conditions and the kinetic parameters of the individual loading domains in PKS homologues.

The PksA TE/CLC monodomain crystal structure was solved to 1.7 Å resolution, and the structure displayed the typical α/β-hydrolase fold in the catalytic closed form [34] (Fig. 5b). Why would nature evolve to use an enzyme fold that is widespread and well known to drive O-C bond formation [96] when that is exactly the reaction that must be prevented in the Claisen-type (C-C bond formation) product release mechanism? Indeed, the TE/CLC domain has maintained its ability to catalyze hydrolysis (O-C bond formation), which likely serves an editing function in the absence of the correct Claisen substrate [27, 34]. Removal of the TE/CLC domain or individual mutation of the catalytic triad residues (analogous to the serine protease triad Ser, His and Asp [97]) destroy both its hydrolytic and C-C bond-forming abilities, leading to spontaneous product derailment as the O-C cyclized norpyrone (**11**) [32, 34]. The key to the native reaction lies in the lid loop or gateway to the catalytic core of the enzyme. The major difference among the known TE domain structures is in the lid loop region. Docking of the ACP domain with its tethered intermediate to the TE/CLC domain presumably induces a conformational change to open the lid for substrate entry. Supporting this notion, NMR titration experiments suggest this analogous region of the TE domain interacts with its peptidyl carrier protein in enterobactin biosynthesis [98, 99].

Based on the catalytic closed-form structure of the TE/CLC domain, the phosphopantetheine arm in the ACP domain must leave the pocket before the lid loop can close (Fig. 5a) [34]. The substrate side chain swings into the space vacated by the departing arm, and when the lid closes, the substrate side chain is locked into the correct position adjacent to the catalytic histidine residue to facilitate C-C bond formation. The hydrophobic binding chamber and lid impose strict substrate positioning, which is the structural basis for regiospecific product release by C–C bond formation and program termination in fungal aromatic polyketide biosynthesis.

Conclusions

Until recently, the fungal iterative PKSs eluded detailed biochemical investigations owing to their large size and the lack of microbiological tools to produce these enzymes in experimentally useful quantities [16]. The development of gene expression methods and the evolutionary relatedness of NR-PKSs to the well-studied animal FASs have greatly accelerated our understanding of iterative enzyme catalysis. The domains of type I iterative PKSs or the free-standing auxiliary proteins that interact with fungal PKSs must bind and accommodate substrates of increasing length, harbour compatible protein-protein interaction surfaces, and maintain favourable kinetics to channel the elongating product to a specific structural outcome. The ACP domain efficiently docks with its client domains in these large synthases to covalently shuttle the growing intermediates through a large number of biosynthetic steps at high effective substrate concentrations. The multidomain type I systems enjoy the same kinetic benefits from covalent interactions (linkers) formed by evolutionary domain fusion events [32, 100]. The linker regions appear not to interfere with programming as *in vitro* activity remains when the linkers are truncated or replaced by His_{6x}-fusion tags.

Because multidomain systems have evolved by fusion events, it is fortunate experimentally, but not surprising, to see wild-type catalytic chemistry faithfully executed by the dissected, free-standing domains. PKSs illustrate the fact that the fundamental chemical potential held by a handful of protein folds can be exploited in different ways to take on new functions. The fundamental chemistry and protein fold in PT domains, for example, is similar to DH domains in animal FASs, but the double hot dog fold has been modified to expand the repertoire of carbonyl chemistry to allow control of the cyclization specificity by fungal aromatic PKSs.

Individual domains or *trans*-acting auxiliary enzymes appear to be qualitatively interchangeable, and the future manipulation of iterative enzymes will be a balancing act between enzyme tolerance and specificity. Even though fungal NR-PKSs have solved the daunting chemical task of controlling the synthesis and cyclization of highly reactive poly-β-keto intermediates, they execute an otherwise relatively simple biosynthetic program in that no reductive processing steps occur during assembly. That domains or accessory enzymes can act *in trans* to access the growing intermediates and selectively modulate activity provides a basis for understanding more complex iterative PKS programming. The PR and HR iterative PKSs in fungi are capable of carrying out selective redox or methylation reactions during assembly to increase the structural diversity of iterative type I PKS products. Presumably, similar forces are at work here to modulate the synthetic programming of these enzymes and to direct the biochemical outcome. In the pioneering work of Hutchinson, Vederas, and co-workers, for example, it was found that a *trans*-acting ER auxiliary protein was necessary to complement the HR-PKS lovastatin nonaketide synthase (LNKS) to properly produce the lovastatin precursor [101]. When the ER protein was lacking, LNKS exhibited aberrant activity *in vivo*. The free-standing ER was subsequently shown *in vitro* to hold a gatekeeper function to modulate LNKS activity [54]. It is likely that such interactions are more commonly required to process reduced polyketides, and auxiliary

enzymes will be required to obtain on-path products in future heterologous expression or *in vitro* biochemical experiments [102]. Similar biochemical approaches will undoubtedly provide more insights into these more complex iterative PKSs.

Filamentous fungi produce many secondary metabolites with biomedical potential, and fungal genome projects indicate that they can produce many more [1, 103-105]. Single hosts have been reported to contain as many as 30 PKSs located in secondary metabolic gene clusters. Of these, there are often multiple homologues in the NR-PKSs group, and they are not redundant. Consequently, there are great opportunities for combining genomics with natural product discovery in filamentous fungi. Understanding the mechanistic basis for the programming of starter unit selection, chain-length control, cyclization specificity and chain termination not only provides a foundation for uncovering the many encoded aromatic polyketides emerging from fungal genome sequencing efforts, but will also be critical for the rational engineering of alternative bioactive products. The genius of the evolutionary flexibility of NR-PKSs is that small changes in enzyme structure in turn affect the programming that is the basis for versatile biosynthetic library construction, as exemplified by the large aromatic polyketide family of natural products.

Acknowledgments

Our work on fungal aromatic polyketide biosynthesis is supported by the US National Institutes of Health (grant ES001670 to C.A.T.). J.M.C. is a Damon Runyon fellow supported by the Damon Runyon Cancer Research Foundation (DRG-2002-09).

References

1. Keller NP, Turner G, Bennett JW. Fungal secondary metabolism - from biochemistry to genomics. Nat Rev Microbiol. 2005; 3:937–47. [PubMed: 16322742]
2. Hoffmeister D, Keller NP. Natural products of filamentous fungi: enzymes, genes, and their regulation. Nat Prod Rep. 2007; 24:393–416. [PubMed: 17390002]
3. Schumann J, Hertweck C. Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes. J Biotechnol. 2006
4. Chiang YM, Lee KH, Sanchez JF, Keller NP, Wang CC. Unlocking fungal cryptic natural products. Nat Prod Commun. 2009; 4:1505–1510. [PubMed: 19967983]
5. Scherlach K, Hertweck C. Triggering cryptic natural product biosynthesis in microorganisms. Org Biomol Chem. 2009; 7:1753–60. [PubMed: 19590766]
6. Smith S, Tsai SC. The type I fatty acid and polyketide synthases: a tale of two megasynthases. Nat Prod Rep. 2007; 24:1041–72. [PubMed: 17898897]
7. Cronan JE, Thomas J. Bacterial fatty acid synthesis and its relationships with polyketide synthetic pathways. Methods Enzymol. 2009; 459:395–433. [PubMed: 19362649]
8. Witkowski A, Joshi AK, Smith S. Mechanism of the beta-ketoacyl synthase reaction catalyzed by the animal fatty acid synthase. Biochemistry. 2002; 41:10877–87. [PubMed: 12196027]
9. Staunton J, Weissman KJ. Polyketide biosynthesis: a millennium review. Nat Prod Rep. 2001; 18:380–416. [PubMed: 11548049]
10. Hertweck C, Luzhetskyy A, Rebets Y, Bechthold A. Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. Nat Prod Rep. 2007; 24:162–90. [PubMed: 17268612]
11. Zhang W, Tang Y. In vitro Analysis of Type II Polyketide Synthase. Methods Enzymol. 2009; 459:367–393. [PubMed: 19362648]
12. Das A, Khosla C. Biosynthesis of aromatic polyketides in bacteria. Acc Chem Res. 2009; 42:631–9. [PubMed: 19292437]
13. Austin MB, Noel JP. The chalcone synthase superfamily of type III polyketide synthases. Nat Prod Rep. 2003; 20:79–110. [PubMed: 12636085]
14. Austin MB, et al. Crystal structure of a bacterial type III polyketide synthase and enzymatic control of reactive polyketide intermediates. J Biol Chem. 2004; 279:45162–74. [PubMed: 15265863]

15. Funa N, Awakawa T, Horinouchi S. Pentaketide resorcylic acid synthesis by type III polyketide synthase from *Neurospora crassa*. J Biol Chem. 2007; 282:14476–81. [PubMed: 17374612]
16. Fujii I. Heterologous expression systems for polyketide synthases. Nat Prod Rep. 2009; 26:155–69. [PubMed: 19177221]
17. Weissman KJ. Introduction to Polyketide Biosynthesis. Methods Enzymol. 2009; 459:3–16. [PubMed: 19362633]
18. Shen B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. Curr Opin Chem Biol. 2003; 7:285–95. [PubMed: 12714063]
19. Muller R. Don't classify polyketide synthases. Chem Biol. 2004; 11:4–6. [PubMed: 15112984]
20. Bingle LE, Simpson TJ, Lazarus CM. Ketosynthase domain probes identify two subclasses of fungal polyketide synthase genes. Fungal Genet Biol. 1999; 26:209–23. [PubMed: 10361035]
21. Nicholson TP, et al. Design and utility of oligonucleotide gene probes for fungal polyketide synthases. Chem Biol. 2001; 8:157–78. [PubMed: 11251290]
22. Hertweck C. The biosynthetic logic of polyketide diversity. Angew Chem Int Ed Engl. 2009; 48:4688–716. [PubMed: 19514004]
23. Cox RJ, Simpson TJ. Fungal Type I Polyketide Synthases. Methods Enzymol. 2009; 459:49–78. [PubMed: 19362635]
24. Klenow H, Henningsen I. Selective elimination of the exonuclease activity of the deoxyribonucleic acid polymerase from *Escherichia coli* B by limited proteolysis. Proc Natl Acad Sci U S A. 1970; 65:168–175. [PubMed: 4905667]
25. Tsukamoto Y, Wong H, Mattick JS, Wakil SJ. The architecture of the animal fatty acid synthetase complex. IV. Mapping of active centers and model for the mechanism of action. J Biol Chem. 1983; 258:15312–22. [PubMed: 6654914]
26. Plate CA, Joshi VC, Wakil SJ. Studies on the mechanism of fatty acid synthesis. XXIV. The acetyl- and malonyltransacylase activities of pigeon liver fatty acid synthetase. J Biol Chem. 1970; 245:2868–75. [PubMed: 4912535]
27. Udwary D, Merski M, Townsend C. A Method for Prediction of the Locations of Linker Regions within Large Multifunctional Proteins, and Application to a Type I Polyketide Synthase. J Mol Biol. 2002; 323:585–598. [PubMed: 12381311]
28. Crawford JM, Dancy BCR, Hill EA, Udwary DW, Townsend CA. Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase. Proc Natl Acad Sci U S A. 2006; 103:16728–16733. [PubMed: 17071746] First publication demonstrating the enzymatic basis for starter unit initiation in fungal aromatic PKSs.
29. Crawford JM, Vagstad AL, Ehrlich KC, Townsend CA. Starter unit specificity directs genome mining of polyketide synthase pathways in fungi. Bioorg Chem. 2008; 36:16–22. [PubMed: 18215412]
30. Crawford JM, Vagstad AL, Whitworth KP, Ehrlich KC, Townsend CA. Synthetic Strategy of Nonreducing Iterative Polyketide Synthases and the Origin of the Classical "Starter-Unit Effect". Chembiochem. 2008; 9:1019–1023. [PubMed: 18338425]
31. Crawford JM, Vagstad AL, Ehrlich KC, Udwary DW, Townsend CA. Acyl-carrier protein-phosphopantetheinyltransferase partnerships in fungal fatty acid synthases. Chembiochem. 2008; 9:1559–63. [PubMed: 18551496]
32. Crawford JM, et al. Deconstruction of iterative multidomain polyketide synthase function. Science. 2008; 320:243–6. [PubMed: 18403714] Describes the dissection and enzymatic reconstitution of an iterative PKS to define domain activities and provides the first direct experimental evidence for hypothetical poly-β-keto intermediates.
33. Crawford JM, et al. Structural basis for biosynthetic programming of fungal aromatic polyketide cyclization. Nature. 2009; 461:1139–43. [PubMed: 19847268] Atomic level structures of the first PT domain reveal how poly-β-keto intermediates are folded and cyclized in fungi.
34. Korman TP, et al. Structure and function of an iterative polyketide synthase thioesterase domain catalyzing Claisen cyclization in aflatoxin biosynthesis. Proc Natl Acad Sci U S A. 2010; 107:6246–51. [PubMed: 20332208] The first crystal structure of a Claisen cyclase-type thioesterase provides a model for C-C bond cyclization in polyketide chain termination.

35. Ma Y, et al. Catalytic relationships between type I and type II iterative polyketide synthases: The *Aspergillus parasiticus* norsolorinic acid synthase. Chembiochem. 2006; 7:1951–8. [PubMed: 17086560]
36. Ma SM, Tang Y. Biochemical characterization of the minimal polyketide synthase domains in the lovastatin nonaketide synthase LovB. Febs J. 2007; 274:2854–64. [PubMed: 17466016]
37. Zhang W, Li Y, Tang Y. Engineered biosynthesis of bacterial aromatic polyketides in *Escherichia coli*. Proc Natl Acad Sci U S A. 2008; 105:20683–8. [PubMed: 19075227] A heterologous expression system employing engineered protein linkers was developed in *E. coli* to produce new polyketide derivatives from fungal and bacterial enzyme components.
38. Birch AJ, Massy-Westropp RA, Moye CJ. Studies in relation to biosynthesis. VII. 2-Hydroxy-6-methylbenzoic acid in *Penicillium griseofulvum* Dierckx. Aus. J. Chem. 1955; 8:539–544.
39. Townsend, C.; Minto, R. Comprehensive Natural Products. Elsevier, Oxford; Sankawa: 1999. Biosynthesis of Aflatoxin; p. 443-471.
40. Yu J, et al. Clustered pathway genes in aflatoxin biosynthesis. Appl Environ Microbiol. 2004; 70:1253–62. [PubMed: 15006741]
41. Yabe K, Nakajima H. Enzyme reactions and genes in aflatoxin biosynthesis. Appl Microbiol Biotechnol. 2004; 64:745–55. [PubMed: 15022028]
42. Chang PK, Cary JW, Yu J, Bhatnagar D, Cleveland TE. The *Aspergillus parasiticus* polyketide synthase gene pksA, a homolog of *Aspergillus nidulans* wA, is required for aflatoxin B1 biosynthesis. Mol Gen Genet. 1995; 248:270–7. [PubMed: 7565588]
43. Watanabe CM, Wilson D, Linz JE, Townsend CA. Demonstration of the catalytic roles and evidence for the physical association of type I fatty acid synthases and a polyketide synthase in the biosynthesis of aflatoxin B1. Chem Biol. 1996; 3:463–9. [PubMed: 8807876]
44. Watanabe C, Townsend C. Initial Characterization of a Type 1 Fatty Acid Synthase and Polyketide Synthase Multienzyme complex NorS in the Biosynthesis of Aflatoxin B1. Chem Biol. 2002; 9:981–988. [PubMed: 12323372]
45. Townsend CA, Christensen S, Trautwein K. Hexanoate as a Starter Unit in Polyketide Biosynthesis. J Am Chem Soc. 1984; 106:3868–3869.
46. McKeown DSJ, McNicholas C, Simpson TJ, Willett NJ. Biosynthesis of norsolorinic acid and averufin: substrate specificity of norsolorinic acid synthase. Chem Commun (Camb). 1996:301–302.
47. Brobst SW, Townsend CA. The potential role of fatty acid initiation in the biosynthesis of the fungal aromatic polyketide aflatoxin B1. Can J Chem. 1994; 72:200–207.
48. Brown DW, Adams TH, Keller NP. *Aspergillus* has distinct fatty acid synthases for primary and secondary metabolism. Proc Natl Acad Sci U S A. 1996; 93:14873–7. [PubMed: 8962148]
49. Hitchman T, et al. Hexanoate Synthase, a Specialized Type I Fatty Acid Synthase in Aflatoxin B1 Biosynthesis. Bioorg Chem. 2001; 29:293–307. [PubMed: 16256699]
50. Ehrlich KC, Li P, Scharfenstein L, Chang PK. HypC, the anthrone oxidase involved in aflatoxin biosynthesis. Appl Environ Microbiol. 2010; 76:3374–7. [PubMed: 20348292]
51. Hussain SP, Schwank J, Staib F, Wang XW, Harris CC. TP53 mutations and hepatocellular carcinoma: insights into the etiology and pathogenesis of liver cancer. Oncogene. 2007; 26:2166–76. [PubMed: 17401425]
52. Linnemannstons P, et al. The polyketide synthase gene pks4 from *Gibberella fujikuroi* encodes a key enzyme in the biosynthesis of the red pigment bikaverin. Fungal Genet Biol. 2002; 37:134–48. [PubMed: 12409099]
53. Ma SM, et al. Enzymatic synthesis of aromatic polyketides using PKS4 from *Gibberella fujikuroi*. J Am Chem Soc. 2007; 129:10642–3. [PubMed: 17696354]
54. Ma SM, et al. Complete reconstitution of a highly reducing iterative polyketide synthase. Science. 2009; 326:589–92. [PubMed: 19900898] Biochemical determination of PKS gate keeper functions for auxiliary enzymes.
55. Gaffoor I, Trail F. Characterization of two polyketide synthase genes involved in zearalenone biosynthesis in *Gibberella zeae*. Appl Environ Microbiol. 2006; 72:1793–9. [PubMed: 16517624]

56. Zhou H, Zhan J, Watanabe K, Xie X, Tang Y. A polyketide macrolactone synthase from the filamentous fungus *Gibberella zeae*. Proc Natl Acad Sci U S A. 2008; 105:6249–54. [PubMed: 18427109]
57. Reeves CD, Hu Z, Reid R, Kealey JT. Genes for the biosynthesis of the fungal polyketides hypothemycin from *Hypomyces subiculosus* and radicicol from *Pochonia chlamydosporia*. Appl Environ Microbiol. 2008; 74:5121–9. [PubMed: 18567690]
58. Wang S, et al. Functional characterization of the biosynthesis of radicicol, an Hsp90 inhibitor resorcylic acid lactone from *Chaetomium chiversii*. Chem Biol. 2008; 15:1328–38. [PubMed: 19101477]
59. Zhou H, et al. Enzymatic synthesis of resorcylic acid lactones by cooperation of fungal iterative polyketide synthases involved in hypothemycin biosynthesis. J Am Chem Soc. 2010; 132:4530–1. [PubMed: 20222707]
60. Wattana-amorn P, et al. Solution structure of an acyl carrier protein domain from a fungal type I polyketide synthase. Biochemistry. 2010; 49:2186–93. [PubMed: 20136099]
61. Keatinge-Clay AT, Maltby DA, Medzihradszky KF, Khosla C, Stroud RM. An antibiotic factory caught in action. Nat Struct Mol Biol. 2004; 11:888–893. [PubMed: 15286722]
62. Li Y, Xu W, Tang Y. Classification, prediction and verification of the regioselectivity of fungal polyketide synthase product template domains. J Biol Chem. 2010 A phylogenetic approach to predicting polyketide cyclization in unknown NR-PKSs.
63. Ma SM, et al. Redirecting the cyclization steps of fungal polyketide synthase. J Am Chem Soc. 2008; 130:38–9. [PubMed: 18072779] First biochemical demonstration of bacterial enzymes intercepting fungal poly-β-keto intermediates and carrying out regiospecific reactions.
64. Bodner MJ, Phelan RM, Freeman MF, Li R, Townsend CA. Non-heme iron oxygenases generate natural structural diversity in carbapenem antibiotics. J Am Chem Soc. 2010; 132:12–3. [PubMed: 20017478]
65. Fischbach MA, Clardy J. One pathway, many products. Nat Chem Biol. 2007; 3:353–5. [PubMed: 17576415]
66. Thomas R. A biosynthetic classification of fungal and streptomycete fused-ring aromatic polyketides. Chembiochem. 2001; 2:612–27. [PubMed: 11828498]
67. Zhou H, Li Y, Tang Y. Cyclization of aromatic polyketides from bacteria and fungi. Nat Prod Rep. 2010; 27:839–68. [PubMed: 20358042]
68. Bringmann G, Gulder TA, Hamm A, Goodfellow M, Fiedler HP. Multiple convergence in polyketide biosynthesis: a third folding mode to the anthraquinone chrysophanol. Chem Commun (Camb). 2009:6810–2. [PubMed: 19885487]
69. Maier T, Leibundgut M, Ban N. The crystal structure of a mammalian fatty acid synthase. Science. 2008; 321:1315–22. [PubMed: 18772430] Atomic level resolution structures of the mammalian FAS, a close relative of fungal PKSs.
70. Dillon SC, Bateman A. The Hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. BMC Bioinformatics. 2004; 5:109. [PubMed: 15307895]
71. Leesong M, Henderson BS, Gillig JR, Schwab JM, Smith JL. Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site. Structure. 1996; 4:253–64. [PubMed: 8805534]
72. Keatinge-Clay A. Crystal structure of the erythromycin polyketide synthase dehydratase. J Mol Biol. 2008; 384:941–53. [PubMed: 18952099]
73. Koski KM, Haapalainen AM, Hiltunen JK, Glumoff T. Crystal structure of 2-enoyl-CoA hydratase 2 from human peroxisomal multifunctional enzyme type 2. J Mol Biol. 2005; 345:1157–69. [PubMed: 15644212]
74. Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon BG. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. Proc Natl Acad Sci U S A. 2003; 100:15670–5. [PubMed: 14676319]
75. Mills SG, Beak P. Solvent effects on keto-enol equilibria: tests of quantitative models. J. Org. Chem. 1985; 50:1216–1224.
76. Hibbert F, Emsley P. Hydrogen bonding and chemical reactivity. Adv. Phys. Org. Chem. 1991; 26:255–379.

77. Ames BD, et al. Crystal structure and functional analysis of tetracenomycin ARO/CYC: implications for cyclization specificity of aromatic polyketides. Proc Natl Acad Sci U S A. 2008; 105:5349–54. [PubMed: 18388203]
78. Iyer LM, Koonin EV, Aravind L. Adaptations of the helix-grip fold for ligand binding and catalysis in the START domain superfamily. Proteins. 2001; 43:134–44. [PubMed: 11276083]
79. Cox RJ. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. Org Biomol Chem. 2007; 5:2010–26. [PubMed: 17581644]
80. Szewczyk E, et al. Identification and characterization of the asperthecin gene cluster of *Aspergillus nidulans*. Appl Environ Microbiol. 2008; 74:7607–12. [PubMed: 18978088]
81. Chooi YH, Cacho R, Tang Y. Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. Chem Biol. 2010; 17:483–94. [PubMed: 20534346]
82. Du L, Lou L. PKS and NRPS release mechanisms. Nat Prod Rep. 2010; 27:255–78. [PubMed: 20111804]
83. Watanabe A, et al. Re-identification of *Aspergillus nidulans* wA Gene to Code for a Polyketide Synthase of Naphthopyrone. Tet Lett. 1999; 40:91–94.
84. Watanabe A, et al. Product identification of polyketide synthase coded by *Aspergillus nidulans* wA gene. Tet Lett. 1998; 39:7733–7736.
85. Fujii I, Watanabe A, Sankawa U, Ebizuka Y. Identification of Claisen cyclase domain in fungal polyketide synthase WA, a naphthopyrone synthase of *Aspergillus nidulans*. Chem Biol. 2001; 8:189–197. [PubMed: 11251292] First biochemical description of Claisen cyclase activity for NR-PKS TE domains.
86. Watanabe A, Ebizuka Y. Unprecedented mechanism of chain length determination in fungal aromatic polyketide synthases. Chem Biol. 2004; 11:1101–6. [PubMed: 15324811]
87. Tsai HF, et al. Pentaketide melanin biosynthesis in *Aspergillus fumigatus* requires chain-length shortening of a heptaketide precursor. J Biol Chem. 2001; 276:29292–8. [PubMed: 11350964]
88. Pihet M, et al. Melanin is an essential component for the integrity of the cell wall of *Aspergillus fumigatus* conidia. BMC Microbiol. 2009; 9:177. [PubMed: 19703288]
89. Langfelder K, Streibel M, Jahn B, Haase G, Brakhage AA. Biosynthesis of fungal melanins and their importance for human pathogenic fungi. Fungal Genet Biol. 2003; 38:143–58. [PubMed: 12620252]
90. Liu GY, Nizet V. Color me bad: microbial pigments as virulence factors. Trends Microbiol. 2009; 17:406–13. [PubMed: 19726196]
91. Fujii I, et al. Heterologous expression and product identification of *Colletotrichum lagenarium* polyketide synthase encoded by the PKS1 gene involved in melanin biosynthesis. Biosci Biotechnol Biochem. 1999; 63:1445–52. [PubMed: 10501004]
92. Fujii I, et al. Enzymatic synthesis of 1,3,6,8-tetrahydroxynaphthalene solely from malonyl coenzyme A by a fungal iterative type I polyketide synthase PKS1. Biochemistry. 2000; 39:8853–8. [PubMed: 10913297]
93. Bardshiri E, Simpson TJ. ^{13}C and ^{2}H Labelling studies on the biosynthesis of scytalone in *Phialaphora lagerbergii*. Tetrahedron. 1983; 39:3539–3542.
94. Wheeler MH, et al. New biosynthetic step in the melanin pathway of *Wangiella (Exophiala) dermatitidis*: evidence for 2-acetyl-1,3,6,8-Tetrahydroxynaphthalene as a novel precursor. Eukaryot Cell. 2008; 7:1699–711. [PubMed: 18676950]
95. Fujii I. Functional analysis of fungal polyketide biosynthesis genes. J Antibiot (Tokyo). 2010
96. Nardini M, Dijkstra BW. Alpha/beta hydrolase fold enzymes: the family keeps growing. Curr Opin Struct Biol. 1999; 9:732–7. [PubMed: 10607665]
97. Carter P, Wells J. Dissecting the catalytic triad of a serine protease. Nature. 1988; 332:564–568. [PubMed: 3282170]
98. Koglin A, et al. Structural basis for the selectivity of the external thioesterase of the surfactin synthetase. Nature. 2008; 454:907–11. [PubMed: 18704089]
99. Frueh DP, et al. Dynamic thiolation-thioesterase structure of a non-ribosomal peptide synthetase. Nature. 2008; 454:903–6. [PubMed: 18704088]

100. Tran L, Tosin M, Spencer JB, Leadlay PF, Weissman KJ. Covalent linkage mediates communication between ACP and TE domains in modular polyketide synthases. Chembiochem. 2008; 9:905–15. [PubMed: 18348128]

101. Kennedy J, et al. Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. Science. 1999; 284:1368–72. [PubMed: 10334994]

102. Belecki K, Crawford JM, Townsend CA. Production of octaketide polyenes by the calicheamicin polyketide synthase CalE8: implications for the biosynthesis of enediyne core structures. J Am Chem Soc. 2009; 131:12564–6. [PubMed: 19689130]

103. Machida M, et al. Genome sequencing and analysis of *Aspergillus oryzae*. Nature. 2005; 438:1157–61. [PubMed: 16372010]

104. Nierman WC, et al. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature. 2005; 438:1151–6. [PubMed: 16372009]

105. Galagan JE, et al. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature. 2005; 438:1105–15. [PubMed: 16372000]

Box 1

PKSs

Type I PKSs are multidomain enzyme systems. Individual domains can be used once in a modular “assembly line” fashion (modular type I) or can be reused multiple times during product assembly in an “iterative” fashion (iterative type I). Fungal PKSs are often iterative type I systems. Type II PKSs represent free-standing proteins that associate to carryout chemistry. Bacterial aromatic PKSs are type II iterative systems. Type III PKSs are single proteins found in plants, bacteria, and fungi. They resemble the KS domain/protein of type I and type II systems, but they utilize acyl-CoA substrates directly in an iterative fashion without the requirement of any auxiliary enzymes.

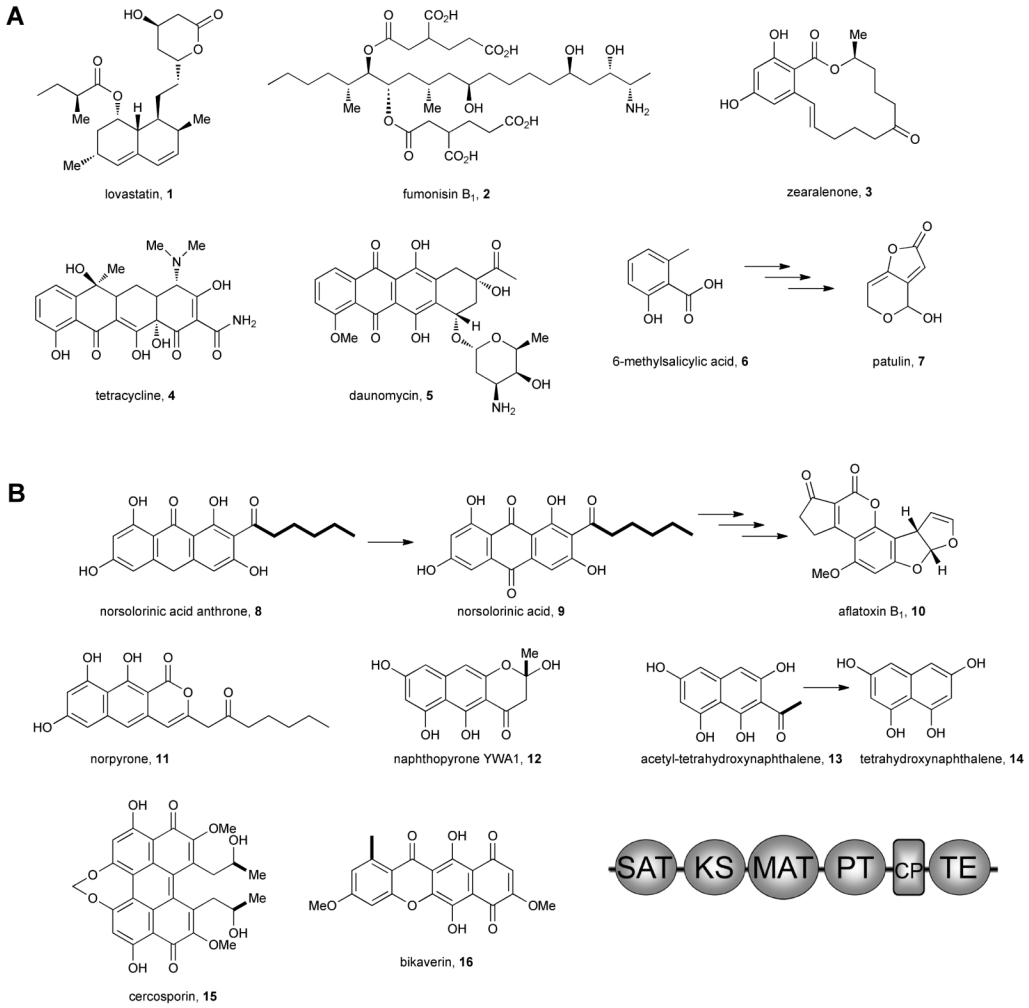

Fig. 1. Representative polyketides

a) Bioactive polyketides from bacterial and fungal sources. b) Examples of fungal aromatic polyketides. The domain architecture for the NR-PKS group, represented as balls (domains) on a string (linkers), is shown: starter unit:ACP transacylase (SAT), ketosynthase (KS), malonyl-CoA:ACP transacylase, product template (PT), acyl-carrier protein (ACP or CP), and thioesterase (TE)/Claisen cyclase (CLC).

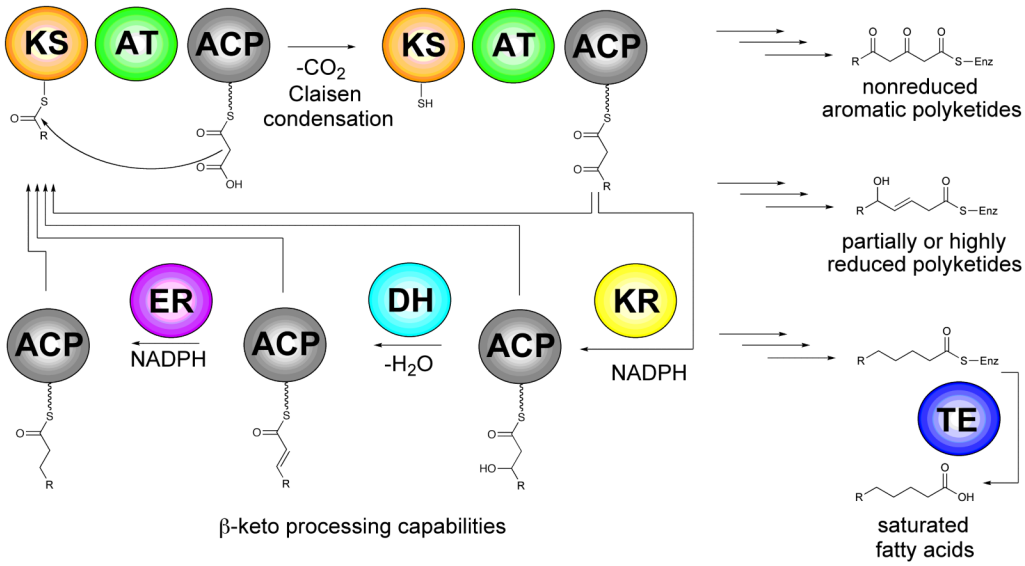

Fig. 2. Basic fatty acid and polyketide processing elements

The growing substrates are passed back and forth between the KS and ACP via thioester transfers (Enz = enzyme: KS or ACP). The acyl-transacylase (AT) loads the ACP from starter or extender acyl-CoA substrates. Product complexity arises from the number of cycles (chain length control) and selective reductive processing that can occur during product assembly by a ketoreductase (KR), dehydrase (DH), or enoyl-reductase (ER). The TE is often used to terminate the program to release the final product. Type II systems are free-standing proteins, as depicted in the image, that associate to form a multi-enzyme complex, whereas type I systems are fused into single multidomain proteins.

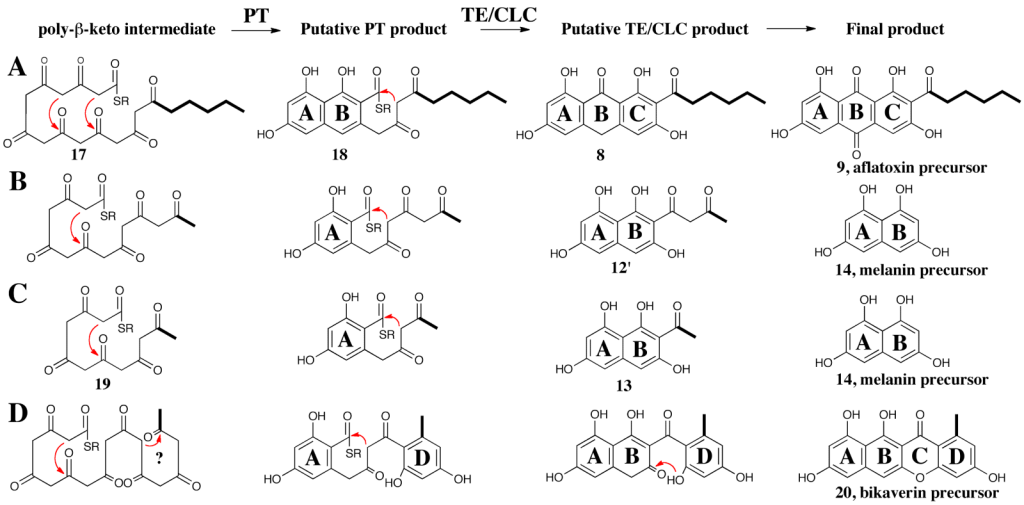

Fig. 3. Model fungal aromatic polyketides and their biosynthetic outcomes

Cyclization of the ACP-tethered linear poly-β-keto intermediates (R group) by the PT domain and TE/CLC domain is shown for the production of norsolorinic acid anthrone (**8**, **A**), YWA1 napthopyrone (**12**, **B**), tetrahydroxynaphthalene (**14**, **C**), and the bikaverin precursor (**20**, **D**).

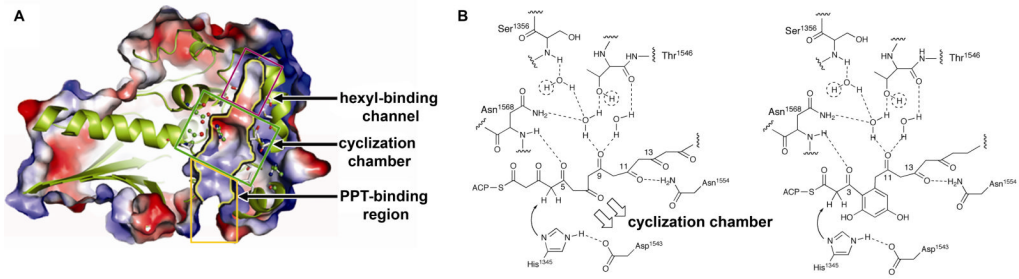

Fig. 4. PT-mediated cyclization of aromatic polyketides

A) The monomer unit of the high-resolution dimeric PT crystal structure is shown, illustrating the linear arrangement of substrate binding sites. **B)** Proposed mechanism for cyclization/aromatization in the PksA PT domain.

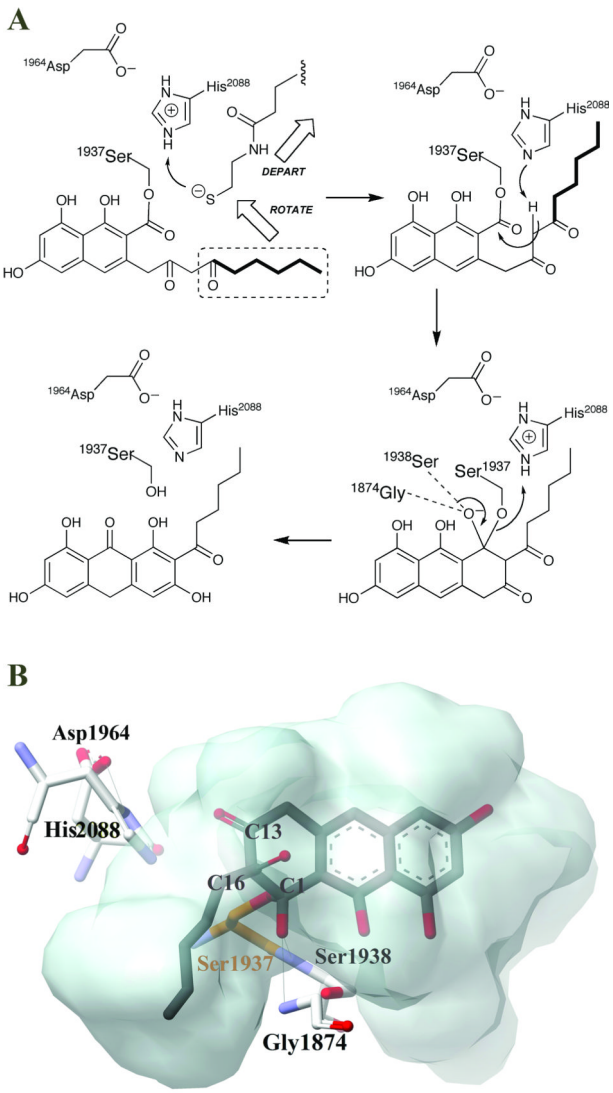

Fig. 5. TE/CLC-mediated chain termination

A) Proposed mechanism for the final Claisen/Dieckmann condensation. The reaction sequence is shown beginning from the covalent TE/CLC-oxyester intermediate, which is analogous to the classical serine protease intermediate. The ACP phosphopantetheine arm must depart before the substrate fatty acyl side chain can swing into correct position for C-C cyclization as opposed to O-C product derailment. **B)** The modeled tetrahedral intermediate is shown filling the active site cavity volume in the TE/CLC cyclization chamber.