JOC The Journal of Organic Chemistry

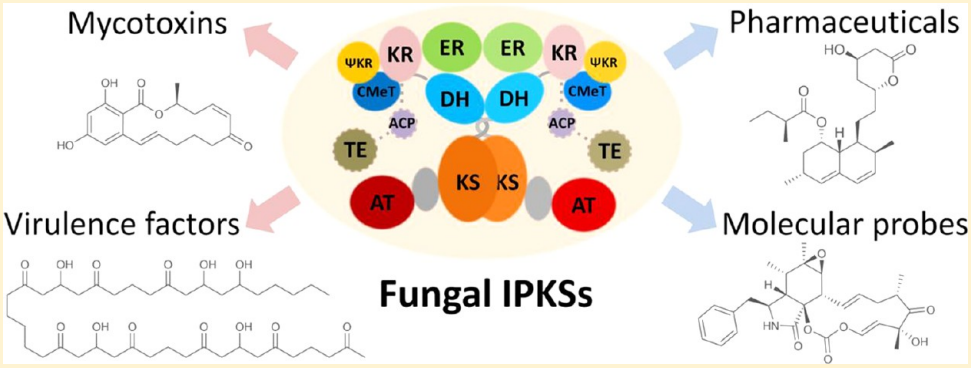
Perspective

pubs.acs.org/joc

Navigating the Fungal Polyketide Chemical Space: From Genes to Molecules

Yit-Heng Chooi[†] and Yi Tang*[,†,‡]

[†]Department of Chemical and Biomolecular Engineering, and [‡]Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States



ABSTRACT: The iterative type I polyketide synthases (IPKSs) are central to the biosynthesis of an enormously diverse array of natural products in fungi. These natural products, known as polyketides, exhibit a wide range of biological activities and include clinically important drugs as well as undesirable toxins. The PKSs synthesize these structurally diverse polyketides via a series of decarboxylative condensations of malonyl-CoA extender units and β-keto modifications in a highly programmed manner. Significant progress has been made over the past few years in understanding the biosynthetic mechanism and programming of fungal PKSs. The continuously expanding fungal genome sequence data have sparked genome-directed discoveries of new fungal PKSs and associated products. The increasing number of fungal PKSs that have been linked to their products along with in-depth biochemical and structural characterizations of these large enzymes have remarkably improved our knowledge on the molecular basis for polyketide structural diversity in fungi. This Perspective highlights the recent advances and examines how the newly expanded paradigm has contributed to our ability to link fungal PKS genes to chemical structures and vice versa. The knowledge will help us navigate through the logarithmically expanding seas of genomic information for polyketide compound discovery and provided opportunities to reprogram these megasynthases to generate new chemical entities.

INTRODUCTION

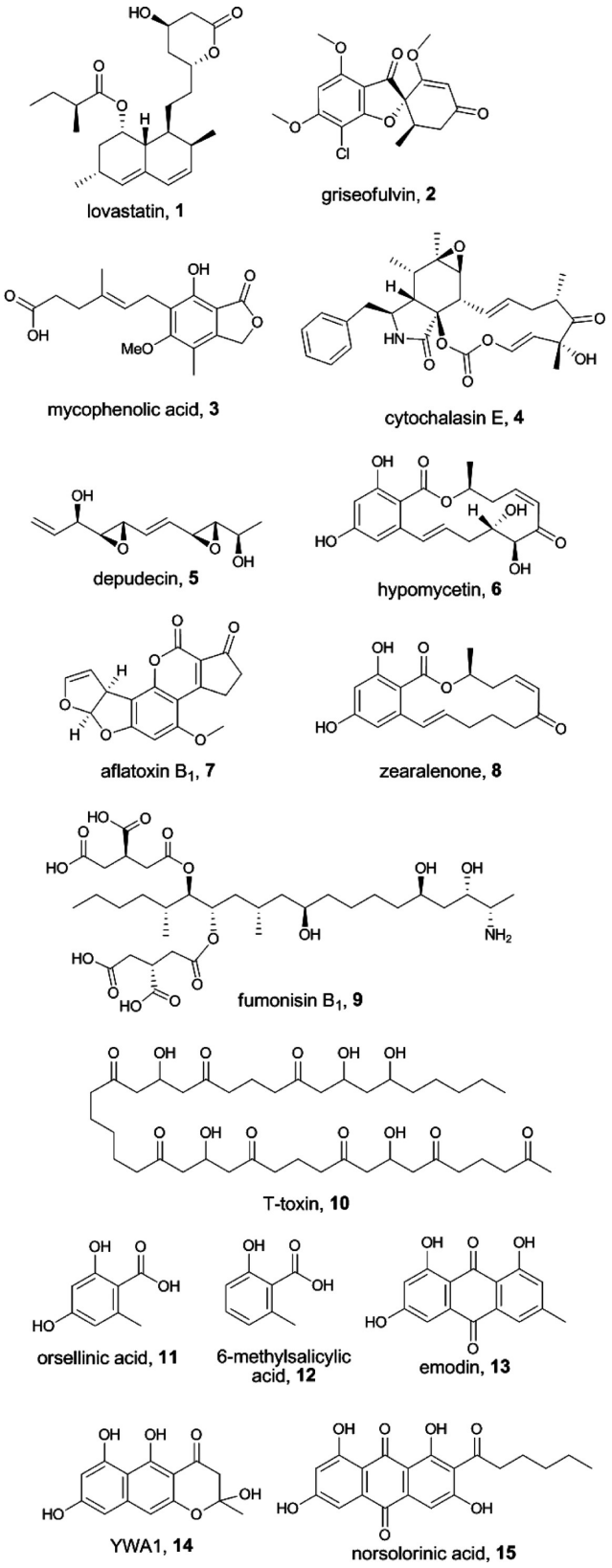
Polyketides encompass a large group of structurally diverse natural products unified by a common biosynthetic origin of the carbon backbones, which are derived from the polymerization of short-chain carboxylic acids units (acetate, propionate and butyrate, etc.). They are widely distributed across plants, bacteria, fungi, and some marine organisms. Their significance to humans has been recognized not only as an important source of medicine but also as toxins and virulence factors of pathogens.[1] As prolific producers of bioactive natural products, fungi are the source for several important polyketide-derived drugs, which include the cholesterol-lowering lovastatin **1**, the antifungal griseofulvin **2**, and the immunosuppressive mycophenolic acid **3**. Other bioactive fungal polyketides, such as the actin and angiogenesis inhibitor cytochalasin E **4**, the protein transport inhibitor brefeldin, the histone deacetylase inhibitor depudecin **5**, and the MAP kinase inhibitor hypomycetin **6** are important chemical probes for investigation of cellular processes. On the other hand, the fungal polyketides are also the source of many mycotoxins, such as aflatoxin B_1 **7**, zearalenone **8**, and fumonisin B_1 **9**, which have detrimental health effects to humans and a significant impact on agriculture.[2] Polyketides from phytopathogens, such as T-toxins (**10**), are important virulence factors in plant host infection.[3]

Historically, fungal polyketides have a special place in their contribution to our understanding of polyketide biosynthesis.[4,5] Classic fungal polyketides such as orsellinic acid **11**, 6-methylsalicylic acid **12**, griseofulvin **2**, and emodin **13** have been key model compounds in the development of the polyketide biosynthetic hypothesis (head-to-tail polyacetate hypothesis) in the 1950s by Birch.[6,7] Later isotopic tracer studies in fungi using ^{14}C-labeled and ^{13}C-labeled acetate with the introduction of NMR spectroscopy have validated the polyketide biosynthetic hypothesis.[8,9] Some of the earliest evidence for involvement of acetyl-CoA thioester in polyketide pathways was demonstrated with cell-free extracts from fungi.[10−12] In fact, the 6-methylsalicyclic acid (**12**) synthase (6MSAS) from the fungus *Penicillium patulum* is the first PKS

Received: August 3, 2012
Published: August 31, 2012

   dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933−9953

to be purified and characterized in vitro.[13] Thus, the fungal polyketides have been central to the early formation of polyketide research.



The following genetic and molecular biology era of polyketide biosynthetic studies (1990s) was dominated by the shifted focus to polyketide pathways in actinomycete bacteria.[4,5] Research in this area was accelerated by the development of genetics and molecular biology tools for these bacteria[4,14] and the findings that biosynthetic genes are clustered on host genomes.[5,15,16] The research progress in fungal polyketide biosynthesis was much slower due to the lack of genetic tools, the absence of functional enzyme preparations, and the general difficulties encountered in working with eukaryotes (i.e., larger genome size, intron-containing genes etc.). It was not until the past decade that we begin to gain more understanding on the molecular basis and mechanisms of the polyketide biosynthesis in fungi.[17] The improved molecular biology genetic tools, better heterologous hosts and recombinant enzyme preparation methods, advanced analytical instruments, and availability of genomic information have all contributed to the accelerated progress. The fungal polyketide chemical space will be loosely defined here as the total chemical diversity that could be generated by all the fungal polyketide pathways found in Nature. The surge of genomic information has revealed that we have only explored a very small fraction of this space compare to that is encoded within fungal genomes; this has sparked a renaissance in natural product discovery from fungi through genome mining of potential gene clusters.[18,19] This Perspective will provide an update on the recent insights into the fungal polyketide biosynthesis at the enzymatic level, with the main focus on the core polyketide synthase (PKS) enzymes. The knowledge accumulated in the past decade will help us navigate through the vast sea of mass genomic information that is continually expanding, and will be invaluable for our continuing quest for new drugs and agrochemicals.

■ POLYKETIDE SYNTHASES

PKSs are central to the biosynthesis of the carbon backbone in polyketides. Using enzymology parallel to that of the fatty acid synthases (FASs), the carbon backbones are built by repetitive decarboxylative Claisen condensation catalyzed by β-ketoacyl synthase (KS) using malonyl thioesters as extender units. The minimal enzyme components involved in chain extension are the KS, malonyl-CoA:ACP transacylase (MAT), and the acyl carrier protein (ACP). During each β-keto extension cycle, the three β-keto processing enzyme components ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) can, respectively, catalyze the reduction of the newly extended β-keto group to a β-alcohol, dehydration to an α,β unsaturated enoyl, and/or reduction of the double bond to a methylene. The PKSs can generally be categorized into type I and type II based on the classification of fatty acid synthases (FASs), as well as the ACP-independent type III PKSs; however, some exceptions that do not fit into this classification are occasionally discovered.[20,21] Introductions into the various types of PKSs have been summarized elsewhere.[1,21,22] A majority of fungal PKSs belong to a subcategory of type I PKSs, which are multidomain monomodular megasynthases. With a few exceptions, most of the fungal type I PKSs use the functional domains iteratively (as opposed to the single iteration multimodular type I PKSs in bacteria); thus, they are also known as the iterative type I PKSs (IPKS). Although some ACP-independent type III PKSs have been recently found in fungi,[23−25] this review will focused on the IPKSs.

One hallmark of type I PKSs is the juxtaposition of different catalytic domains in a single large polypeptide; hence, they are also referred as megasynthases. The IPKSs employed by fungi use the combinations of the active sites repeatedly in each cycle of chain extension and β-keto processing. The fungal IPKSs are

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933−9953

evolutionarily more related to the mammalian FASs (mFASs) and share significant protein sequence similarity and domain architecture. Currently, no X-ray crystal structure of a full-length IPKS has been solved, but the mFAS structure has been solved at 3.2 Å.[26] This has allowed us to speculate that the organization of IPKS domains in three-dimensional space should adopt a similar head-to-head intertwined dimeric structure. The mFAS structure also provided insights into how the ACP-tethered substrate may be shuttled between different functional domains in PKSs. However, unlike mFASs, which produce saturated fatty acids, the IPKSs are more programmed and can alter the extent of α- and β-carbon processing in different iterations to yield nonreduced, partially reduced, and highly reduced products. Furthermore, the IPKSs possess additional functional domains that facilitate various auxiliary modifications, such as cyclization, methylation, and amidation with amino acid moieties. Therefore, IPKSs possess a much higher level of programming complexity,[17] of which the molecular and biochemical basis are just beginning to be understood. Decoding how the different set of catalytic instructions are programmed into the IPKSs is the key to achieving the three overarching goals: (1) to be able to understand how variable substrate modifications are achieved in a possessive, multidomain enzyme; (2) to be able to accurately predict the polyketide metabolite structures based on primary sequences of the IPKSs; and (3) to be able to rationally engineer the IPKSs to produce new or hybrid polyketide scaffolds.

BRIDGING THE CONNECTIONS BETWEEN IPKS GENES, PROTEINS, AND METABOLITES

The new generation of natural product biosynthetic research has a strong focus on linking genes to molecules.[27] Over the past two decades, we have seen significant advancement in identifying and cloning the IPKS gene that encodes a specific metabolite in a fungal genome. The first sequenced fungal PKS gene encoding 6MSAS was cloned in 1990 by immunological screening of a genomic DNA expression library with polyclonal antibody.[28] In the earlier studies, the connections between the PKS genes and their products were more commonly established by complementation of blocked mutant with cosmid library. For pigment compounds, such as YWA1 **14**, this approach has been particularly successful.[29] The localization of the lovastatin nonaketide synthase (LNKS) gene, however, was a herculean task, which involved the screening of 6000 clones by TLC/HPLC due to lack of visible phenotype.[30] The availability of the handful of IPKS genes has leads to the development of degenerate primer PCR that target the highly conserved KS domain to generate homologous probe for DNA library screening. This led to the discovery of IPKS genes, such as the *pksL1* that encodes the norsolorinic acid (**15**) synthase (NSAS) in aflatoxin **7** pathway[31] and *fum5* that encodes fumonisin (**9**) PKS.[32] The realization that similarities in metabolite structures is reflected in the IPKS protein sequences has sparked the development of degenerate primers that can target specific classes of PKS genes.[33,34] These degenerate primers have facilitated the discovery of numerous PKS genes[35,36] and evaluation of PKS gene diversity in various fungi.[37−39]

There has been a significant increase in the number of polyketide compounds that has been linked to the fungal IPKS genes over the past 5 years. Fungal genome sequences have revealed the complete inventory of IPKS genes that are encoded in each genome and shown that there is much untapped polyketide biosynthetic potential in fungi. For example, in the genus *Aspergillus*, each organism encodes 13−35 PKSs,[40] far exceeding the number of metabolites previously isolated from these species. The majority of these are orphan PKS genes, in which their products are unknown, and most of them are silent under laboratory culture conditions.[18,19] This discovery has propelled genome-oriented polyketide compound discovery, especially in model organisms like *A. nidulans*, and leads to the discovery of known and novel polyketide compounds that have previously never been isolated in this species such as aspyridone,[41] asperthecin,[42] asperfuranone **18**,[43] F9775A/B,[44,45] etc. These are facilitated largely by the development of versatile high DNA recombination frequency strains[46] as well as approaches for activation of silent gene clusters.[18,19,41,44,45] Furthermore, the recent introduction of next-generation sequencing technologies, such as the Roche 454, Illumina GAIIx, and the more recent Illumina Hiseq2000, has dramatically lowered the cost for fungal genome sequencing; thus, it will soon be common place for individual laboratories to undertake natural product research-motivated fungal genome sequencing projects as exemplified in our recent discovery of two IPKS gene cluster for biosynthesis of **2** and viridicatumtoxin **17**.[47] Nevertheless, knowledge on the detail inter-relationships between the PKS genes, proteins, and the metabolite structures is required to efficiently harness the power of the new genomic technologies for polyketide compound discovery.

It is now well-accepted that the fungal IPKSs can be functionally and phylogenetically grouped into three major classes, i.e., the nonreducing (NR), the highly reducing (HR), and the partial-reducing (PR) PKSs.[17,34,48] This classification based on the extent of β-keto reduction catalyzed by the IPKSs is useful as it is reflected in the PKS domain architecture as well as the produced metabolites. The aromatic polyketides are produced by NR-PKSs by cyclization of the nonreduced reactive poly-β-keto chains into mono- or polycyclic compounds, such as **11**,[49] **13**,[44] and **15**.[31] The HR-PKSs produce a large diversity of linear and cyclic nonaromatic compounds including well-known examples such as, **1**,[50] **5**,[51] **9**,[32] and **10**.[52] The PR-PKSs have so far only been linked to production of **12**,[28,53] which requires a single reduction of a specific keto group in a tetraketide backbone. Phylogenetic analyses have shown that the 6-MSASs and other uncharacterized PR-PKSs are less related to the NR- and HR-PKSs but more related to the bacterial modular type I PKSs, supporting an origin from actinobacteria by horizontal gene transfer.[39,48] Research in our laboratory has focused on the enzymology of NR-PKSs and HR-PKSs, which will be predominantly discussed here.

NONREDUCING POLYKETIDE SYNTHASES

Besides containing the minimal PKS domains (KS, AT, and ACP), the NR-PKSs lack the trio of β-keto processing domains (KR, DH, and ER) that are present in the HR-PKSs and mFASs. On the other hand, there are additional functional domains that are unique to the NR-PKSs, including the starter unit:ACP transacylase (SAT) domain, product template (PT) domain, and thioesterase (TE) releasing domain. Our understanding of the enzymology of the fungal NR-PKSs has remarkably improved recently via various in vitro enzymatic studies. Along with that, a significant number of aromatic polyketides has also been linked to NR-PKS genes recently (Table 1). Reconstitution of purified biosynthetic enzymes (in

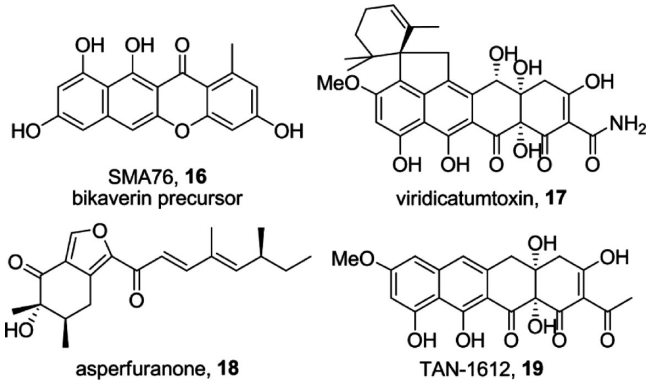
dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933−9953

Table 1. Selected Metabolites That Have Been Linked to NR-PKSs in the Past 5 Years

final product	first isolable product/ intermediate*	organisms	genes/locus	PKS	domain architecture	ref
asperthecin	acetoatrochrysone	*A. nidulans*	AN6000.3	AptA	SAT-KS-AT-PT-ACP	42, 54
atrochrysone	atrochrysone	*A. terreus*	ATEG_08451	ACAS	SAT-KS-AT-PT-ACP	55
hypothemycin, **6**	7′,8′-dehydrozeralenol (DHZ)	*Hypomyces subiculosus*	*hpm3*	Hpm3-NR	SAT-KS-AT-PT-ACP-TE	56
radicicol	monocillin II, **22**	*Pochonia chlamydosporia/ Chaetomium chiversii*	*rdc1, ccRads2*	Rdc1-NR	SAT-KS-AT-PT-ACP-TE	56, 57
asperfuranone, **18**	benzaldehyde intermediate	*A. nidulans*	AN1034.3	AfoE	SAT-KS-AT-PT-ACP-CMeT-R	43
F9775A/Blecanoric acid	orsellinic acid	*A. nidulans*	AN7909.3	OAS	SAT-KS-AT-PT-ACP-TE	44, 45, 49
viridicatumtoxin, **17**	5-dehydroxyanthrotainin	*P. aethiopicum*	*vrtA*	VrtA	SAT-KS-AT-PT-ACP	47, 58
griseofulvin, **2**	norlichexanthone	*P. aethiopicum*	*gsfA*	GsfA	SAT-KS-AT-PT-ACP	47
TAN-1612, **19**	TAN-1612	*A. niger*	ASPNIDRAFT_	AdaA	SAT-KS-AT-PT-ACP	54
austinol	3,5-dimethylorsellinic acid	*A. nidulans*	AN8383.3	AusA	SAT-KS-AT-PT-ACP-CMeT-EST	59
terretonin	3,5-dimethylorsellinic acid	*A. terreus*	*trt4*/ ATEG_10080	Trt4	SAT-KS-AT-PT-ACP-CMeT-EST	60
mycophenolic acid, **3**	5-methylorsellinic acid	*P. brevicompactum*	*mpaC*	MpaC	SAT-KS-AT-PT-ACP-CMeT-EST	61, 62
alternariol	alternariol	*A. nidulans*	AN7071.3	AtnS	SAT-KS-AT-PT-ACP	63
tropolone	3-methylorcinaldehyde acid	*Talaromyces stipitatus*	*tropA*	TropA	SAT-KS-AT-PT-CMeT-R	64
O-methyldiaporthin	hexaketide isocoumarin	*A. oryzae*	*aoiG*/ AOR_1_74024	AoiG	SAT-KS-AT-PT-ACP-ACP-TE	65

both dissociated and intact forms) enables the most direct analysis of the functions and products of the megasynthases and has been a key approach in understanding the bacterial PKSs and NRPSs (noribosomal peptide synthetases).[66] One approach is via "enzyme deconstruction" where the fungal IPKSs are expressed as individual dissociated catalytic domains for in vitro enzymatic assays.[67−69] Our laboratory has employed heterologous systems for expression and in vitro reconstitution of these large megasynthases in intact form.[54,70−76] In vivo NR-PKS domain swapping experiments have also illuminated important insights into the programming rules in these megasynthases.[77−83] Furthermore, the crystal and solution structures of several functional domains of NSAS have provided important structural and mechanistic insights into NR-PKSs.[84−86] It is now possible to reasonably predict the product structures (size, number of aromatic rings, etc.) based on the primary sequence of the NR-PKSs. A majority of the recent enzymology studies have been performed on two PKS systems of similar domain architecture, the *Aspergillus flavus* PKSA (norsolorinic acid synthase, NSAS) that produces **15** and *Gibberella fujikuroi* PKS4 (GfPKS4) that produces SMA76 (**16**), the intermediate of bikaverin. The recent insights into the enzymology of NR-PKSs have been reviewed.[87] We will discuss here the roles of the different functional domains in NR-PKSs, including the recent findings in our laboratory and other less well-characterized NR-PKS domains.

Loading of Starter Unit. The molecular basis for the incorporation of starter unit has been discovered by the Townsend group in the context of the NSAS.[67] Compound **15** is the first isolable intermediate in the biosynthesis of the aflatoxins (e.g., **7**).[88,89] The biosynthesis of **15** by the NSAS requires a hexanoate starter unit produced by a dedicated fungal FAS (HexA/HexB) (Figure 1A).[90,91] By using an "enzyme deconstruction" approach guided by the Udwary−Merski algorithm in silico,[92] the hexanoyl starter unit loading function in the NSAS was mapped to an *N*-terminal conserved domain, which exhibits secondary-structural similarity to MATs domains



and was named starter unit:ACP transacylase (SAT) domain.[67] The SAT domain transfers a starter acyl unit to the ACP, which then primes the KS domain to initiate the first cycle of extension (Figure 1A). In NSAS, the SAT domain use thioester chemistry, having a GXCXG active site, which is conserved in *A. nidulans* WA (produces **14**),[93] *C. lagenarium* ClPKS1 (tetrahydroxynaphthalene),[94] *Cercospora nicotianae* CTB1 (cercosporin),[95] and many other NR-PKSs. This discovery of the molecular basis of starter unit loading has been extended to rationalize the classical acetyl "starter-unit effect" observed in early isotope tracer studies of fungal aromatic polyketide compounds and the selectivity of several NR-PKS SAT domains toward acetyl-CoA was demonstrated for WA, ClPKS1, CTB1, and GfPKS4 (Figure 1B).[96]

The SAT domain of NR-PKSs also plays an important role in biosynthesis of resorcylic lactones (RALs),[97] an important group of polyketides that include members such as the estrogenic mycotoxin **8**, the mitogen-activated protein (MAP)-kinase inhibitor **6**, and the heat-shock protein (HSP)-90 inhibitor radicicol.[56,57,71,73,74] The biosynthesis of RALs involved the synthesis of highly reduced starter unit by HR-PKSs and direct transfer to the SAT domain of NR-PKSs for further chain extension, cyclization, and release (Figure 1D).

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933−9953

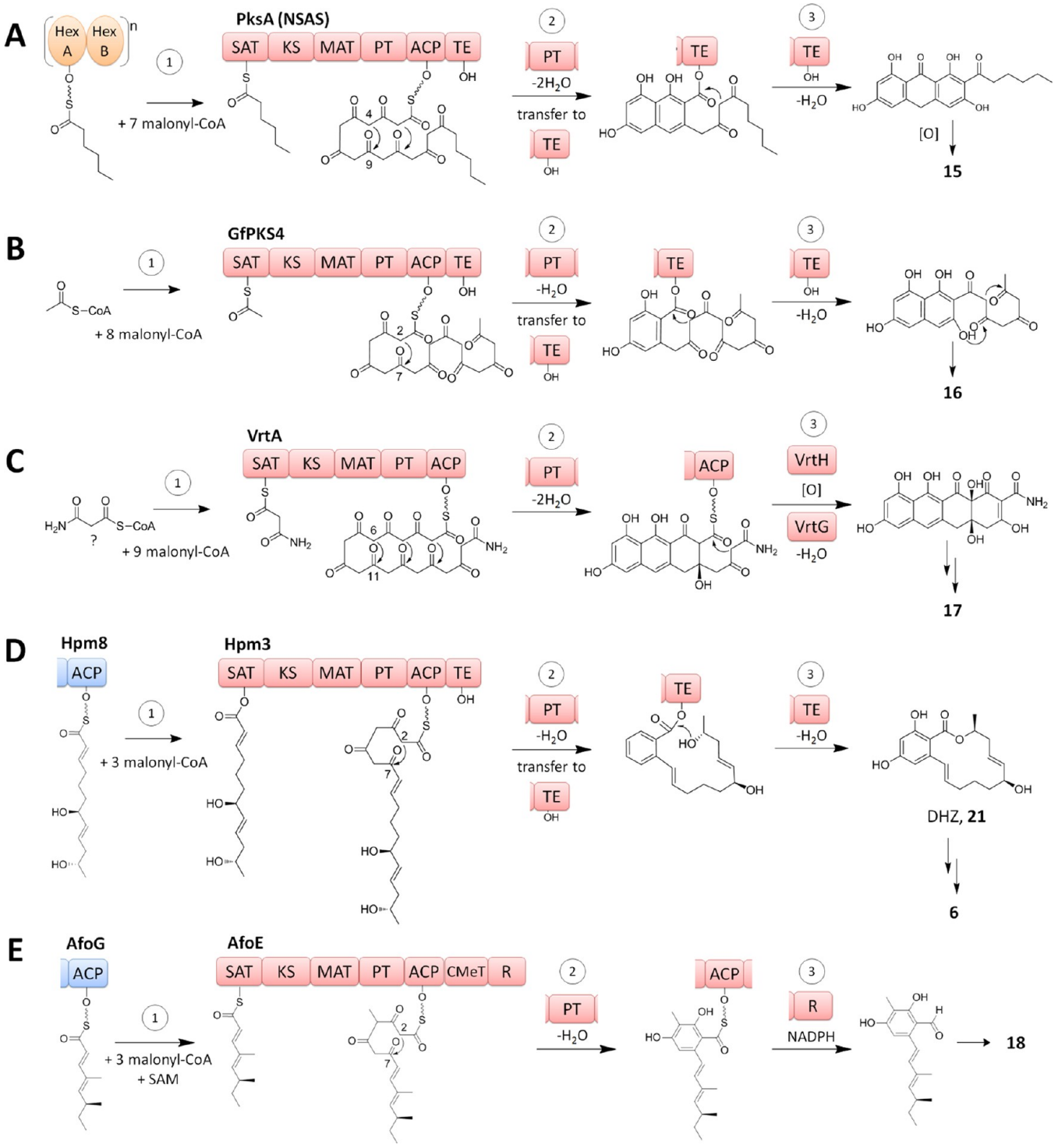


Figure 1. Biosynthesis of aromatic polyketides by various NR-PKSs highlighting the variations in (1) loading of starter units, (2) PT-mediated aldol cyclization, and (3) product release mechanisms.

This was demonstrated in the complete reconstitution of the two PKSs involved in biosynthesis of **6** and radicicol.[71,74] Interestingly, the SAT domain of RAL NR-PKSs contains a GXSXG active site in the SAT domain instead of GXCXG; point mutation of the serine residue resulted in loss of activity for GzPKS13 and Hpm3.[71,73] Thus, the RAL NR-PKSs use oxyester chemistry common to MAT and other acyltransferases. Similarly, the SAT domain is involved in the biosynthesis of asperfuranone **18** (Figure 1E).[43] Swapping of the SAT domain of the asperfuranone NR-PKS (AfoE) with the SAT domain from NSAS (StcA) in an *A. nidulans* Δ*stcA* mutant strain resulted in the production of a hexanoyl primed 1,4-naphthaquinone derivative.[79] Besides the more common fatty acid and polyketide-derived acyl starter units, the SAT domain of VrtA was proposed to prime the biosynthesis of **15** with a unique malonamate stater unit (Figure 1C).[47]

An alternative SAT-independent priming mechanism was also observed in vitro when the RAL NR-PKSs of the **9** and **6**

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933−9953

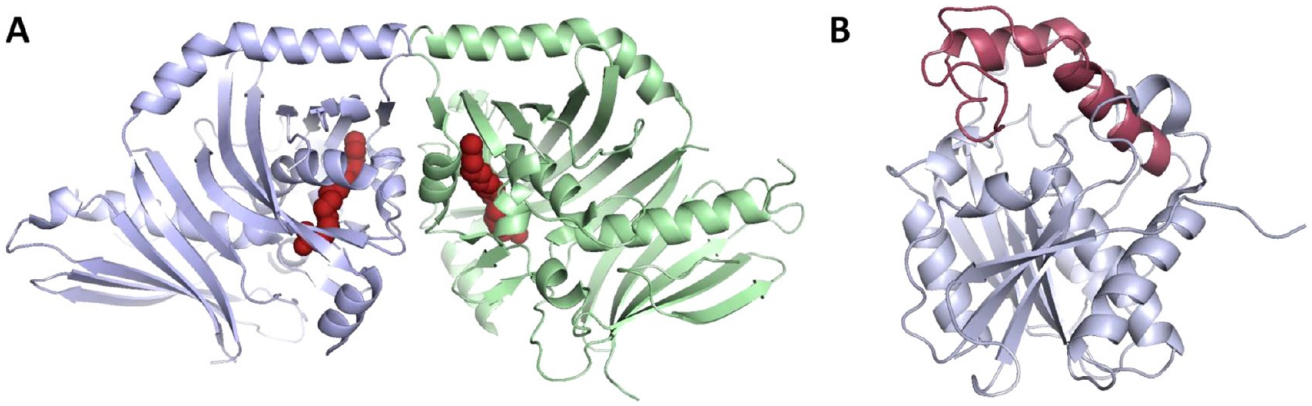


Figure 2. Ribbon diagram of (A) the PksA dimeric PT structure, showing the palmitate (red) bound in the deep pocket; (B) PksA TE domain structure in close configuration showing the lid loop (magenta).

(GzPKS13 and Hpm3) were assayed with acyl-SNAC (*N*-acetylcysteamine thioester) substrates.[71,73] In the case of GzPKS13, acyl-SNACs of different chain lengths were tested and the NR-PKS exhibits a preference toward C_6−C_{10} acyl-SNACs (the native starter unit is C_{12}).[71] The ability to utilize the acyl-SNAC starter units of these NR-PKSs is not affected by mutation of the active site serine at the SAT domain, suggesting the acyl-SNACs are load directly onto the KS domain.[71,73] Similar phenomenon was observed in the precursor directed biosynthesis of octanoyl-primed benzopyrone derivatives using GfPKS4, where octanoyl-CoA can be incorporated in a SAT-independent manner.[70,96]

Extension and Chain Length Control. The extension component of PKSs consists of the minimal PKS domains, KS, MAT, and ACP. There is evidence that indicates the chain length control of NR-PKSs primarily resides in the KS domain, likely limited by its cavity size. In the precursor-directed biosynthesis by GfPKS4, incorporation of a longer starter unit leads to reduced number of β-keto extension cycles and yielded polyketide derivatives that are similar in carbon chain length to the native product **16**.[70] More substantial evidence comes from another study in our laboratory where the minimal PKS extension engine of GfPKS4 was reassembled to produce bacterial-type aromatic polyketides in *E. coli*.[98] The KS-MAT didomain and a linked or stand-alone ACP domain was demonstrated to synthesize the expected nonaketide chain in the absence of SAT, PT, and TE.[98] Similarly, the hybrid NR-PKS AfoE with the hexanoyl-priming SAT domain from NSAS catalyzes an additional extension cycle to yield the 1,4-naphthaquinone derivative that has the identical chain length as asperfuranone.[79] The involvement of MAT or ACP in controlling the chain length is less plausible. The MAT domain has been shown to only responsible for the loading of malonyl-CoA to the ACP and is not involved in the transport of the growing poly-β-keto intermediate.[69] Additionally, the recent solved ACP solution structure of NSAS showed that the ACP appeared to have minimal interaction with tethered acyl thioester.[86]

Polyketide Folding and Aldol Cyclization. The ability of NR-PKSs to stabilize the highly reactive nascent poly-β-keto chains and control the regioselectivity of the cyclization was one of the mysteries of these megasynthases. The three common first-ring cyclization patterns of fungal aromatic polyketides are C2−C7, C4−C9, and C6−C11 (Figure 1). To determine the enzymatic basis of fungal polyketide cyclization, the Townsend group dissected the NSAS and reconstituted the polyketide synthesis activity in vitro using a combination of individual domains or didomains of the NSAS.[68] The study localized the cyclization function of NR-PKSs to a previously unknown conserved domain, named product template (PT) domain, immediately upstream of the ACP (Figure 1A).[68] The NSAS PT domain was shown to mediate the consecutive C4−C9 and C2−C11 aldol cyclization reactions. Complementary experiments in our lab have shown that removal of the PT domain disrupted the C2−C7 cyclization function of GfPKS4 and lead to spontaneous cyclization of the backbone into various shunt products.[98,99] Furthermore, the poly-β-keto chain synthesized by the minimal PKS can be captured by various bacterial cyclases to form different "S-folded" bacterial-type aromatic compounds.[98]

Recently, the crystal structure of the dimeric PT domain of NSAS was solved at 1.8 Å.[84] The high-resolution structure showed that the PT domain exists as dimer and exhibits a modified "double hot dog" fold and is structurally similar to the DH domain in HR-PKSs and mFASs (Figure 2A). Thus, the PT domain could be evolved from DH domain of ancient HR-PKSs.[84] A deep substrate binding pocket is present in the PT domain structure, where it could be divided into (1) the outer phosphopantetheinyl (PPT)-binding region that binds to the PPT-arm extended from the ACP, (2) the core cyclization chamber where the corresponding C4 α-carbon and the C9 carbonyl of the poly-β-keto intermediate are brought to proximity for regiospecific aldol condensation, and (3) a hexyl starter unit binding pocket.[84] In the same study, the oligomeric status of the NSAS domains was also examined using native PAGE mobility experiments; both the KS domain and the PT domain were shown to exist in dimer, while the C-terminal ACP-TE didomain and the N-terminal SAT domain are likely to be monomer. The dimeric nature of PT domain suggests that it also plays a role, along with the KS domain, in the dimerization of the NR-PKS.

The different forms of first-ring cyclization (C2−C7, C4−C9, and C6−C11) are likely mediated by the shape of the internal pocket of the PT domains. Indeed, phylogenetic analysis shows that the PT domains of NR-PKSs (excluding CMeT domain-containing NR-PKSs) can be divided into five major groups, which correlate with the cyclization regioselectivity and the length of the poly-β-ketone portion of the final products.[99] The regioselectivity of C6−C11 PT domains (group V) were confirmed by assaying excised PT domains (from AptA and VrtA) using the engineered minimal GfPKS4 (PKS_WJ) that lacks the native C2−C7 PT domain.[99] Addition

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933−9953

of the heterologous PT domain *in trans* fixed the C6−C11 cyclization pattern of the nonaketide produced by PKS_WJ. Besides providing a convenient "read-out" for PT domains of unknown regioselectivity, PKS_WJ has also been used to study bacterial cyclases; it was demonstrated that the bacterial cyclases (e.g., TcmN and ZhuI) can direct the cyclization of the unreduced nonaketide chain produced by PKS_WJ.[98] Combining PKS_WJ with *gris* CYC1, *oxy* CYC2, and *act* KR, the anthraquinone SEK26 can be successfully produced in vitro and in *E. coli*.[98] Bioinformatic and functional PT domain studies also complemented the previous *vrtA* gene deletion study[47] and confirmed that synthesis of the tetracyclic carboxamide precursor of **17** proceeds via a C6−C11 cyclization (Figure 1C).[99] Furthermore, the PT domain cyclization prediction also led to the discovery of the *A. niger ada* cluster that synthesize TAN-1612 **19**.[54] AdaA is the NR-PKS known to synthesize the longest polyketide chain (C20) so far. Using a similar approach to GfPKS4/PKS_WJ,[98] the minimal PKS of AdaA may be potentially hijacked to synthesize bacterial decaketides, such as the anticancer doxorubicin, the antibiotics tetracyclines and angucyclines, in *E. coli*.

The very recent *A. nidulans* genome-wide NR-PKS genes promoter-swapping experiment revealed the aromatic polyketides derived from different cyclization patterns encoded by the *A. nidulans* NR-PKSs.[63] In the same study, phylogenetic analysis of the NR-PKSs is extended to include the CMeT domain-containing NR-PKSs (cover 71 fungal NR-PKSs available at the Broad Institute Aspergillus Comparative Database and other characterized NR-PKSs). Interestingly, the phylogenetic analyses showed that the topology of the PT domain tree and the tree constructed using whole NR-PKSs are nearly identical, in which the NR-PKSs can be grouped into 7 clades (the CMeT domain-containing NR-PKSs are divided into two groups, VI and VII).[63] Thus, the PT domains have likely coevolved with the entire NR-PKSs and no PT domain recombination event occurs after the initial divergence from the putative common ancestor into these groups of NR-PKSs.

Besides the more common C2−C7, C4−C9, and C6−C11 first-ring aldol cyclization regiospecificities, there also exists the more unusual cyclization pattern, such as the one occur in the biosynthesis of **2** and the related norlichexanthone, which involved a C1−C6 Claisen condensation and a C8−C13 aldol cyclization. The griseofulvin (*gsf*) gene cluster in *P. aethiopicum* has been discovered by our group recently via shotgun genome sequencing.[47] The PKS GsfA bears a typical NR-PKS domain organization but lacks the TE domain. Since no enzyme encoded in the *gsf* cluster or additional domain in the GsfA appears to catalyze the C1−C6 Claisen condensation, we propose that the GsfA PT domain is involved in both cyclization steps. The recent review on cyclization of aromatic polyketides from bacteria and fungi relates the different cyclization patterns to various fungal aromatic polyketides.[100]

Claisen Cyclization/Macrolactonization and Product Release. The typical function of TE domain in fatty acid biosynthesis is hydrolytic release of fatty acid from mFASs. In bacterial type I modular PKSs and NRPSs, the C-terminal TE domains catalyze macrolactonization or macrolactamization reactions.[101] Similar TE domains are present at the C-terminal of many fungal NR-PKSs but have been shown to catalyze cyclization via intramolecular Claisen condensation (Dieckmann condensation). The role of NR-PKS C-terminal TE domain in Claisen cyclization was first identified in the biosynthesis of the pigment YWA1 by *A. nidulans* WA PKS via in vivo heterologous expression and mutagenesis experiments; the TE domain was renamed as Claisen-like cyclase domain (CLC).[93] Later, in vitro assays with intact NR-PKS and dissected domains further confirmed the role of TE/CLC domain in catalyzing the Claisen cyclization in the formation of **16** by GfPKS4 (Figure 1B)[70] and **15** by NSAS (Figure 1A).[68] The TE/CLC domain of GfPKS4 when expressed and purified as standalone protein can interact *in trans* with the GfPKS4 that lacks the TE/CLC domain or contain a point mutation at the active serine site.[70]

The 1.7 Å crystal structure of the PksA (NSAS) TE/CLC domain exhibits an α/β-hydrolase fold and contains the expected catalytic triad (Ser-His-Asp) typical of serine hydrolases.[85] The crystal structure was captured in a closed catalytic conformation (Figure 2B). It was proposed that docking of the loaded ACP domain will induce conformational change to the open form for insertion of the ACP tethered intermediate.[85] More recently, the Townsend group revealed that the TE domain plays an additional role in biosynthetic editing by removal of stalled intermediates.[102] Using mass spectrometry techniques coupled with proteolysis, catalytic snapshots of individual domains/active sites with the covalently bound intermediates were obtained for PksA. It was observed that the presence of covalent species with incorrectly bound intermediates is markedly reduced in the presence of TE.

Unlike the NR-PKSs that synthesize fused-ring compounds, the TE domains of RAL-type NR-PKSs catalyze macrolactonization (Figure 1D),[71,73,74] a releasing mechanism commonly seen in bacterial type I modular PKSs.[103] The macrolactonization function of the TE domain of RAL-type NR-PKSs was first demonstrated in GzPKS13, the NR-PKS of **8** pathway.[71] The GzPKS13 TE domain exhibits less than 20% identity to aforementioned TE/CLC and bacterial macrolactonization TE domains.[104] Equipped with the intact recombinant GzPKS13 and the 11-hydroxyundecanyl-SNAC substrate mimic, the entire catalytic repertoire of GzPKS13 can be reconstituted in vitro.[71] The ω-hydroxyl group of the 11-hydroxyundecanoyl moiety was able to mimic the terminal alcohol in the natural substrates and yielded a corresponding RAL analogue. Subsequent biochemical characterization of the excised GzPKS13-TE domain using SNAC substrate mimics confirmed that the macrolactonization function of GzPKS13 resides in the TE domain.[104] In addition to macrocyclization, GzPKS13-TE can perform esterification in the presence of primary alcohols, amidation in the presence of amides and hydrolysis in the absence of other nucleophiles.[104]

Instead of TE domain, a C-terminal reductase (R) release domain is found in some NR-PKSs, mostly from the group previously classified as Clade III NR-PKSs[48] that contain a CMeT domain downstream of ACP (Figure 1E). The function of the R domain was first demonstrated by the Cox group in the heterologous expression of the *Acremonium strictum* AsPKS1 in *Aspergillus oryzae*, in which 3-methylorcinaldehyde was identified as the PKS product.[105] The release of the polyketide product as an aldehyde is indicative of a reductive release instead of a hydrolytic release. Subsequent study by heterologous expression of the C-terminal truncated AsPKS1 (AsPKS1-ΔR) confirmed the role of R domain in catalyzing the reductive release.[106] Interestingly, expression of the AsPKS1-ΔR in *A. oryzae* resulted in production of 3-methylorsellinic acid. Thus, there is still much to discover about these Clade III NR-PKS systems.

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933−9953

In the biosynthesis of **18**, the aldehyde from reductive release catalyze by the R domain of AfoE (a Clade III NR-PKS) is key to the formation of the furan ring (Figure 1E).[43] More recently, a similar 3-methylorcinaldehyde synthase from *Talaromyces stipitatus* was found to be involved in biosynthesis of tropolone.[64] On the other hand, the *A. niger* silent gene cluster that encodes biosynthesis of azaphilones was recently discovered in our laboratory.[107] Activation of the *aza* gene cluster led to the production of several azaphilones and an early benzaldehyde intermediate presumably produced by the Clade III NR-PKS AzaA. In vitro assay of the benzaldehyde intermediate with a recombinant flavin-dependent monooxygenase AzaH (encoded in the aza cluster) revealed the formation of the characteristic pyrano-quinone core in azaphilones via a hydroxylation-mediated pyran cyclization between the aldehyde and an intramolecular carbonyl carbon.[107]

More recently, several CMeT-containing NR-PKSs, including the MpaC in pathway for **3**,[61,62] AN8383.3 in austinol biosynthesis,[59] the Trt4 in terretonin pathway,[60] and the 5-methylorsellinic acid synthase CHGG_10128,[108] have been identified to contain a C-terminal esterase/lipase-like (EST) domain.[108] These EST domains belong to the family of serine hydrolases and contain the GXSXG active site motif, but only exhibits low sequence similarity (~10% identity to NSAS/PksA-TE) and appears slightly larger than the TE/CLC domains. On the basis of the structure of the direct PKS product from these NR-PKSs, which are 3,5-methylorsellinic acid for AN8383.3 and Trt4, and 5-methylorsellinic acid for MpaC and CHGG_10128, the EST domains function as a hydrolase instead of a TE/CLC. The low homology of EST domain to TE/CLC domain and their almost exclusive presence in the Clade III NR-PKSs imply that the EST and TE/CLC originated from separate gene fusion events, and explains the location of CMeT domain between the ACP and EST domains in these Clade III NR-PKSs.

Numerous NR-PKSs lack a releasing domain, especially those containing a PT domain that is classified as group V (C6−C11 regioselectivity).[48] The *A. nidulans* NR-PKS AptA (AN6000.3) was shown to be involved in biosynthesis of asperthecin.[42] Based on gene deletion and bioinformatics analysis, a standalone metallo-β-lactamase (MβL)-like hydrolase encoded by *aptB* was proposed to be involved in hydrolytic release of the polyketide product from AptA. Ensuing study on the atrochrysone pathway in *A. terreus* provided the biochemical evidence that the MβL-like hydrolase (atrochrysone carboxylic acid thioesterase, ACTE) essentially acts as a hydrolytic thioesterase *in trans* to release the mature polyketide product from the NR-PKS ACP as a carboxylic acid.[55] A surprise finding comes from our recent study of TAN-1612 **19** biosynthesis that such MβL-type TE can participate in a Claisen-like cyclization to form tetracyclic products.[54] Using pathway reconstitution in yeast and in vitro assays, it was demonstrated that the fourth-ring Claisen cyclization of **19** catalyzed by MβL-TE (AdaB) requires a prior hydroxylation at the carbon α to the thioester by a flavin-dependent monooxygenase (AdaC). In the absence of Cα hydroxylation, AdaB instead catalyzes a hydrolytic release.[54] As all known tetracyclic fungal polyketides (**17**, hypomycetin, anthrotainin) contain a similarly positioned hydroxyl group, it is likely that identical fourth ring Claisen cyclization mechanism involving analogs of AdaB and AdaC takes place (Figure 1C). This strategy is markedly different from the cyclization of bacterial tetracyclic polyketides, which involved multiple discrete cyclases.[100,109,110]

In conclusion, based on bioinformatic and biochemical studies of different domains in NR-PKSs, we can now reasonably predict the type of structures that may be produced from these IPKSs. While the PT domain grouping offers insight into the regioselectivity of the initial cyclization events, the identity of chain releasing domains/enzymes yield information on the overall structure of the product. More accurate prediction of exact chain lengths, which is not possible at this point, will require structural information of the KS domains.

■ HIGHLY REDUCING POLYKETIDE SYNTHASES

The HR-PKSs exhibit a much higher degree of complexity in their biosynthetic programming compared to the NR-PKSs. Besides domains that control chain extension and release, the HR-PKSs contain three β-keto processing domains that could selectively catalyze up to three levels of β-keto reduction (ketoreduction, dehydration and enoylreduction) during each extension cycles. In addition, many HR-PKSs also possess a CMeT domain for methylation of α-carbon on the extended polyketide chain using *S*-adenosylmethionine (SAM) as cofactor. For example, biosynthesis of the lovastatin decalin core dihydromonacolin L **31** by LNKS (LovB) requires 35 or more highly orchestrated sequence of catalytic steps.[72] Furthermore, the ability of some HR-PKSs like LovB to catalyze the formation of fused nonaromatic ring systems via intramolecular cycloaddition (such as Diels−Alder reaction) can introduce significant structural complexity from seemingly simple linear carbon backbones. Some HR-PKSs are also fused with a complete C-terminal nonribosomal peptide synthetase (NRPS) module, forming a PKS-NRPS hybrid megasynthetase that can incorporate an amino acid into the polyketide product. The increased complexity and the larger size of HR-PKSs, as well as difficulties associated with the analysis of acyclic, nonaromatic compounds, have impeded the progress on detailed biochemical investigation of HR-PKSs. Another unique feature is that most HR-PKSs lack a built-in releasing domain and require an additional releasing enzyme that is often not identified. The incomplete understanding of HR-PKS releasing mechanism has also complicated efforts in the heterologous reconstitution of HR-PKS systems. Recent advances in multigene expression tools and in vitro reconstitution methods have improved our understanding of the HR-PKS systems; however, the biochemical and structural aspects of these complex biosynthetic machineries have remained enigmatic.

Recent progress on the biochemistry of HR-PKSs has been made using enzymes involved in biosynthesis of **1**, **9**, RALs (e.g., **6** and **8**), and tetramic acid-derived products. The discussion will divided into sections of HR-PKSs that synthesize linear-chain products, decalin-forming HR-PKSs, and the PKS-NRPS hybrids.

Linear Chain HR-PKSs. Most of the HR-PKSs that produce reduced linear polyketide chain so far appear to contain a functional ER domain and do not require a *trans*-acting ER enzyme. One of the most simple HR-PKS products is the methylated diketide (α-*S*-methyl butyrate) synthesized by the 277 kDa lovastatin diketide synthase (LDKS or LovF).[50] LovF contain the archetypical HR-PKS organization with the complete set of chain extension (KS-AT and ACP) and β-keto processing domains (KR, DH, and ER) as well as a CMeT domain. Biosynthesis of the α-*S*-methyl butyrate unit involves all the domains in a single iteration, which represents the most

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933−9953

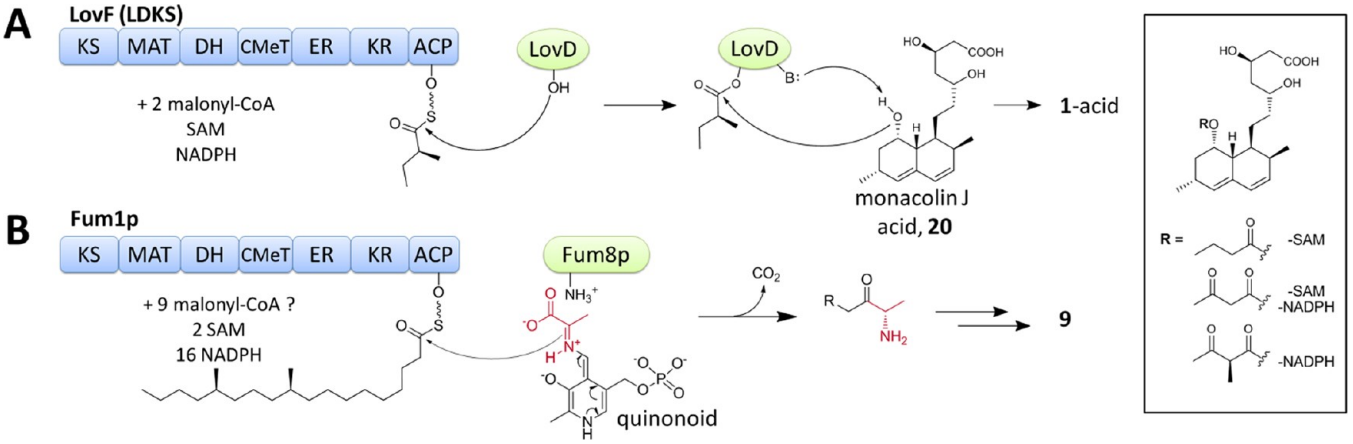


Figure 3. Off-loading of reduced linear polyketide product synthesized by HR-PKSs. (A) LovD acyltransferase-mediated offloading in biosynthesis of **1**. Lovastatin derivatives obtained by in vitro feeding of diketide SNAC substrates shown in the box. (B) FUM8p-mediated offloading in biosynthesis of **9**.

simple programing rule. By in vitro reconstitution of LovF from a yeast expression system BJ5464-NpgA, it was demonstrated that the acyltransferase LovD is essential for off-loading of the α-*S*-methyl butyrate product.[122] LovD catalyzes the acyl transfer of the diketide substrate tethered to the ACP of LovF to the C-8 hydroxyl group on the acyl acceptor monacolin J to form **1**. Further kinetic experiments comparing the catalytic efficiency of LovD toward α-*S*-methylbutyryl attached to CoA, SNAC, stand-alone LovF-ACP, and intact LovF demonstrated that protein−protein interactions between LovD and LovF-ACP play a key role in efficient catalysis.[122] Protein−protein interactions between large PKS and a smaller accessory enzyme is a recurring theme in fungal IPKSs. Interestingly, when NADPH, SAM, or both was omitted from the in vitro assays with LovF, LovD, and monacolin J, derivatives of **1** with modified side chains were obtained (Figure 3).

The squalestatin tetraketide synthase (SQTKS), involved in the synthesis of squalestatin in *Phoma sp.* C2932,[123] is a close homologue of LovF in which its polyketide product has been identified.[35] It has the exact domain architecture as LovF but exhibits a higher level of programming complexity in that it catalyzes two more iterations of extension, during which the tailoring domains are selectively used. When heterologously expressed in *A. oryzae,* the dimethylated tetraketide product of SQTKS was recovered.[35] Since SQTKS lacks a releasing domain, the ability to detect the tetraketide product in the culture may result from hydrolytic release catalyzed by endogenous enzyme in *A. oryzae*. The complete gene cluster of squalestatin is yet to be identified; however, an acyltransferase similar to that of LovD should be expected to catalyze the transfer of the tetraketide to the core structure. The attempt to engineer the SQTKS to generate an unmethylated tetraketide analog by mutating conserved residues in the C-MeT domain, either led to complete loss of product or no effect on production of the tetraketide.[124] It was suggested that the expected unmethylated product in the culture may be degraded by β-oxidation pathway in the host.[124]

The role of HR-PKS FUM1p (encoded by *fum1*) in the biosynthesis of mycotoxin fumonisin was first identified in *G. fujikuroi* (anamorph: *Fusarium verticillioides*).[32,125] The HR-PKS consists of the same domain architecture as LovF and SQTKS, but produces a much longer, dimethylated, and saturated nonaketide (C18) chain (Figure 3). All of the hydroxyls (C3, C5, C14, and C15) on the polyketide chain of **9** are derived from molecular oxygen, except the hydroxyl vicinal to the amino group at C3, which originated from alanine.[126] Therefore, the KR, DH, and ER domains of FUM1p are programmed to completely reduce the β-keto group after every chain extension. Du's group developed a sophisticated gene manipulation system in the fumonisin-producing *F. verticilliodes* to study the HR-PKS by mutagenesis and domain swapping experiments in vivo.[80,81,127] To study the releasing mechanism, a homologue of L-serine palmitoyltransferase (FUM8p) was expressed in yeast, and the isolated microsomes was shown to selectively offloaded the correct C_{18} polyketide chain through condensation with L-alanine in a PLP-dependent reaction (Figure 3).[128] Therefore, the accessory-releasing enzyme FUM8p plays an important role in terminating PKS function and controlling the chain length, which may explain the production of **9** with identical chain length when the KS domain of FUM1p in *F. verticilliodes* is swapped with that from *C. heterostrophus* ChPKS1 of T-toxin pathway.[80]

The FUM1p KS domain was also swapped with that from LovF in an effort to determine the chain length control mechanism of HR-PKSs.[81] The *F. verticilliodes* expressing the chimeric FUM1p produced aromatic dihydroisocoumarins instead of the expected diketide product of LovF. An intriguing possibility proposed by the authors is that these dihydroisocoumarins could be formed from the condensations between one tetraketide and one pentaketide product of the HR-PKS.[81] Thus, these polyketide products from the chimeric FUM1p (shorter than nonaketide but longer than diketide) may suggest that chain length of HR-PKS product is in part controlled by the KS domain, while the downstream domains/enzymes play additional roles. However, this result also illustrates the unpredictable nature of such swapping experiments and further underscores our lack of understanding in these highly programmed machineries.

Several HR-PKSs are known to synthesize linear reduced polyketides as advanced starter units for a collaborating NR-PKS (see later in the Perspective), such as those involved in biosynthesis of RALs,[56,57,129,130] asperfuranone,[43] and azaphilones.[43,131] Similarly, the HR-PKS EasB is involved in the biosynthesis of the linear starter unit for a NRPS to produce emericellamide.[113] The programming rule of one of these HR-

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933−9953

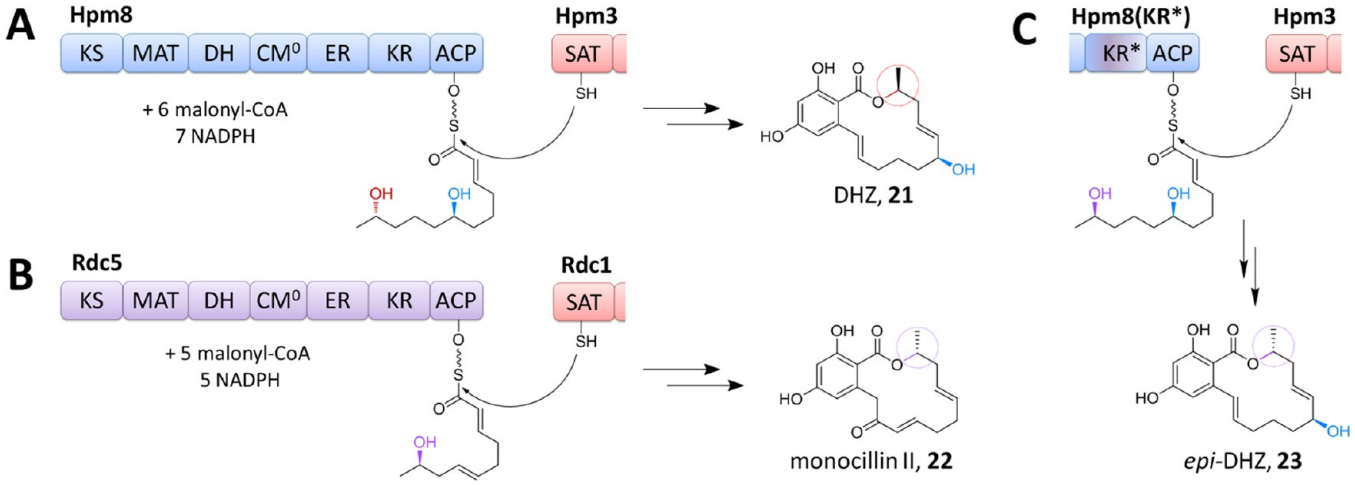


Figure 4. Mapping the substrate-tuned stereospecificity of Hpm8 KR domain. A portion of KR domain in (A) Hpm8 was swapped with that of (B) Rdc5 to generate (C) a chimeric Hpm8(KR*) that produces *epi*-DHZ **23**.

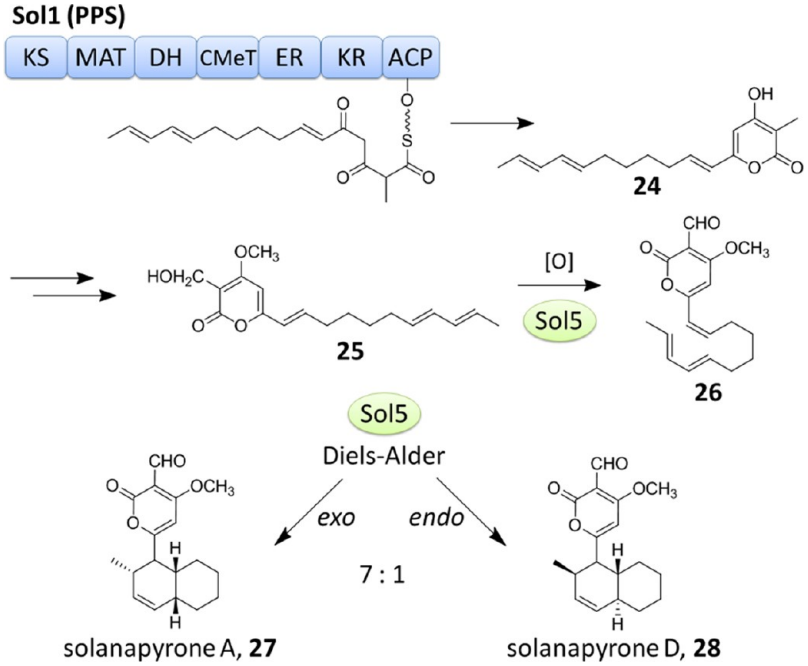


Figure 5. Biosynthesis of solanapyrones in *A. solani*.

PKSs, Hpm8 involved synthesis of the reduced hexaketide portion of hypothemycin **6**,[73] was investigated in detail using heterologous and in vitro reconstitution.[73,75] It was revealed that in addition to the programming rules that control permutative use of the tailoring domains, the KR domain in Hpm8 has another built-in level of complexity that introduce stereochemistry variations at reduced β-carbons. On the basis of the absolute structure of 7′,8′-dehydrozearalenol (DHZ) **21**, the first isolable intermediate of **6**, it could be deduced that the macrolactone ring is derived from a hexaketide chain that contain two alcohols of opposite stereochemistry (Figure 4A). Hence, of the five ketoreduction steps catalyzed by the Hpm8 KR domain during the chain extension, the β-ketoreduction occurred at diketide and tetraketide stage are apparently of opposite stereospecificities. In the bacterial type I modular PKS systems, it is generally established that while KR domains have relaxed substrate selectivity toward different β-keto substrates, the ketoreduction is always catalyzed with stringent stereospecificity to afford a single stereoisomer.[132,133] Thus, the ability of Hpm8 KR domain to exhibit opposite stereospecificities at different chain length is very surprising.

In collaboration with the Vederas group, our laboratory set out to understand this stereochemical puzzle by first assaying the reconstituted Hpm8 PKS with β-ketoacyl-SNAC substrates of various chain lengths. Except for the diketide substrate (acetoacetyl-SNAC), which gives a β-hydroxyl product with the L configuration, the Hpm8-KR catalyzes the β-reduction to give a D-hydroxyl product for all the β-ketoacyl-SNAC substrates from triketide to hexaketide chain lengths. These results verified the substrate-tuned stereospecificity of the Hpm8 KR. In contrast to Hpm8, the same diketide intermediate is reduced with the opposite stereochemistry to give a D-alcohol by the KR domain of Rdc5, the HR-PKS involved in the biosynthesis of a related RAL **22**, which is the intermediate en route to radicicol (Figure 4B).[74] Guided by homology models, the secondary structural elements that contribute to the stereoselectivity

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933−9953

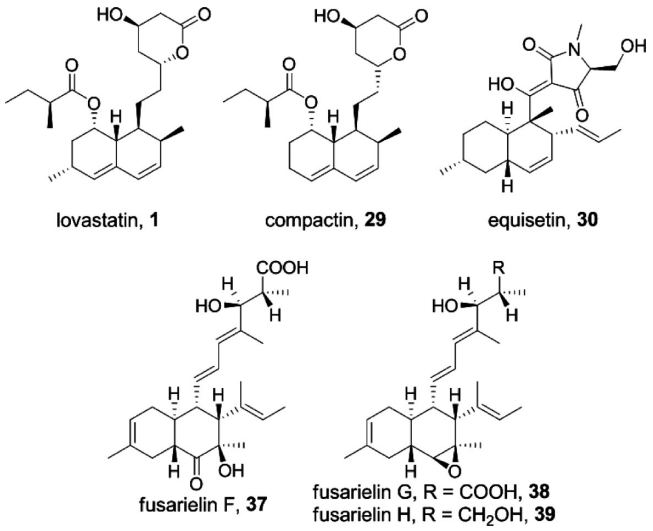
Table 2. Selected Metabolites That Have Been Linked to HR-PKSs and PKS-NRPSs in the Past 5 Years

final product	first isolable product	organisms	genes/locus	PKS	domain architecture	ref
aspyridone A, **44**	preaspyridone	*A. nidulans*	*apdA*/AN8414	ApdA	KS-AT-CMeT-DH-ER^0-KR-ACP-C-A-T-R*	41, 76
pseurotin, **42**	N.D.	*A. fumigatus*	*psoA*/Afu8g00540	PsoA	KS-AT-CMeT-DH-ER^0-KR-ACP-C-A-T-R	111
chaetoglobosin A, **56**	N.D.	*P. expansum*	*cheA*	CheA	KS-AT-CMeT-DH-ER^0-KR-ACP-C-A-T-R	112
emericellamide	N.D.	*A. nidulans*	*easB*/AN2547.3	EasB	KS-AT-CMeT-DH-ER-KR-ACP	113
asperfuranone, **18**	methylbenzaldehyde..	*A. nidulans*	*afoG*/AN1036.3	AfoG	KS-AT-CMeT-DH-ER-KR-ACP	43
hypomycetin, **6**	7′,8′-dehydrozeralenol (DHZ)	*H. subiculosus*	*hpm8*	Hpm8-HR	KS-AT-DH-ER-KR-ACP	56
radicicol	monocillin II	*P. chlamydosporia/ C. chiversii*	*rdc5, ccRads1*	Rdc5-HR	KS-AT-DH-ER-KR-ACP	56, 57
cyclopiazonic acid, **41**	cyclo-acetoacetyl-L-tryptophan (cAATrp)	*A. oryzae/A. flavus*	AO090001000277/ *AFLA_139490/cpaA*	CpaS	KS-AT-CMeT^0-DH^0-ER^0-KR^0-ACP-C-A-T-R*[a]	114, 115
depudecin	N.D.	*Alternaria brassicicola*	*dep5*	DepS	KS-AT-DH-ER-KR-ACP	51
solanapyrone A, **27**	prosolanapyrone I	*Alternaria solani*	*sol1*	PSS	KS-AT-DH-ER-KR-ACP	116
ACT-toxin	N.D.	*Alternaria alternata*	*actts3*	ACTTS3	KS-AT-DH-ER-KR-ACP	117
pyripyropene	4-hydroxy-6-(3-pyridinyl)-2*H*-pyran-2-one	*A. terreus*	*pyr2*/AFUA_6G13930	Pyr2	KS-AT-CMeT^0-DH^0-ER^0-KR^0-ACP[a]	118
isoflavipucine, **46**	N.D.	*A. terreus*	ATEG_00325	ATEG00325	KS-AT-CMeT^0-DH^0-ER^0-KR^0-ACP-C-A-T-R[a]	119
desmethylbassianin A, **45**	preDMB A	*Beauveria bassiana*	*dmbS*	DmbS	KS-AT-CMeT-DH-ER^0-KR-ACP-C-A-T-R*	78, 82
cytochalasin E/K, **4**, **57**	N.D.	*A. clavatus*	*ccsA*/ACLA_078660	CcsA	KS-AT-CMeT-DH-ER^0-KR-ACP-C-A-T-R	120
fusarielin F−H, **37**−**39**	N.D.	*Gibberella zeae*	FGSG_10464/*fsl1*	FSL1/ GzPKS9	KS-AT-CMeT-DH-ER^0-KR-ACP	121

[a]HR-PKSs and PKS-NRPSs that have lost the catalytic functions for all the β-keto reductive domains. N.D., not determined.

difference between the two KR domains were mapped to a $\beta-\alpha-\alpha$ structural motif in the catalytic active site. The $\beta-\alpha-\alpha$ structural motif may contribute to positioning of the β-carbon at the diketide stage for a L configuration β-reduction, while the β-carbon tilted back to the default position for D configuration β-reduction as the chain grows longer. Using a Hpm8/Rdc5 chimeric HR-PKS and the NR-PKS Hpm3, we were able to synthesize the diastereomeric *epi*-DHZ **23** in *S. cerevisiae*.

Solanapyrones **27** and **28** are decalin ring-containing polyketides, but unlike LovB in lovastatin biosynthesis (see next section), formation of the decalin rings in **27** and **28** are not catalyzed by the HR-PKS. The HR-PKS (Sol1), prosolanapyrone synthase (PPS), produces a linear triene prosolanapyrone precursor **24** (Figure 5).[116] Thus, it is not surprising that Sol1 is evolutionary more related to HR-PKSs that synthesizes linear reduced polyketide such as SQTKS and FUM1p, than to LovB or EqiS that produce decalin compounds. A bifunctional flavin-dependent oxidase was identified as the solapyrone synthase (Sol5) that can catalyze both an alcohol dehydrogenation of **25** to form **26** and a [4 + 2] cycloaddition reaction of the intermediate **26** to form the *exo* adduct in solanapyrone A **27** (Figure 5).[116,134,135] HR-PKSs that produce similar methylated and nonmethylated pyrones have also been identified in *A. solani*.[136,137] The release of these pyrones from the HR-PKSs is likely via spontaneous lactonization of a tri-β-keto intermediate, as observed in the in vitro assays of LovB (see below). Other interesting linear polyketides produced by HR-PKSs, including the histone deacetylase inhibitor depudecin[51] and ACT-toxin,[117] are listed in Table 2.



Decalin-Forming HR-PKSs. HR-PKSs are known to produce several decalin-containing products including **1**, compactin **29**, and equisetin **30**. The HR-PKSs in the pathway of **1** (LovB)[30,50] and **29** (MlcB)[138] are highly similar and are terminated with a condensation-like (CON) domain, while that of **30** (EqiS)[139] is fused to an intact NRPS module. All three HR-PKSs contain an inactive ER domain as evident from the lack of consensus active site residues, and recruit a dissociated *trans*-acting ER enzyme encoded in the gene cluster to perform the required enoylreduction steps. Heterologous expression in *A. nidulans* showed that LovB and LovC (the *trans* ER) are required for production of the decalin intermediate **31**.[140]

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933−9953

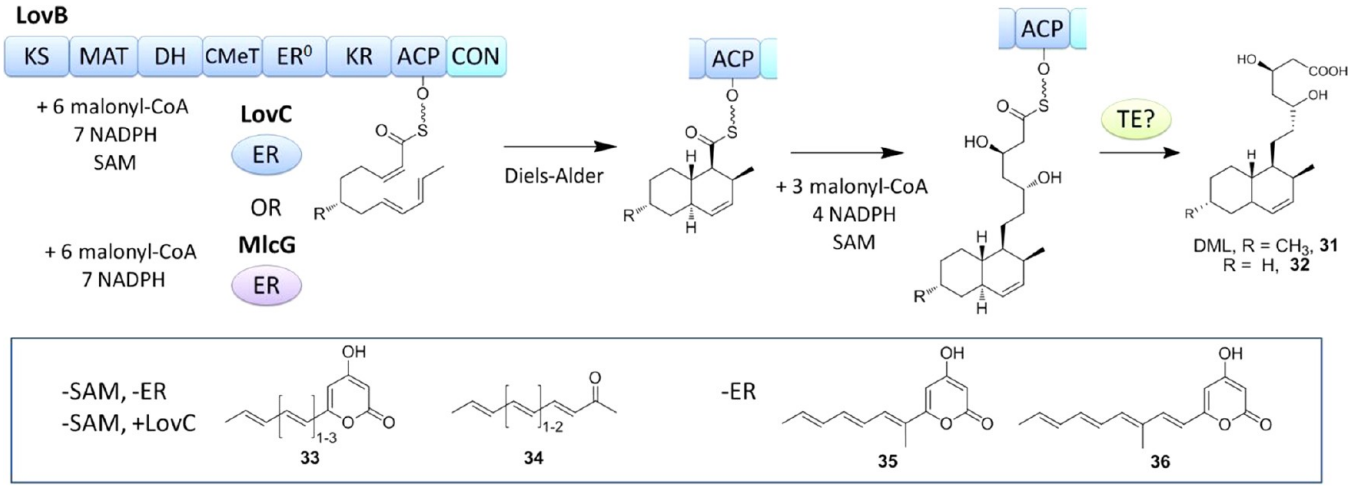


Figure 6. In vitro reconstitution of LovB for production of dihydromonacolin L (DML) acid **31** with LovC and its desmethyl derivative **32** with MlcG. Box: derailment products when SAM or *trans*-ER (LovC/MlcG) or both was omitted from the assays.

Deletion of *lovC* in *A. terreus* resulted in loss of production of **1**, which can be restored by feeding of **31**, demonstrating that the *trans*-ER is indispensable for the biosynthesis of **31**.[141]

Using the protease-deficient strain *S. cerevisiae* host BJ5464:NpgA, which contains a chromosomally integrated copy of the *A. nidulans* phosphopantetheinyl transferase gene *npgA*,[142] the intact 335 kDa LovB was solubly expressed and functionally reconstituted in vitro.[72] The recombinant LovB and LovC (from *E. coli*) are able to synthesize **31** in the presence of malonyl-CoA, NADPH and SAM (Figure 6). However, offloading of the polyketide product from the ACP domain of LovB required base hydrolysis. This suggests that a yet-to-be identified releasing enzyme in *A. terreus* is required for the product release from LovB. The heterologous production of **31** in *A. nidulans* expressing LovB and LovC may therefore due to the hydrolytic activity of endogenous TE enzymes from an unrelated pathway. This hypothesis was corroborated by the findings that the heterologous GfPKS4-TE and GzPKS13-TE can offload the product **31** from LovB in vitro.[72]

On the basis of the in vitro system, further insights on the iterative catalysis and the complex programming of LovB were obtained by selectively omitting the individual cofactors or LovC from the assays.[72] In the absence of NADPH, LovB alone was unable to extend the polyketide chain beyond triketide stage and instead off-loaded the derailed product as triketide lactone (4-hydroxy-6-methyl-2*H*-pyran-2-one), via spontaneous cyclization that is commonly observed for unreduced tri-β-keto thioester. This demonstrates that HR-PKSs such as LovB lack the mechanism to stabilize poly-β-keto chain to prevent lactonization, which is a hallmark of NR-PKSs. Addition of NADPH, thus enabling the KR and subsequently DH functions, resulted in offloading of derailment products (**33** and **34**) by LovB (Figure 6), but no nonaketide product was observed. The lack of methylation and enoylreduction at the tetraketide stage disrupts the normal programmed steps and promotes offloading of polyene products from LovB ACP, which occurs either through α-pyrone formation (**33**, major route) or by hydrolysis and decarboxylation (**34**, minor route). When both SAM and NADPH were added to the reaction in the absence of LovC, methylated hexa- and heptaketide (**35** and **36**) were formed. The position of the methyl groups in these two products, which were also isolated from the *A. nidulans* heterologously expressing only LovB,[50] is consistent with its timing at the tetraketide stage.

Surprisingly, addition of LovC to the reaction in the absence of SAM did not yield the C6-desmethyl derivative of DML **32**. Instead, derailment products formed due to the lack of enoylreduction at the tetraketide stage, such as **33** and **34** were observed. This suggests that LovC has stringent substrate specificity toward the α-methylated tetraketide and can be considered a "proofreader" to ensure the required methylation had occurred. Supporting this notion, in vitro assay of LovB with *trans*-ER MlcG of **29** pathway, which naturally prefers an unmethylated tetraketide intermediate, resulted in production of **32** in the absence of SAM.[72] Unexpectedly, **31** can be obtained when SAM is added to the assay with LovB and MlcG, demonstrating the tolerance of MlcG toward the methylated tetraketide as well.

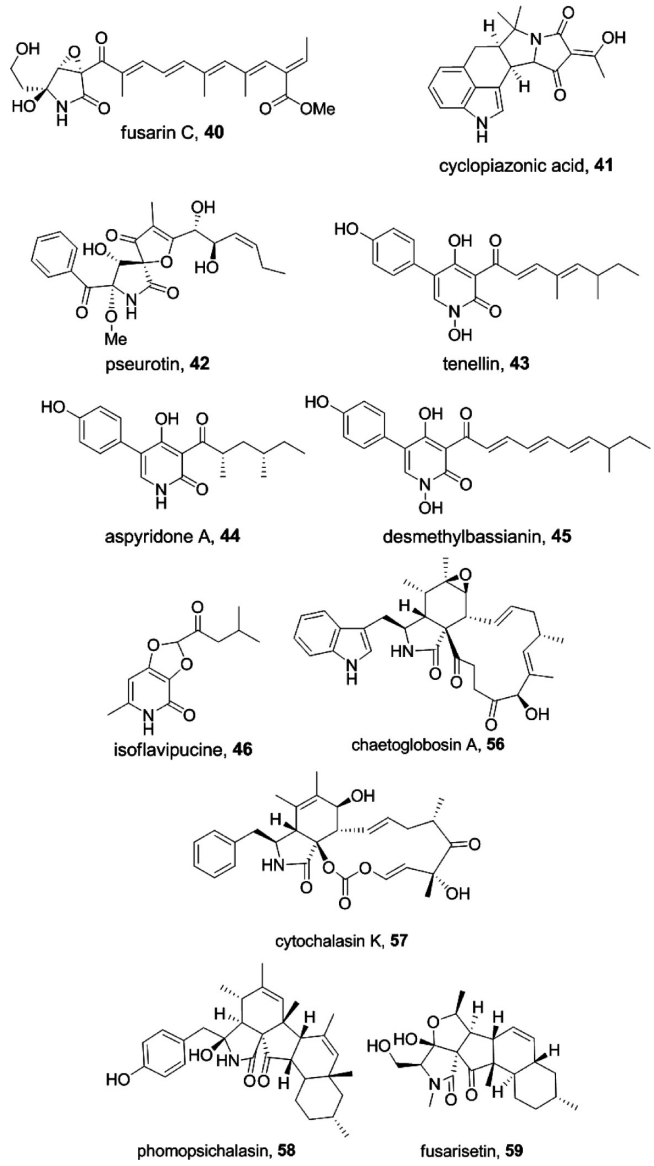
The precise role of LovC in reducing three out of eight possible β-enoyl intermediates during biosynthesis of **31** is important for forming the correct triene hexaketide for the Diels–Alder cyclization (Figure 6). LovC was shown to complex strongly with LovB in gel shift assays and gel filtration.[143,144] Recently, the high resolution crystal structure of LovC at 1.89 Å was solved in collaboration with the Tsai, the Vederas, and our groups, showing a unique medium-chain dehydrogenase/reductase (MDR) fold in the monomeric form (most MDR proteins are dimers).[144] The monomeric form of LovC was proposed to be important for its interaction with LovB. The crystal structure facilitates the mapping of the potential ACP docking site, which is an electropositive surface patch of LovC and may include a flexible, extended loop unique to *trans*-acting ER. Further in silico docking analyses with short to long polyketide intermediates at the active site of LovC provided a possible explanation for the substrate specificity of LovC.[144] The terminal CON domain of LovB remained the most enigmatic as this domain, which initially appears to be unrelated to the iterative functions of the PKS, was shown to be essential for formation of **31**. When the CON domain was removed to yield the truncated LovB (LovB-ΔC), the enzyme was unable to synthesize **31** in the presence of LovC and all the cofactors. In the absence of LovC, LovB-ΔCON produced **35** and **36**, indicating that the chain extension, β-keto processing and C-MeT domains of LovB-ΔCON are fully functional. The ability to synthesize **31** by LovB-ΔCON can be restored by

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933–9953

complementation with the standalone CON domain *in trans*. While the exact role of the CON domain in the synthesis of **31** has not been conclusively established, it is tantalizing to speculate that it is involved in facilitating the *endo*-selective [4 + 2] Diels–Alder cycloaddition at the hexaketide stage (Figure 6).[116,145] The CON domain, with its sequence similarly to canonical condensation (C) domains from NRPS, is proposed to have a sufficiently large active site cavity that can serve a structural role to mold the hexaketide into the correct configuration to yield the otherwise unfavorable decalin regioisomer.

Equisetin **30** is another decalin-containing compound that has been linked to an IPKS.[139] Unlike LovB, however, equisetin synthase (EqiS) is a typical PKS–NRPS hybrid, which contains a complete NRPS module at the C-terminus. The decalin ring system is highly similar to that in **1** and could originate from a similar Diels–Alder reaction as proposed in the biosynthesis of **1**. Besides the decalin ring, **30** also contains a tetramic acid moiety, which is formed by the action of the C-terminal NRPS module, and will be discussed in more detail in the next section. Other than the complete NRPS module, the primary sequence and domain architecture of LovB are otherwise highly similar to EqiS.[139] Very recently, an HR-PKS gene cluster responsible for the biosynthesis of decalin-containing fusarielin F–H (**37**-**39**) was discovered in *Fusarium graminearum*.[121] The decalin-ring system of **37**-**39** is similar to that in **30**, while the HR-PKS, FSL1, is similar to LovB and EqiS (37% and 43% protein identity). Like LovB, the ER domain of FSL1 is inactive and a *trans*-ER (FSL5) is encoded in vicinity of the gene cluster. Intriguingly, FSL1 lacks the C-terminal CON domain found in LovB (Table 2).

PKS-NRPS Hybrids. Fungal PKS–NRPS hybrid mega-synthases are HR-PKSs that contain a complete C-terminal NRPS module (C–A–T–R) including the condensation (C), adenylation (A), thiolation or peptidyl carrier protein (T), and the reduction/Dieckmann cyclase (R) domains. There are also several PKS-NRPS hybrids that have lost the catalytic functions for all the β-keto reductive domains, such as cyclopiazonic acid (**41**) synthase (CpaS)[146] and isoflavipucine **46** PKS–NRPS



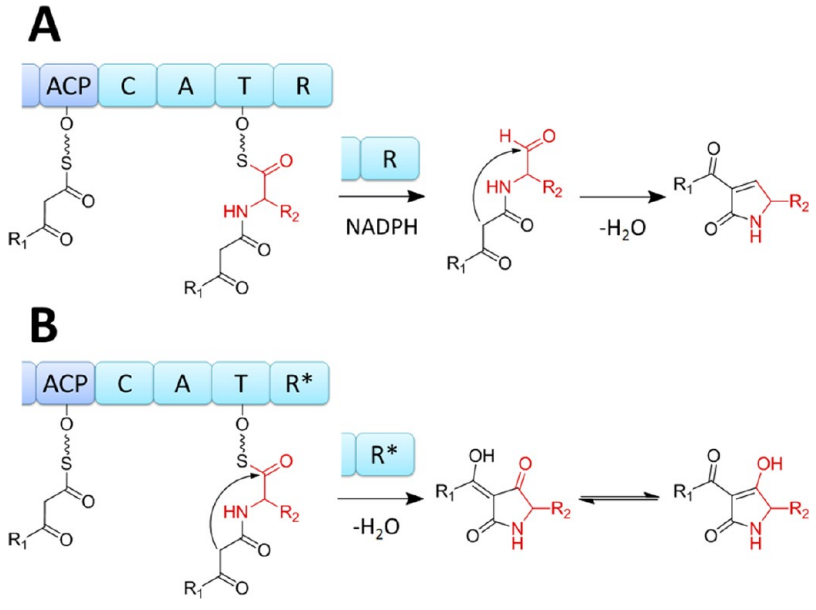


Figure 7. Product release catalyzed by R (reductive release) or R* (Dieckmann cyclization) domain of PKS–NRPSs.

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933–9953

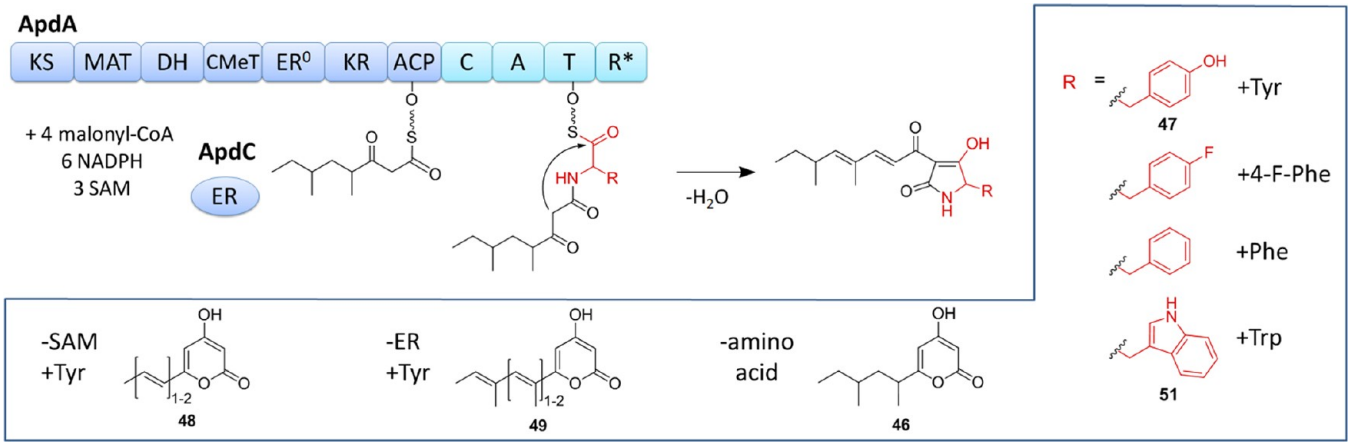


Figure 8. In vitro reconstitution of preaspyridone A **47** biosynthesis. Box: derailment products and analogues of **47** generated from using different amino acids in assays.

(ATEG00325)[119] (Table 2). Selected PKS–NRPSs that have been linked to the metabolites recently are listed in Table 2. As PKS–NRPSs have a built-in releasing mechanism via the R domain, and can form a 2-pyrrolidone or a tetramic acid chromophore that can be easily detected, these enzymes are good models for understanding the complex functions of HR-PKSs. One of the earliest PKS-NRPS that was linked to a product is the fusarin C (**40**) synthase (FUSS).[147] Around the same time, the PKS–NRPS EqiS was linked to the production of **30**.[139]

The C-terminus NRPS component resembles a typical termination module in bacterial multimodular NRPS systems that release a peptidyl aldehyde or alcohol products. The R domain in these NRPS could either catalyze a two-electron reductive release to yield an aldehyde (e.g., gramicidin),[148] or a four electron reductive release to afford a peptidyl alcohol (e.g., lyngbyatoxin).[149] Reductive peptidyl release by an R domain is also observed among fungal NRPSs, most notably among those producing peptaibol antibiotics.[150] Hence the R domain of fungal PKS-NRPS was originally proposed to catalyze the reductive release of the aminoacylated polyketide product as an aldehyde (Figure 7A).[41,147] However, an alternative mechanism was later proposed for the tenellin (**43**) synthase (TENS) R domain.[151] When heterologous expressed in *A. oryzae*, TENS synthesized pretenellin A **48** that contains a 2,4-pyrrolidinedione (tetramic acid) moiety (Figure 9A), which is at a higher oxidation state than it would be catalyzed by the Knoevenagel condensation that affords a 2-pyrrolinone as shown in Figure 7A. This result suggested that the R domain may catalyze a Dieckmann condensation to form the tetramic acid without reducing the aminoacyl bound on the T domain (Figure 7B).[151]

The first direct biochemical evidence that supports this hypothesis comes from the in vitro enzymatic study of the EqiS R domain by the Schmidt group.[152] The recombinant EqiS R domain was shown to catalyze the cyclization of a synthetic acetoacetyl-Ala thioester into a tetramic acid in an NADPH-independent manner (Figure 7B). Further insights into the biochemical basis of the tetramate-forming Dieckmann cyclization were provided by the Walsh group on characterization of the R domain of cyclopiazonic acid (**41**) synthase CpaS.[114,153] which synthesizes the tetramic acid cyclo-acetoacetyl-L-tryptophan.[115] The CpaS R domain was shown to cyclize the T-domain-tethered acetoacetyl-Trp-thioester into the tetramic acid in an NADPH-independent manner.[153] These R domains that do not catalyze reductive release reaction are redesignated as R* domains.[153] Further sequence alignment and mutagenesis experiments in the study showed that the canonical Ser-Tyr-Lys catalytic triad of SDR family of enzymes is not involved, but a conserved aspartic acid residue is important for the catalytic activity by generating the required α-carbanion.[153] Nonetheless, some PKS-NRPSs are likely to have retained a redox active R domain to produce a 2-pyrrolinone intermediate. These include the PKS–NRPSs that synthesize pseurotin **42**,[111] isoflavipucine **46**,[119,154] as well as the cytochalasins (see below). Recently, the R domain of *A. terreus* ATEG00325 that synthesize **46** has been demonstrated to catalyze a reductive release of T domain tethered aminoacyl thioesters as aldehyde in a NADPH-dependent manner.[154] Surprisingly, in the presence of hydrophobic amino acids and free thiols, the reconstituted NRPS module of ATEG00325 can synthesize a wide range of thiopyrazines.

Several 2-pyridone-containing fungal metabolites, such as tenellin **43**,[155] aspyridone **44**,[41] and desmethylbassianin **45**,[82] have been linked to PKS-NRPSs.[41,82,155] The 2-pyridone moiety of tenellin has been shown to be derived from an oxidative ring expansion of a tetramic acid intermediate, pretenellin A **52**, produced by TENS.[156] Similar oxidative ring expansion from a tetramic acid intermediate is likely involved in biosynthesis of **44** and **45**. All of these PKS-NRPSs contain the R* domain that catalyzes Dieckmann cyclization. In our laboratory, insights into the programming of PKS-NRPS are obtained from the in vitro reconstitution of the PKS–NRPS ApdA,[76] involved in biosynthesis of preaspyridone **47** (Figure 8), a tetramic acid intermediate of **44**.[41] ApdA (439 kDA) was expressed from BJ5474-NpgA and together with the *trans*-ER partner ApdC, synthesized **47** in an in vitro assay that contains all the required cofactors and building blocks. ApdA therefore represents one of the largest multidomain megasynthases reconstituted in vitro. ApdC also shows stringent specificity toward the properly methylated substrates as derailed pyrones **48** are observed in the absence of SAM. The study also discovered that the NRPS module of ApdA (ApdA–NRPS) plays an important role in controlling the chain length and the fidelity of the PKS portion of ApdA (ApdA–PKS). Incorrectly processed tetraketide precursors are not accepted by the C domain of ApdA–NRPS and instead recaptured by the PKS module and further elongated and eventually offloaded as α-pyrones **48** and **49**. Omitting the amino acid tyrosine from the

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933–9953

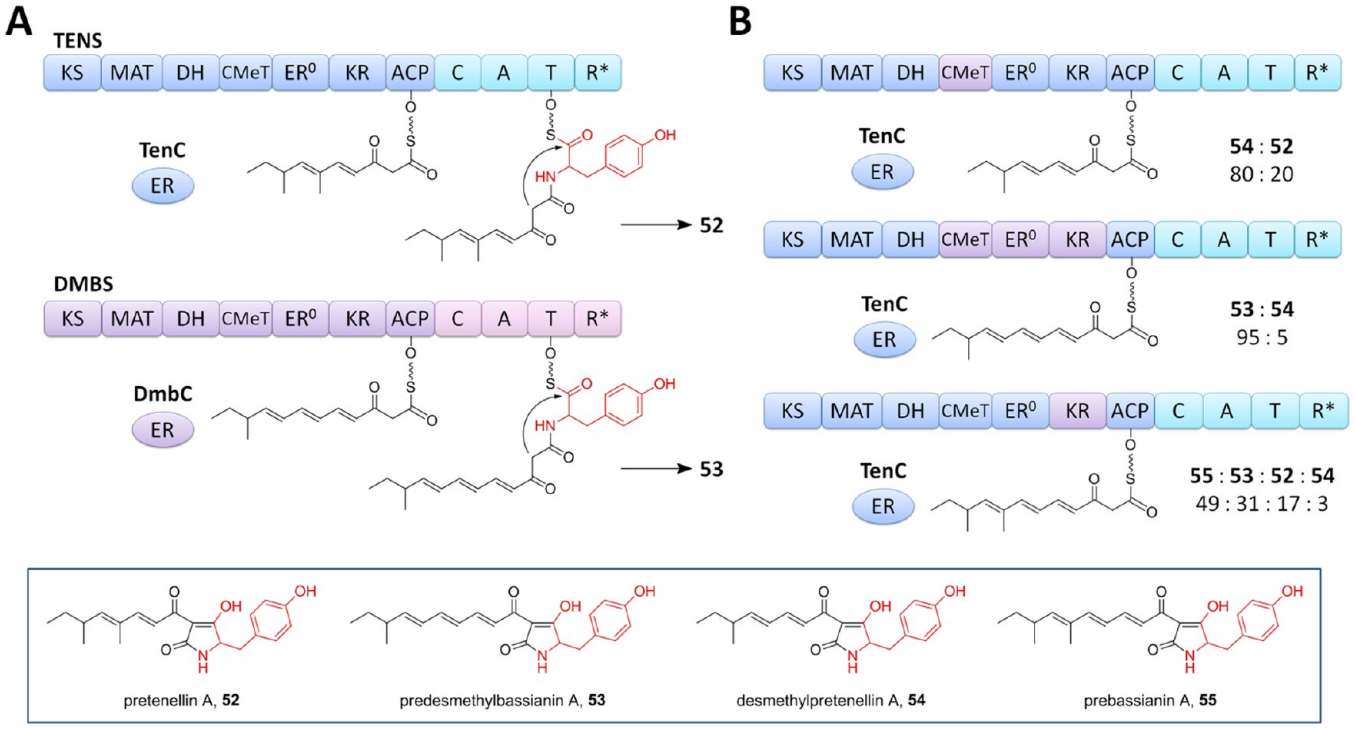


Figure 9. Domain swaps between TENS and DMBS: (A) biosynthesis of **52** and **53** by TENS and DMBS, respectively; (B) minimal domain swaps constructs that significantly altered the product outcome. The product ratio and the polyketide chain of the most abundant product are shown.

assay or using the truncated ApdA–PKS portion only resulted in production of a dimethylated pentaketide pyrone **50**, indicating that the ApdA–PKS can function correctly without the NRPS portion and behaves essentially as a standalone HR-PKS. Although the C domain of ApdA–NRPS has stringent specificity toward the polyketide substrate at the donor site, the acceptor site of the C domain and the A domain appear to have more relaxed substrate specificity. The intact ApdA was able to accept phenylalanine, fluorinated phenylalanine and tryptophan to produce derivatized preaspyridone (Figure 8), albeit at lower yield.[76] The dissection of the PKS and NPRS modules, followed by successful recombination *in trans* to produce **47**, enabled the mix-and-matching of modules from different sources. The CpaS NRPS, which catalyzes the activation and condensation of tryptophan, can weakly complement ApdA–PKS to produce the tryptophan derivative of preaspyridone **51**.[76]

The programming rules of PKS–NRPS are also explored by the Cox group via PKS domain and *trans*-ER swaps between TENS that produces **52** and desmethylbassianin synthase (DMBS) that produces **53** (Figure 9A).[78,82] The TENS and DMBS are two highly similar PKS–NRPSs (87% identity) and likely diverged relatively recently to produce compounds that are different by one chain extension and one methylation. Numerous components of the two PKS–NRPSs are found to be interchangeable, including the *trans*-ER partners. The NRPS modules from both systems can also be interchanged efficiently to produce tetramic acids. The chain lengths of the polyketide portions of the products are not affected by the NRPS swap and remain the same as the native systems.[82] Hence, it appears that the polyketide chain length is solely controlled by the PKS modules of TENS and DMBS, which differs from a more NRPS-controlled mechanism observed in ApdA.

The similarities between TENS and DMBS systems prompted a more systematic domain swapping study aimed at interrogating the programming codes of the PKS modules.[78] These experiments used TENS as the "acceptor" of the different domains "donated" from DMBS, beginning with KS–AT didomain and incorporated one additional domain from the DMBS each time until the entire PKS portion of the TENS has been substituted with DMBS–PKS. This allowed identification and dissection of domains that contribute to the structural differences between **52** and **53.** The chimeric TENS with KS-AT-DH-CMeT from DMBS produces a desmethyl derivative **54** and a lower amount of **52**. This demonstrated that the HR-PKS methylation pattern is controlled by the CMeT domain and was subsequently verified by a single DMBS–CMeT domain swap into TENS (Figure 9B). More importantly, it was found the TENS chimera continued to produce pentaketide **52** efficiently, instead of the hexaketide **53**, until the swapping extended to the KR domain. When the domain swaps extended to KR domain, the chimeric TENS produced solely the **53**, indicating the chain length control mechanism was at least partially, encoded in the KR domain. This was further confirmed in a single KR swap experiment, in which TENS containing the DMBS KR produced four compounds, with the hexaketide **55** as the major compound. It is likely the TENS KR is unable to reduce the β-keto pentaketide intermediate, which is immediately amidated by the C domain and offloaded as **52**. On the other hand, the DMBS KR is able to reduce the β-keto pentaketide intermediate, which is not a substrate of the C domain and is returned to the KS domain for one additional elongation. The resulting β-keto hexaketide is then unreduced and forward to the NRPS module. This unexpected finding that subtle differences in tailoring domain specificity can influence overall polyketide chain length further underscores the complexity of the HR-PKSs, and how surprising, new levels of fidelity control can be uncovered from comparative studies of closely related systems.

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933–9953

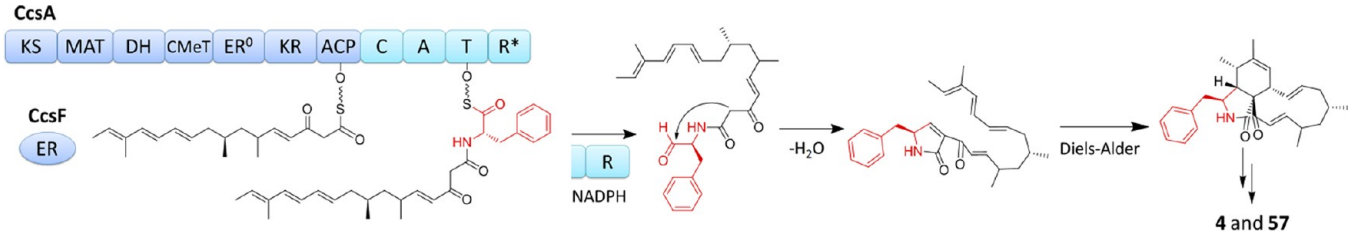


Figure 10. Biosynthesis of cytochalasin E and K via a Diels–Alder cyclization of a pyrrolinone-containing pentaene intermediate.

Another important group of compounds that have been linked to the PKS-NRPSs is the cytochalasins,[112,120] a family of actin inhibitors characterized by the presence of an isoindolone moiety fused with a macrocyclic ring, which can be either a carbocycle, a lactone or a cyclic carbonate (e.g., **4**, **56**, and **57**).[157] It has been proposed that an intramolecular [4 + 2] Diels–Alder cycloaddition between a diene residing in the polyketide portion (as a result of selective tailoring), and the pyrrolinone dienophile is involved in the formation of the characteristic tricyclic isoindolone-fused macrocarbocycle (Figure 10).[158−160] The involvement of a PKS-NRPS (CheA) in biosynthesis of cytochalasins was first demonstrated by RNA-silencing experiment in the chaetoglobosin A **56** producer *Penicillium expansum*.[112] Recently, our laboratory discovered a PKS–NRPS (CcsA)-containing biosynthetic gene cluster in *Aspergillus clavatus* NRRL 1 is involved in the biosynthesis of cytochalasin E **4** and K **57**.[120] The requirement of a pyrrolinone dienophile for the proposed Diels–Alder cycloaddition suggests that the R domain of CheA and CcsA are likely to be redox competent to produce an aminoacyl aldehyde intermediate, which is cyclized via an intramolecular Knoevenagel condensation to form the dienophile (Figures 7A and 10). The Diels–Alder cycloaddition in cytochalasin biosynthesis is expected to be catalyzed by a dedicated enzyme, since biomemetic synthesis of cytochalasin led to mixture of isoindolone stereoisomers derived from both *endo* and *exo* intermediates.[161] Within this group of compounds, there are several structurally intriguing compounds that do not contain the macrocyclic ring,[157] including the pentacyclic phomopsichalasin **58**[162,163] and chaetochalasin A.[164] These compounds are likely to be synthesized via two intramolecular Diels–Alder reactions. Recently, a structurally related compound, fusarisetin A **59**, with acinar morphogenesis inhibitory activity was isolated from *Fusarium* sp. FN080326.[165] Compound **59** features a new pentacyclic scaffold, in which the perhydroisoindolone scaffold typical of cytochalasins is replaced with an oxygen-containing isoindol tricyclic moiety (5/5/5).

■ DUAL PKS SYSTEMS

The biosynthesis of numerous fungal polyketides requires the functions of two PKSs, each responsible for the synthesis of different portions of the molecule. The first such discovery is the LovB-LovF pair in the aforementioned pathway of **1**. The origins of many compounds with a highly reduced polyketide portion connected to an aromatic portion (e.g., **6**, **8**, and **18**) are now known to be products of collaborative effort between an HR-PKS and NR-PKS.[43,73,74] Thus, such division of labor between the HR-PKS and NR-PKS is a clever strategy for Nature to expand the polyketide chemical diversity without the need to introduce further programming into the already complex iterative catalysts. A less-obvious division of labor is in the biosynthesis of T-toxin, in which two PKSs are required for the biosynthesis of the long, contiguous reduced product.[52] The biochemical basis for linking two reduced polyketide chain together to form T-toxin is yet to be elucidated.

In general, the PKS partnerships can be classified into two different modes: (1) the sequential mode, in which an upstream PKS produces a polyketide product that is transferred to a downstream PKS to be further elongated and processed; (2) the convergent mode, where two PKSs function independently in parallel to produce two polyketide products, which are ultimately connected via accessory enzymes. For the sequential PKS-PKS partnership, the biochemical basis of linking a reduced polyketide chain to a nonreducing polyketide has been attributed to the interactions between the upstream HR-PKS ACP domain and the downstream NR-PKS SAT domain.[67,73] The structural basis of this interaction is not known at this point. In the Hpm8-Hpm3 reconstitution study, we showed that a chimeric Hpm3 swapped with a GzPKS13 SAT domain was not able to capture the reduced starter unit from Hpm8, indicating that the GzPKS13 SAT domain may not recognize the noncognate Hpm8-ACP.[73] Similarly, NSAS (PksA) has been shown to complex strongly with the hexanoyl FAS (HexA/HexB) to facilitate the efficient transfer of hexanoyl starter unit to NSAS in the biosynthesis of **15** (Figure 1A).[90] The docking site on NSAS is located on the SAT domain, as confirmed by the in vitro visualization of the acyltransfer between HexA–ACP and PKSA–SAT domain using radiochemical assays.[166] Mapping the protein–protein interactions between ACP and SAT domains will allow us to exploit this fungal version of modularity among PKSs.

In the convergent mode of PKS–PKS partnership, typically an acyltransferase is involved in transferring a polyketide chain synthesized by one HR-PKS to a finished polyketide product synthesized by a different PKS. Protein–protein interactions between the acyltransferase and HR-PKS have also been shown to be important, such as in the case between the acyltransferase LovD and LovF ACP.[122] In the case for the azaphilone (*aza*) pathway discovered recently in *A. niger*,[107] the gene cluster encodes two PKSs that showed significant homology to the HR-PKS (AfoG) and NR-PKS (AfoE), which act sequentially in the asperfuranone pathway.[43] Yet, the two *aza* PKSs form a convergent mode of partnership in generating *O*-acylated azaphilones. An acyltransferase gene (*azaD*) is also present in the *aza* cluster; however, it has not been determined if AzaD is responsible for transfer of the reduced polyketide chain produced by the HR-PKS (AzaB) or other common acyl groups, such as acetyl or benzoyl. Thus, it is still not possible to predict the mode of collaboration between an NR-PKS and HR-PKS encoded in a single gene cluster based on primary sequences. Since dual PKS gene clusters are widely spread in sequenced fungal genomes, elucidation of genetic and molecular basis that account for the different PKS–PKS

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933–9953

partnerships will enhance our abilities in predicting the metabolite structures based on gene sequences.

CONCLUDING REMARKS AND OUTLOOK

Substantial knowledge and insight has been acquired regarding the genetics and enzymology of fungal PKSs compared to as near as five years ago, especially for the NR-PKSs, but there remain many unknowns. The fungal IPKSs used to generate the large diversity of polyketides is derived from a limited number of protein architectures and IPKS gene lineages. Thus, it is expected that we will come to a point in the future where we have identified almost all possible core polyketide scaffolds. Nevertheless, research into fungal IPKSs in the next decade is likely to continue give us new insights and surprises about the catalytic prowess encoded in these megasynthases. A recent example is from the PR-PKS systems, which we are unable to cover in details in this perspective. After 20 years since the 6-MSAS gene was cloned, it was discovered that the conserved domain previously assigned as a DH domain in 6-MSAS is in fact a hydrolase that catalyze the release of the 6-methylsalicyclic acid product.[167] A region (ID domain) important for dimeric interactions between 6MSAS monomer has also been located in an earlier study.[168] Genome-wide analysis of PKS genes in multiple fungi has revealed the presence of genes encode PR-PKSs with a NRPS-like A−T didomain at the N-termini (e.g., *C. heterostrophus* PKS24).[48] These NRPS/PKSs are different from the highly reducing PKS−NRPSs, but are more related to 6MSAS and other bacterial type I modular PKSs. A more recent phylogenetic analysis presented a case for horizontal gene transfer of such NRPS/PKS genes from bacteria and discovered their distribution across different fungal genera.[169] To date, no product from these NRPS/PKSs has been identified, but they are likely to produce partially reduced polyketide products primed with an amino acid.

The recent discoveries of the IPKSs for biosynthesis of **41** (CpaS),[115] **46** (ATEG00325),[119] and pyripyropene (Pyr2)[118] have also challenged our existing paradigm on HR-PKSs and HR-PKS−NRPSs. The KR, DH, and ER domains (and CMeT domain) in these three PKS−NRPSs have evolved to nonfunctional pseudodomains (Table 2). Further surprise comes from the ability of HR-PKS Pyr2 to accept a nicotinyl-CoA starter unit in the biosynthesis of pyripyropene.[118] The selectivity for an aromatic advanced starter unit is unprecedented to both fungal HR-PKSs and NR-PKSs. While the genetic and biochemical basis for starter unit selectivity in NR-PKSs are known, nothing is currently known about the origin of starter unit specificity in HR-PKSs. Hence, it is clear that the diversity of IPKSs in fungi is far from exhaustively explored and may continually give us new surprises.

Sequence and phylogenetic analyses have shown that the current diversity of NR-PKSs and HR-PKSs in the filamentous ascomycetes was predominantly diversified further by point mutation and deletion, rather than by domain recombination within fungi.[48] This is in contrast to the bacterial modular type I PKSs where domain duplication, recombination and horizontal gene transfer dominate the PKS diversification.[170,171] This conclusion has important implications to our overarching goal toward prediction of metabolite structures from primary fungal PKS sequences. Such an evolution model dictates that the gaining or swapping of functional domains in the PKS rarely happens. For example, the gain of a C-terminal NRPS module by the HR-PKSs is believed to happen only once during the early divergence of fungal PKSs.[48] This model, where the IPKS diversification are dominated by mutations and deletion, is also evident in the recent findings in which certain functional domains have evolved to adopt completely new functions (e.g., PT domain, CON domain, and R* domain). Therefore, in the investigation of the enzymology and attempt to devise common programming rules of fungal PKSs, it should be kept in mind that some traits may only be generalized to a small subset of closely related PKSs and should be guided by their evolutionary relationship. For biosynthetic engineering efforts, the implications are that domain swapping between closely related fungal IPKSs will likely be more fruitful, while fine-tuning of the whole megasynthase, perhaps by directed evolution, will be required for swapping between less related IPKSs.

The recent in vitro biochemical characterization of fungal IPKS, as dissociate domains and in intact form, have revolutionized our understanding of their enzymology and mechanisms. Overall, these recent studies have painted a unifying (albeit still fuzzy) picture where these megasynthases do not use a linear set of catalytic instructions; rather, the programmed catalysis is achieved by the intertwined substrate specificities and carefully tuned reaction kinetics embedded in each of the functional domains, and facilitated by allosteric interdomain and protein−protein interactions. Such a model is reflected in the recent study that utilized advanced MS techniques to interrogate the covalent bound intermediates on the NR-PKS NSAS.[102] The catalytic snapshots from MS analysis have provided us a glimpse (based on the abundance of each covalent bound species) of how the kinetics and interdomain interactions may be in play in the biosynthesis of **15**.[102] The programming of HR-PKSs is much more complex and our understanding is still lacking. A similar MS approach for HR-PKSs is expected to shed light into the mechanism of their programmed iterative catalysis. Likewise, bioorganic approaches where thioester substrate mimics of the polyketide intermediates in each extension cycles are used to probe the specificity and reaction kinetics of each functional domain and combination of domains, such as that shown in the recent study of KR domain in Hpm8,[75] will be an increasingly useful tool to understand the programming issue of fungal IPKSs. There also remains a lack of insight into the organization, interactions and structures of domains within the intact megasynthases, which hinders our ability to reprogram the IPKSs effectively. Thus, structural studies by X-ray crystallography, NMR and electron microscopy will continue to play important roles, especially for our understanding of the domains in HR-PKSs and interdomain interactions within intact IPKSs.

In broader aspects, it is important that we continue establish the connections between diverse metabolites and IPKSs. This is particularly exciting with the large in-flux of orphan IPKS genes contributed by fungal genome sequencing and the availability of genome mining tools, which include manipulation of the regulatory networks in the host[12,35,41,44] and whole-pathway heterologous expressions.[54,172] Each of these PKS-metabolite connections will serve as a lighthouse that illuminates the nearby chemical space achievable by related IPKSs and contributes to the overarching goal for accurate prediction of metabolite structures from primary IPKS sequences. There are already several bioinformatics programs being developed with the ultimate goal to predict metabolite structures from PKS and NRPS sequences and whole gene clusters, such as the antiSMASH,[173] SBSPKS,[174] ASMPKS,[175] and PKS/NRPS

dx.doi.org/10.1021/jo301592k | J. Org. Chem. 2012, 77, 9933−9953

analysis.[176] However, most of these programs currently only work well with bacterial secondary metabolite gene sequences that are governed by the colinearity rule; thus, more experimental information is still required to be able to reliably predict the product from fungal PKSs and the associated cluster. Finally, it will be the shared aspirations of biochemists and biomolecular engineers that we will ultimately be able to generate a mathematical or computational model for the programmed iterative catalysis that reflects the actual dynamics of these megasynthases and be able to effectively expand the polyketide chemical space encoded by fungal IPKS to realms that have yet to be explored by Nature.

AUTHOR INFORMATION

Corresponding Author

*E-mail: yitang@ucla.edu.

Notes

The authors declare no competing financial interest.

Biography

Yi Tang received his Ph.D. in Chemical Engineering from Caltech in 2002 under the mentoring of Prof. David Tirrell and did his postdoctoral research with Prof. Chaitan Khosla at Stanford University. He is currently Professor in the Department of Chemical Engineering and Biomolecular Engineering (since 2004) and the Department of Chemistry and Biochemistry (since 2009) at UCLA. His research laboratory is focused on natural product biosynthesis, biocatalysis, and nanobiotechnology.

ACKNOWLEDGMENTS

The collection of work on fungal IPKSs described in this Perspective from our laboratory was performed by a number of outstanding former and current graduate students and postdoctoral fellows. Y.T. is indebted to them for their hard work and dedication. We have had the privilege of working with some of the most brilliant scientists and their research groups on biosynthesis of fungal natural products, most importantly with Prof. John Vederas (University of Alberta) and Prof. Christopher Walsh (Harvard Medical School). We also acknowledge the contributions of other collaborators, including Prof. Pieter Dorrestein (UC San Diego), Prof. Sheryl Tsai (UC Irvine), Prof. Nancy Da Silva (UC Irvine), Prof. Kenji Watanabe (Shizuoka University) Prof. Todd Yeates (UCLA) and Prof. Neil Garg (UCLA). Y.T. thanks Prof. Chaitan Khosla (Stanford University) for introducing him to natural product biosynthesis and Prof. David Tirrell (Caltech) for graduate training at the interface of chemistry and biology. Work in our laboratory on natural product biosynthesis is supported by the NIH, NSF, the Alfred Sloan Foundation, the David and Lucile Packard Foundation, the Camille and Henry Dreyfus Foundation, and funds from UCLA Henry Samueli School of Engineering and Applied Science. The image of the fungus Aspergillus terreus, on the cover, was provided by Professor Lynne Sigler, University of Alberta Microfungus Collection and Herbarium, Edmonton, Alberta, Canada. Thanks goes to Leah Martin-Visscher for her design of the cover.

REFERENCES

(1) Hertweck, C. *Angew. Chem., Int. Ed. Engl.* **2009**, *48*, 4688.
(2) Huffman, J.; Gerber, R.; Du, L. *Biopolymers* **2010**, *93*, 764.
(3) Mobius, N.; Hertweck, C. *Curr. Opin. Plant Biol.* **2009**, *12*, 390.
(4) Bentley, R.; Bennett, J. W. *Annu. Rev. Microbiol.* **1999**, *53*, 411.
(5) Staunton, J.; Weissman, K. J. *Nat. Prod. Rep.* **2001**, *18*, 380.
(6) Birch, A. J.; Donovan, F. W. *Aust. J. Chem.* **1953**, *6*, 360.
(7) Birch, A. J. *Science* **1967**, *156*, 202.
(8) Bentley, R. *Annu. Rev. Biochem.* **1962**, *31*, 589.
(9) Simpson, T. J. *Biosynthesis* **1998**, *195*, 1.
(10) Bassett, E. W.; Tanenbaum, S. W. *Biochim. Biophys. Acta* **1960**, *40*, 535.
(11) Tanenbaum, S. W.; Bassett, E. W. *Biochim. Biophys. Acta* **1962**, *59*, 524.
(12) Gatenbeck, S.; Hermodsson, S. *Acta Chem. Scand.* **1965**, *19*, 65.
(13) Lynen, F.; Tada, M. *Angew. Chem.* **1961**, *73*, 513.
(14) Hopwood, D. A. *Chem. Rev.* **1997**, *97*, 2465.
(15) Malpartida, F.; Hopwood, D. A. *Nature* **1984**, *309*, 462.
(16) Hopwood, D. A.; Bibb, M. J.; Bruton, C. J.; Chater, K. F.; Feitelson, J. S.; Gil, J. È. A. *Trends Biotechnol.* **1983**, *1*, 42.
(17) Cox, R. J. *Org. Biomol. Chem.* **2007**, *5*, 2010.
(18) Brakhage, A. A.; Schroeckh, V. *Fungal Genet. Biol.* **2011**, *48*, 15.
(19) Chiang, Y. M.; Lee, K. H.; Sanchez, J. F.; Keller, N. P.; Wang, C. C. *Nat. Prod. Commun.* **2009**, *4*, 1505.
(20) Muller, R. *Chem. Biol.* **2004**, *11*, 4.
(21) Shen, B. *Curr. Opin. Chem. Biol.* **2003**, *7*, 285.
(22) Weissman, K. J. *Methods Enzymol.* **2009**, *459*, 3.
(23) Funa, N.; Awakawa, T.; Horinouchi, S. *J. Biol. Chem.* **2007**, *282*, 14476.
(24) Seshime, Y.; Juvvadi, P. R.; Fujii, I.; Kitamoto, K. *Biochem. Biophys. Res. Commun.* **2005**, *331*, 253.
(25) Yu, D.; Xu, F.; Zeng, J.; Zhan, J. *IUBMB Life* **2012**, *64*, 285.
(26) Maier, T.; Leibundgut, M.; Ban, N. *Science* **2008**, *321*, 1315.
(27) Walsh, C. T.; Fischbach, M. A. *J. Am. Chem. Soc.* **2010**, *132*, 2469.
(28) Beck, J.; Ripka, S.; Siegner, A.; Schiltz, E.; Schweizer, E. *Eur. J. Biochem.* **1990**, *192*, 487.
(29) Mayorga, M. E.; Timberlake, W. E. *Mol. Gen. Genet.* **1992**, *235*, 205.
(30) Hendrickson, L.; Davis, C. R.; Roach, C.; Nguyen, D. K.; Aldrich, T.; McAda, P. C.; Reeves, C. D. *Chem. Biol.* **1999**, *6*, 429.
(31) Feng, G. H.; Leonard, T. J. *J. Bacteriol.* **1995**, *177*, 6246.
(32) Proctor, R. H.; Desjardins, A. E.; Plattner, R. D.; Hohn, T. M. *Fungal Genet. Biol.* **1999**, *27*, 100.
(33) Bingle, L. E.; Simpson, T. J.; Lazarus, C. M. *Fungal Genet. Biol.* **1999**, *26*, 209.
(34) Nicholson, T. P.; Rudd, B. A.; Dawson, M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. *Chem. Biol.* **2001**, *8*, 157.
(35) Cox, R. J.; Glod, F.; Hurley, D.; Lazarus, C. M.; Nicholson, T. P.; Rudd, B. A.; Simpson, T. J.; Wilkinson, B.; Zhang, Y. *Chem. Commun. (Camb.)* **2004**, 2260.
(36) Shimizu, T.; Kinoshita, H.; Ishihara, S.; Sakai, K.; Nagai, S.; Nihira, T. *Appl. Environ. Microbiol.* **2005**, *71*, 3453.
(37) Grube, M.; Blaha, J. *Mycol. Res.* **2003**, *107*, 1419.
(38) Lee, T.; Yun, S. H.; Hodge, K. T.; Humber, R. A.; Krasnoff, S. B.; Turgeon, G. B.; Yoder, O. C.; Gibson, D. M. *Appl. Microbiol. Biotechnol.* **2001**, *56*, 181.
(39) Schmitt, I.; Martin, M. P.; Kautz, S.; Lumbsch, H. T. *Phytochemistry* **2005**, *66*, 1241.

(40) Sanchez, J. F.; Somoza, A. D.; Keller, N. P.; Wang, C. C. *Nat. Prod. Rep.* **2012**, *29*, 351.
(41) Bergmann, S.; Schumann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. *Nat. Chem. Biol.* **2007**, *3*, 213.
(42) Szewczyk, E.; Chiang, Y. M.; Oakley, C. E.; Davidson, A. D.; Wang, C. C.; Oakley, B. R. *Appl. Environ. Microbiol.* **2008**, *74*, 7607.
(43) Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Keller, N.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2009**, *131*, 2965.
(44) Bok, J. W.; Chiang, Y. M.; Szewczyk, E.; Reyes-Dominguez, Y.; Davidson, A. D.; Sanchez, J. F.; Lo, H. C.; Watanabe, K.; Strauss, J.; Oakley, B. R.; Wang, C. C.; Keller, N. P. *Nat. Chem. Biol.* **2009**, *5*, 462.
(45) Schroeckh, V.; Scherlach, K.; Nutzmann, H. W.; Shelest, E.; Schmidt-Heck, W.; Schuemann, J.; Martin, K.; Hertweck, C.; Brakhage, A. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 14558.
(46) Nayak, T.; Szewczyk, E.; Oakley, C. E.; Osmani, A.; Ukil, L.; Murray, S. L.; Hynes, M. J.; Osmani, S. A.; Oakley, B. R. *Genetics* **2006**, *172*, 1557.
(47) Chooi, Y. H.; Cacho, R.; Tang, Y. *Chem. Biol.* **2010**, *17*, 483.
(48) Kroken, S.; Glass, N. L.; Taylor, J. W.; Yoder, O. C.; Turgeon, B. G. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 15670.
(49) Sanchez, J. F.; Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Ahuja, M.; Elizabeth Oakley, C.; Woo Bok, J.; Keller, N.; Oakley, B. R.; Wang, C. C. *Mol. Biosyst.* **2010**, *6*, 587.
(50) Kennedy, J.; Auclair, K.; Kendrew, S. G.; Park, C.; Vederas, J. C.; Hutchinson, C. R. *Science* **1999**, *284*, 1368.
(51) Wight, W. D.; Kim, K. H.; Lawrence, C. B.; Walton, J. D. *Mol. Plant. Microbe Interact.* **2009**, *22*, 1258.
(52) Baker, S. E.; Kroken, S.; Inderbitzin, P.; Asvarak, T.; Li, B. Y.; Shi, L.; Yoder, O. C.; Turgeon, B. G. *Mol. Plant. Microbe Interact.* **2006**, *19*, 139.
(53) Fujii, I.; Ono, Y.; Tada, H.; Gomi, K.; Ebizuka, Y.; Sankawa, U. *Mol. Gen. Genet.* **1996**, *253*, 1.
(54) Li, Y.; Chooi, Y. H.; Sheng, Y.; Valentine, J. S.; Tang, Y. *J. Am. Chem. Soc.* **2011**, *133*, 15773.
(55) Awakawa, T.; Yokota, K.; Funa, N.; Doi, F.; Mori, N.; Watanabe, H.; Horinouchi, S. *Chem. Biol.* **2009**, *16*, 613.
(56) Reeves, C. D.; Hu, Z.; Reid, R.; Kealey, J. T. *Appl. Environ. Microbiol.* **2008**, *74*, 5121.
(57) Wang, S.; Xu, Y.; Maine, E. A.; Wijeratne, E. M.; Espinosa-Artiles, P.; Gunatilaka, A. A.; Molnar, I. *Chem. Biol.* **2008**, *15*, 1328.
(58) Chooi, Y. H.; Wang, P.; Fang, J.; Li, Y.; Wu, K.; Tang, Y. *J. Am. Chem. Soc.* **2012**, *134*, 9428.
(59) Lo, H. C.; Entwistle, R.; Guo, C. J.; Ahuja, M.; Szewczyk, E.; Hung, J. H.; Chiang, Y. M.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2012**, *134*, 4709.
(60) Itoh, T.; Tokunaga, K.; Radhakrishnan, E. K.; Fujii, I.; Abe, I.; Ebizuka, Y.; Kushiro, T. *ChemBioChem* **2012**, *13*, 1132.
(61) Hansen, B. G.; Salomonsen, B.; Nielsen, M. T.; Nielsen, J. B.; Hansen, N. B.; Nielsen, K. F.; Regueira, T. B.; Nielsen, J.; Patil, K. R.; Mortensen, U. H. *Appl. Environ. Microbiol.* **2011**, *77*, 3044.
(62) Regueira, T. B.; Kildegaard, K. R.; Hansen, B. G.; Mortensen, U. H.; Hertweck, C.; Nielsen, J. *Appl. Environ. Microbiol.* **2011**, *77*, 3035.
(63) Ahuja, M.; Chiang, Y. M.; Chang, S. L.; Praseuth, M. B.; Entwistle, R.; Sanchez, J. F.; Lo, H. C.; Yeh, H. H.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2012**, *134*, 8212.
(64) Davison, J.; Al Fahad, A.; Cai, M.; Song, Z.; Yehia, S. Y.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 7642.
(65) Nakazawa, T.; Ishiuchi, K.; Praseuth, A.; Noguchi, H.; Hotta, K.; Watanabe, K. *ChemBioChem* **2012**, *13*, 855.
(66) Sattely, E. S.; Fischbach, M. A.; Walsh, C. T. *Nat. Prod. Rep.* **2008**, *25*, 757.
(67) Crawford, J. M.; Dancy, B. C.; Hill, E. A.; Udwary, D. W.; Townsend, C. A. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 16728.
(68) Crawford, J. M.; Thomas, P. M.; Scheerer, J. R.; Vagstad, A. L.; Kelleher, N. L.; Townsend, C. A. *Science* **2008**, *320*, 243.
(69) Ma, Y.; Smith, L. H.; Cox, R. J.; Beltran-Alvarez, P.; Arthur, C. J.; Simpson, F. R. S. T. *ChemBioChem* **2006**, *7*, 1951.
(70) Ma, S. M.; Zhan, J.; Watanabe, K.; Xie, X.; Zhang, W.; Wang, C. C.; Tang, Y. *J. Am. Chem. Soc.* **2007**, *129*, 10642.
(71) Zhou, H.; Zhan, J.; Watanabe, K.; Xie, X.; Tang, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 6249.
(72) Ma, S. M.; Li, J. W.; Choi, J. W.; Zhou, H.; Lee, K. K.; Moorthie, V. A.; Xie, X.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C.; Tang, Y. *Science* **2009**, *326*, 589.
(73) Zhou, H.; Qiao, K.; Gao, Z.; Meehan, M. J.; Li, J. W.; Zhao, X.; Dorrestein, P. C.; Vederas, J. C.; Tang, Y. *J. Am. Chem. Soc.* **2010**, *132*, 4530.
(74) Zhou, H.; Qiao, K.; Gao, Z.; Vederas, J. C.; Tang, Y. *J. Biol. Chem.* **2010**, *285*, 41412.
(75) Zhou, H.; Gao, Z.; Qiao, K.; Wang, J.; Vederas, J. C.; Tang, Y. *Nat. Chem. Biol.* **2012**, *8*, 331.
(76) Xu, W.; Cai, X.; Jung, M. E.; Tang, Y. *J. Am. Chem. Soc.* **2010**, *132*, 13604.
(77) Watanabe, A.; Ebizuka, Y. *Chem. Biol.* **2004**, *11*, 1101.
(78) Fisch, K. M.; Bakeer, W.; Yakasai, A. A.; Song, Z.; Pedrick, J.; Wasil, Z.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. *J. Am. Chem. Soc.* **2011**, *133*, 16635.
(79) Liu, T.; Chiang, Y. M.; Somoza, A. D.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2011**, *133*, 13314.
(80) Zhu, X.; Yu, F.; Bojja, R. S.; Zaleta-Rivera, K.; Du, L. *J. Ind. Microbiol. Biotechnol.* **2006**, *33*, 859.
(81) Zhu, X.; Yu, F.; Li, X. C.; Du, L. *J. Am. Chem. Soc.* **2007**, *129*, 36.
(82) Heneghan, M. N.; Yakasai, A. A.; Williams, K.; Kadir, K. A.; Wasil, Z.; Bakeer, W.; Fisch, K. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J.; Lazarus, C. M. *Chem. Sci.* **2011**, *2*, 972.
(83) Watanabe, A.; Ebizuka, Y. *Tetrahedron Lett.* **2002**, *43*, 843.
(84) Crawford, J. M.; Korman, T. P.; Labonte, J. W.; Vagstad, A. L.; Hill, E. A.; Kamari-Bidkorpeh, O.; Tsai, S. C.; Townsend, C. A. *Nature* **2009**, *461*, 1139.
(85) Korman, T. P.; Crawford, J. M.; Labonte, J. W.; Newman, A. G.; Wong, J.; Townsend, C. A.; Tsai, S. C. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 6246.
(86) Wattana-amorn, P.; Williams, C.; Ploskon, E.; Cox, R. J.; Simpson, T. J.; Crosby, J.; Crump, M. P. *Biochemistry* **2010**, *49*, 2186.
(87) Crawford, J. M.; Townsend, C. A. *Nat. Rev. Microbiol.* **2010**, *8*, 879.
(88) Yabe, K.; Nakajima, H. *Appl. Microbiol. Biotechnol.* **2004**, *64*, 745.
(89) Yu, J.; Chang, P. K.; Ehrlich, K. C.; Cary, J. W.; Bhatnagar, D.; Cleveland, T. E.; Payne, G. A.; Linz, J. E.; Woloshuk, C. P.; Bennett, J. W. *Appl. Environ. Microbiol.* **2004**, *70*, 1253.
(90) Watanabe, C. M.; Townsend, C. A. *Chem. Biol.* **2002**, *9*, 981.
(91) Watanabe, C. M.; Wilson, D.; Linz, J. E.; Townsend, C. A. *Chem. Biol.* **1996**, *3*, 463.
(92) Udwary, D. W.; Merski, M.; Townsend, C. A. *J. Mol. Biol.* **2002**, *323*, 585.
(93) Fujii, I.; Watanabe, A.; Sankawa, U.; Ebizuka, Y. *Chem. Biol.* **2001**, *8*, 189.
(94) Fujii, I.; Mori, Y.; Watanabe, A.; Kubo, Y.; Tsuji, G.; Ebizuka, Y. *Biosci. Biotechnol. Biochem.* **1999**, *63*, 1445.
(95) Choquer, M.; Dekkers, K. L.; Chen, H. Q.; Cao, L.; Ueng, P. P.; Daub, M. E.; Chung, K. R. *Mol. Plant. Microbe Interact.* **2005**, *18*, 468.
(96) Crawford, J. M.; Vagstad, A. L.; Whitworth, K. P.; Ehrlich, K. C.; Townsend, C. A. *ChemBioChem* **2008**, *9*, 1019.
(97) Winssinger, N.; Barluenga, S. *Chem. Commun. (Cambridge)* **2007**, 22.
(98) Zhang, W.; Li, Y.; Tang, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 20683.
(99) Li, Y.; Xu, W.; Tang, Y. *J. Biol. Chem.* **2010**, *285*, 22764.
(100) Zhou, H.; Li, Y.; Tang, Y. *Nat. Prod. Rep.* **2010**, *27*, 839.
(101) Du, L.; Lou, L. *Nat. Prod. Rep.* **2010**, *27*, 255.
(102) Vagstad, A. L.; Bumpus, S. B.; Belecki, K.; Kelleher, N. L.; Townsend, C. A. *J. Am. Chem. Soc.* **2012**, *134*, 6865.
(103) Gokhale, R. S.; Hunziker, D.; Cane, D. E.; Khosla, C. *Chem. Biol.* **1999**, *6*, 117.

(104) Wang, M.; Zhou, H.; Wirz, M.; Tang, Y.; Boddy, C. N. *Biochemistry* **2009**, *48*, 6288.
(105) Bailey, A. M.; Cox, R. J.; Harley, K.; Lazarus, C. M.; Simpson, T. J.; Skellam, E. *Chem. Commun. (Cambridge)* **2007**, 4053.
(106) Fisch, K. M.; Skellam, E.; Ivison, D.; Cox, R. J.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J. *Chem. Commun. (Cambridge)* **2010**, *46*, 5331.
(107) Zabala, A. O.; Xu, W.; Chooi, Y. H.; Tang, Y. *Chem. Biol.* **2012**, *19*, 1049.
(108) Ishiuchi, K.; Nakazawa, T.; Ookuma, T.; Sugimoto, S.; Sato, M.; Tsunematsu, Y.; Ishikawa, N.; Noguchi, H.; Hotta, K.; Moriya, H.; Watanabe, K. *ChemBioChem* **2012**, *13*, 846.
(109) Thompson, T. B.; Katayama, K.; Watanabe, K.; Hutchinson, C. R.; Rayment, I. *J. Biol. Chem.* **2004**, *279*, 37956.
(110) Zhang, W.; Watanabe, K.; Wang, C. C.; Tang, Y. *J. Biol. Chem.* **2007**, *282*, 25717.
(111) Maiya, S.; Grundmann, A.; Li, X.; Li, S. M.; Turner, G. *ChemBioChem* **2007**, *8*, 1736.
(112) Schumann, J.; Hertweck, C. *J. Am. Chem. Soc.* **2007**, *129*, 9564.
(113) Chiang, Y. M.; Szewczyk, E.; Nayak, T.; Davidson, A. D.; Sanchez, J. F.; Lo, H. C.; Ho, W. Y.; Simityan, H.; Kuo, E.; Praseuth, A.; Watanabe, K.; Oakley, B. R.; Wang, C. C. *Chem. Biol.* **2008**, *15*, 527.
(114) Tokuoka, M.; Seshime, Y.; Fujii, I.; Kitamoto, K.; Takahashi, T.; Koyama, Y. *Fungal Genet. Biol.* **2008**, *45*, 1608.
(115) Seshime, Y.; Juvvadi, P. R.; Tokuoka, M.; Koyama, Y.; Kitamoto, K.; Ebizuka, Y.; Fujii, I. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 3288.
(116) Kasahara, K.; Miyamoto, T.; Fujimoto, T.; Oguri, H.; Tokiwano, T.; Oikawa, H.; Ebizuka, Y.; Fujii, I. *ChemBioChem* **2010**, *11*, 1245.
(117) Miyamoto, Y.; Masunaka, A.; Tsuge, T.; Yamamoto, M.; Ohtani, K.; Fukumoto, T.; Gomi, K.; Peever, T. L.; Tada, Y.; Ichimura, K.; Akimitsu, K. *Mol. Plant. Microbe Interact.* **2010**, *23*, 406.
(118) Itoh, T.; Tokunaga, K.; Matsuda, Y.; Fujii, I.; Abe, I.; Ebizuka, Y.; Kushiro, T. *Nat. Chem.* **2010**, *2*, 858.
(119) Gressler, M.; Zaehle, C.; Scherlach, K.; Hertweck, C.; Brock, M. *Chem. Biol.* **2011**, *18*, 198.
(120) Qiao, K.; Chooi, Y. H.; Tang, Y. *Metab. Eng.* **2011**, *13*, 723.
(121) Sorensen, J. L.; Hansen, F. T.; Sondergaard, T. E.; Staerk, D.; Lee, T. V.; Wimmer, R.; Klitgaard, L. G.; Purup, S.; Giese, H.; Frandsen, R. J. *Environ. Microbiol.* **2012**, *14*, 1159.
(122) Xie, X.; Meehan, M. J.; Xu, W.; Dorrestein, P. C.; Tang, Y. *J. Am. Chem. Soc.* **2009**, *131*, 8388.
(123) Dawson, M. J.; Farthing, J. E.; Marshall, P. S.; Middleton, R. F.; O'Neill, M. J.; Shuttleworth, A.; Stylli, C.; Tait, R. M.; Taylor, P. M.; Wildman, H. G.; et al. *J. Antibiot. (Tokyo)* **1992**, *45*, 639.
(124) Skellam, E. J.; Hurley, D.; Davison, J.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. *Mol. Biosyst.* **2010**, *6*, 680.
(125) Desjardins, A. E.; Munkvold, G. P.; Plattner, R. D.; Proctor, R. H. *Mol. Plant. Microbe Interact.* **2002**, *15*, 1157.
(126) Caldas, E. D.; Sadilkova, K.; Ward, B. L.; Jones, A. D.; Winter, C. K.; Gilchrist, D. G. *J. Agric. Food. Chem.* **1998**, *46*, 4734.
(127) Yu, F.; Zhu, X.; Du, L. *FEMS Microbiol. Lett.* **2005**, *248*, 257.
(128) Gerber, R.; Lou, L.; Du, L. *J. Am. Chem. Soc.* **2009**, *131*, 3148.
(129) Gaffoor, I.; Trail, F. *Appl. Environ. Microbiol.* **2006**, *72*, 1793.
(130) Kim, Y. T.; Lee, Y. R.; Jin, J.; Han, K. H.; Kim, H.; Kim, J. C.; Lee, T.; Yun, S. H.; Lee, Y. W. *Mol. Microbiol.* **2005**, *58*, 1102.
(131) Somoza, A. D.; Lee, K. H.; Chiang, Y. M.; Oakley, B. R.; Wang, C. C. *Org. Lett.* **2012**, *14*, 972.
(132) Siskos, A. P.; Baerga-Ortiz, A.; Bali, S.; Stein, V.; Mamdani, H.; Spiteller, D.; Popovic, B.; Spencer, J. B.; Staunton, J.; Weissman, K. J.; Leadlay, P. F. *Chem. Biol.* **2005**, *12*, 1145.
(133) Valenzano, C. R.; Lawson, R. J.; Chen, A. Y.; Khosla, C.; Cane, D. E. *J. Am. Chem. Soc.* **2009**, *131*, 18501.
(134) Katayama, K.; Kobayashi, T.; Oikawa, H.; Honma, M.; Ichihara, A. *Biochim. Biophys. Acta* **1998**, *1384*, 387.
(135) Katayama, K.; Kobayashi, T.; Chijimatsu, M.; Ichihara, A.; Oikawa, H. *Biosci. Biotechnol. Biochem.* **2008**, *72*, 604.
(136) Fujii, I.; Yoshida, N.; Shimomaki, S.; Oikawa, H.; Ebizuka, Y. *Chem. Biol.* **2005**, *12*, 1301.
(137) Kasahara, K.; Fujii, I.; Oikawa, H.; Ebizuka, Y. *ChemBioChem* **2006**, *7*, 920.
(138) Abe, Y.; Suzuki, T.; Ono, C.; Iwamoto, K.; Hosobuchi, M.; Yoshikawa, H. *Mol. Genet. Genomics* **2002**, *267*, 636.
(139) Sims, J. W.; Fillmore, J. P.; Warner, D. D.; Schmidt, E. W. *Chem. Commun. (Cambridge)* **2005**, 186.
(140) Sorensen, J. L.; Vederas, J. C. *Chem. Commun. (Cambridge)* **2003**, 1492.
(141) Sorensen, J. L.; Auclair, K.; Kennedy, J.; Hutchinson, C. R.; Vederas, J. C. *Org. Biomol. Chem.* **2003**, *1*, 50.
(142) Lee, K. K.; Da Silva, N. A.; Kealey, J. T. *Anal. Biochem.* **2009**, *394*, 75.
(143) Auclair, K.; Sutherland, A.; Kennedy, J.; Witter, D. J.; Van den Heever, J. P.; Hutchinson, C. R.; Vederas, J. C. *J. Am. Chem. Soc.* **2000**, *122*, 11519.
(144) Ames, B. D.; Nguyen, C.; Smith, P. T.; Xu, W.; Ma, S. M.; Wong, E.; Wong, S.; Xie, X.; Li, J. W.; Vederas, J. C.; Tang, Y.; Tsai, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 11144.
(145) Moore, R. N.; Bigam, G.; Chan, J. K.; Hogg, A. M.; Nakashima, T. T.; Vederas, J. C. *J. Am. Chem. Soc.* **1985**, *107*, 3694.
(146) Chang, P. K.; Horn, B. W.; Dorner, J. W. *Fungal Genet. Biol.* **2009**, *46*, 176.
(147) Song, Z.; Cox, R. J.; Lazarus, C. M.; Simpson, T. J. *ChemBioChem* **2004**, *5*, 1196.
(148) Kessler, N.; Schuhmann, H.; Morneweg, S.; Linne, U.; Marahiel, M. A. *J. Biol. Chem.* **2004**, *279*, 7413.
(149) Read, J. A.; Walsh, C. T. *J. Am. Chem. Soc.* **2007**, *129*, 15762.
(150) Daniel, J. F.; Filho, E. R. *Nat. Prod. Rep.* **2007**, *24*, 1128.
(151) Halo, L. M.; Marshall, J. W.; Yakasai, A. A.; Song, Z.; Butts, C. P.; Crump, M. P.; Heneghan, M.; Bailey, A. M.; Simpson, T. J.; Lazarus, C. M.; Cox, R. J. *ChemBioChem* **2008**, *9*, 585.
(152) Sims, J. W.; Schmidt, E. W. *J. Am. Chem. Soc.* **2008**, *130*, 11149.
(153) Liu, X.; Walsh, C. T. *Biochemistry* **2009**, *48*, 8746.
(154) Qiao, K.; Zhou, H.; Xu, W.; Zhang, W.; Garg, N.; Tang, Y. *Org. Lett.* **2011**, *13*, 1758.
(155) Eley, K. L.; Halo, L. M.; Song, Z.; Powles, H.; Cox, R. J.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J. *ChemBioChem* **2007**, *8*, 289.
(156) Halo, L. M.; Heneghan, M. N.; Yakasai, A. A.; Song, Z.; Williams, K.; Bailey, A. M.; Cox, R. J.; Lazarus, C. M.; Simpson, T. J. *J. Am. Chem. Soc.* **2008**, *130*, 17988.
(157) Scherlach, K.; Boettger, D.; Remme, N.; Hertweck, C. *Nat. Prod. Rep.* **2010**, *27*, 869.
(158) Merifield, E.; Thomas, E. J. *J. Chem. Soc., Perkin Trans. 1* **1999**, 3269.
(159) Thomas, E. J. *Acc. Chem. Res.* **1991**, *24*, 229.
(160) Oikawa, H.; Murakami, Y.; Ichihara, A. *J. Chem. Soc., Perkin Trans. 1* **1992**, 2955.
(161) Kelly, W. L. *Org. Biomol. Chem.* **2008**, *6*, 4483.
(162) Horn, W. S.; Simmonds, M. S. J.; Schwartz, R. E.; Blaney, W. M. *Tetrahedron* **1995**, *51*, 3969.
(163) Brown, S. G.; Jansma, M. J.; Hoye, T. R. *J. Nat. Prod.* **2012**, DOI: 10.1021/np300248w.
(164) Oh, H.; Swenson, D. C.; Gloer, J. B.; Wicklow, D. T.; Dowd, P. F. *Tetrahedron Lett.* **1998**, *39*, 7633.
(165) Jang, J. H.; Asami, Y.; Jang, J. P.; Kim, S. O.; Moon, D. O.; Shin, K. S.; Hashizume, D.; Muroi, M.; Saito, T.; Oh, H.; Kim, B. Y.; Osada, H.; Ahn, J. S. *J. Am. Chem. Soc.* **2011**, *133*, 6865.
(166) Foulke-Abel, J.; Townsend, C. A. *ChemBioChem* **2012**, DOI: 10.1002/cbic.201200267.
(167) Moriguchi, T.; Kezuka, Y.; Nonaka, T.; Ebizuka, Y.; Fujii, I. *J. Biol. Chem.* **2010**, *285*, 15637.
(168) Moriguchi, T.; Ebizuka, Y.; Fujii, I. *ChemBioChem* **2008**, *9*, 1207.
(169) Lawrence, D. P.; Kroken, S.; Pryor, B. M.; Arnold, A. E. *PLoS One* **2011**, *6*, e28231.
(170) Jenke-Kodama, H.; Sandmann, A.; Muller, R.; Dittmann, E. *Mol. Biol. Evol.* **2005**, *22*, 2027.

(171) Ridley, C. P.; Lee, H. Y.; Khosla, C. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4595.
(172) Heneghan, M. N.; Yakasai, A. A.; Halo, L. M.; Song, Z.; Bailey, A. M.; Simpson, T. J.; Cox, R. J.; Lazarus, C. M. *ChemBioChem* **2010**, *11*, 1508.
(173) Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R. *Nucleic Acids Res.* **2011**, *39*, W339.
(174) Anand, S.; Prasad, M. V.; Yadav, G.; Kumar, N.; Shehara, J.; Ansari, M. Z.; Mohanty, D. *Nucleic Acids Res.* **2010**, *38*, W487.
(175) Tae, H.; Kong, E. B.; Park, K. *BMC Bioinformatics* **2007**, *8*, 327.
(176) Bachmann, B. O.; Ravel, J. *Methods Enzymol.* **2009**, *458*, 181.