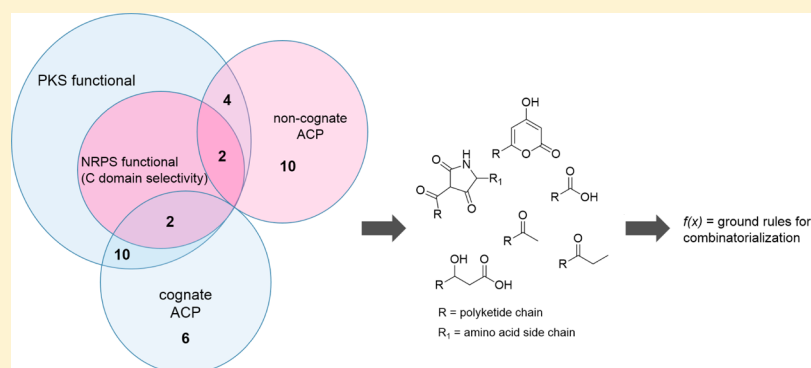# Combinatorialization of Fungal Polyketide Synthase–Peptide Synthetase Hybrid Proteins

Thomas B. Kakule, Zhenjian Lin, and Eric W. Schmidt*

Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, United States

S *Supporting Information*

ABSTRACT: The programming of the fungal polyketide synthase (PKS) is quite complex, with a simple domain architecture leading to elaborate products. An additional level of complexity has been found within PKS-based pathways where the PKS is fused to a single module nonribosomal peptide synthetase (NRPS) to synthesize polyketides conjugated to amino acids. Here, we sought to understand the communication between these modules that enable correct formation of polyketide-peptide hybrid products. To do so, we fused together the genes that are responsible for forming five highly chemically diverse fungal natural products in a total of 57 different combinations, comprising 34 distinct module swaps. Gene fusions were formed with the idea of testing the connection and compatibility of the PKS and NRPS modules mediated by the acyl carrier protein (ACP), condensation (C) and ketoreductase (KR) domains. The resulting recombinant gene fusions were analyzed in a high-yielding expression platform to avail six new compounds, including the first successful fusion between a PKS and NRPS that make highly divergent products, and four previously reported molecules. Our results show that C domains are highly selective for a subset of substrates. We discovered that within the highly reducing (hr) PKS class, noncognate ACPs of closely related members complement PKS function. We intercepted a pre-Diels–Alder intermediate in lovastatin synthesis for the first time, shedding light on this canonical fungal biochemical reaction. The results of these experiments provide a set of ground rules for the successful engineering of hr-PKS and PKS-NRPS products in fungi.

INTRODUCTION

Fungal polyketide synthase (PKS)–nonribosomal peptide synthetase (NRPS) hybrid proteins synthesize a diverse array of biomedically and agriculturally important natural products.[1,2] These include the prescribed anticholesterol drug lovastatin, the biologically useful cytochalasins,[3] compounds such as macrocidins[4] that are critical in fungal pathogenesis of plants, and many other agents. The PKS-NRPS hybrid products are among the most common bioactive compounds isolated from filamentous fungi. The first fungal PKS-NRPS to be characterized was the fusarin C synthetase (FUSS),[5] and since that time many related genes have been identified. The many known fungal PKS-NRPS genes are very similar to each other, but their chemical products are not. It has thus been of great interest to learn the biochemical rules governing product formation and to exploit these rules in synthesizing new derivatives via genetic engineering.[6–8] However, success in these endeavors has so far been quite limited.

Fungal PKS-NRPS products are produced via the action of two modules. A single type I, highly reducing (hr) PKS module acts iteratively to synthesize a complex polyketide core.[1] In fungi, iterative hr-PKS proteins exert exquisite control over regiochemistry, synthesizing polyketides with variable reduction at each acetate extension step despite using the same set of reductase domains repeatedly.[9] How each domain can differentially recognize products of different chain elongation steps has remained a mystery. Some fungal iterative PKSs also require auxiliary proteins. For example, several known decalin-containing polyketides require the cooperation of the PKS and an auxiliary enoyl reductase (ER).[10–13] Upon completion of synthesis of the acetate-derived chain on the PKS, the enzyme bound intermediate is transferred to the NRPS module, at which point an amino acid is appended to the compound.[1,2]
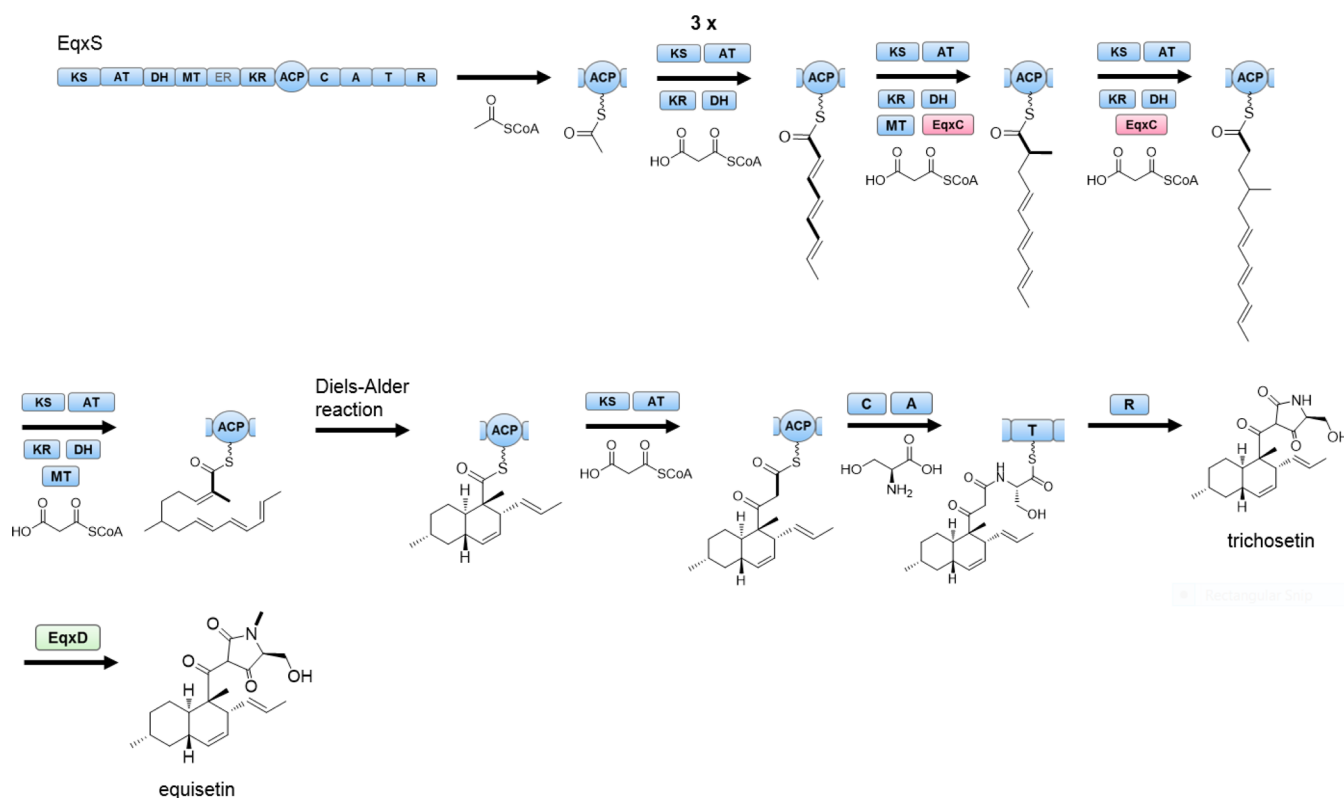
Figure 1. Biogenesis of equisetin showing the complex programming of the PKS where domains on a single polypeptide are utilized iteratively to synthesize the polyketide core. An acyl transferase (AT) domain selects the malonyl units, the ketosynthase (KS) catalyzes the decarboxylative condensation, the C-methyltransferase (MT) performs α-C-methylations, the ketoreductase (KR) reduces keto groups to hydroxyls and the dehydratase domain (DH) eliminates the hydroxyl groups to form olefins, which are then reduced to single bonds by the trans ER EqxC. The synthesized polyketide core is then transferred to the NRPS module, composed of the adenylation domain, which selects the amino acid that gets loaded onto the thiolation domain (T). Conjugation of the polyketide with amino acid is catalyzed by the condensation (C) domain. The R domain catalyzes a Dieckmann reaction to release the product as a tetramic acid, trichosetin.[25,29,30] Further N-methylation by EqxD forms equisetin.[11] Whether the Diels−Alder reaction occurs on-enzyme or after product release is still unknown (it has been drawn on-enzyme for clarity).

The NRPS is responsible for selecting the amino acid, synthesizing the peptide bond, and sometimes performing other steps, such as Dieckmann cyclization to produce tetramic acids. The complexity of the PKS-NRPS biosynthetic process is exemplified in Figure 1, showing the biogenesis of equisetin.

Many different types of fungal polyketide-peptide hybrids have been isolated. In the case of lovastatin and relatives, only a fragment of the NRPS is present, and amino acids are not loaded.[10] The remaining NRPS domains are thought to be responsible for potentially catalyzing a Diels−Alder reaction that cyclizes the linear polyketide chain to produce decalins.[6,10] Other products include the cyclopiazonic acid precursor, cAATrp, which contains the shortest possible polyketide chain fused as a tetramic acid to tryptophan (Figure 2).[14] Prepseurotin is derived from a longer, partially reduced polyketide fused to phenylalanine.[15,16] Equisetin has a decalin polyketide structure, like lovastatin, but is fused to serine as a tetramic acid.[11,17] Fusaridione A is a long, linear polyene fused to tyrosine.[11] Tenellin and desmethylbassianin are ring-expanded tetramic acid derivatives; their biosynthetic genes are among the best characterized fungal PKS-NRPSs.[7,18−20] Many other classes of PKS-NRPS products are known.[21]

A key to combinatorializing fungal PKS-NRPS proteins would thus be to hybridize PKS modules, producing diverse acetate-derived scaffolds, with NRPS modules, activating diverse amino acids. Indeed, this has been tried on three occasions with mixed success.[6−8] Two major domains are of
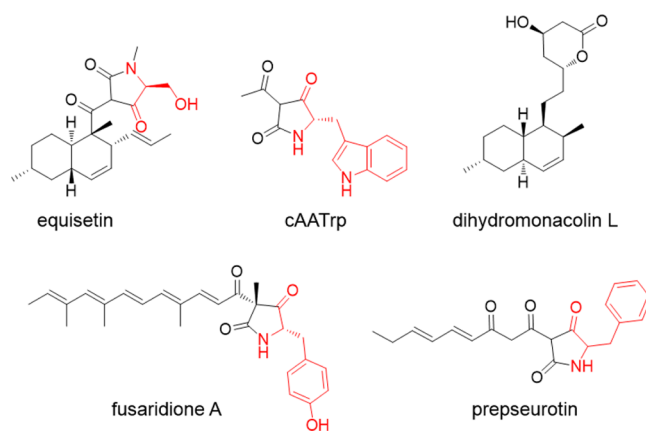

Figure 2. Structures of PKS-NRPS products showing the polyketide chain synthesized by the PKS module (black), and the amino acid appended by the NRPS (red). Dihydromonacolin L is not an amidated product because its biosynthetic enzyme, LovB, possesses a truncated NRPS module.

primary importance in successful fusions: the acyl carrier protein (ACP) from the PKS module, and the peptide bond-forming condensation (C) domain from the NRPS module. It was anticipated that the C domain might exhibit some substrate selectivity, where only certain polyketide products might be acceptable in amide bond formation. The lynchpin was

considered to be the ACP. This small (~70−100 amino acids) but critical domain is covalently tethered to all PKS intermediates and serves to ferry them between different catalytic sites in the PKS.[9] The ACP protein itself must form productive protein−protein interactions with the five or more protein domains in the PKS. The ACP also must form a protein−protein interaction with the C domain, which accepts the incoming polyketide substrate for elongation. Any protein fusions between NRPS and PKS modules therefore pose both a substrate selectivity problem (C domain) and a protein interaction problem (ACP domain).

The major goal of this project was to determine basic programming rules that would enable successful fusion of fungal PKS-NRPS proteins. The study was designed to disentangle the protein−protein interaction and substrate selectivity questions in this highly complex system. To do so, we proposed the hypothesis that fungal ACP domains contain specific sequence elements enabling interaction with PKS and NRPS modules. A series of gene fusions were made using different PKS and NRPS proteins, exploring several types of ACP interactions, with the goal of maintaining normal PKS function and forging productive ACP/C domain interactions. The resulting genes were expressed, and the chemical products were analyzed. In the event, we delineated protein−protein interaction rules and showed that fungal C domains are highly substrate-selective. The secondary goal was to understand key reactions catalyzed by PKS-NRPS proteins, such as the Diels−Alder reaction that leads to decalin products. With the exception of the Diels−Alder reaction, which in the case of lovastatin is clearly catalyzed by the native C domain, we found that fungal PKSs operate independently of NRPS modules and make the products that are expected from the natural PKS-NRPS proteins. Taken together, these experiments provide an integrated view of the function and engineering potential of fungal PKS-NRPS hybrids.

■ RESULTS

Experimental Design and Mutant Construction. To differentiate the effects of protein−protein interactions and C-domain selectivity, we fused PKS and NRPS proteins at different points, using four different types of ACP connections (Figure 3). First, we left the PKS ACP intact, fusing it directly with the second C domain. Second, we replaced the ACP with the ACP natively fused to the acceptor NRPS. Third, we fused the N-terminal region of the ACP from the PKS with the C-
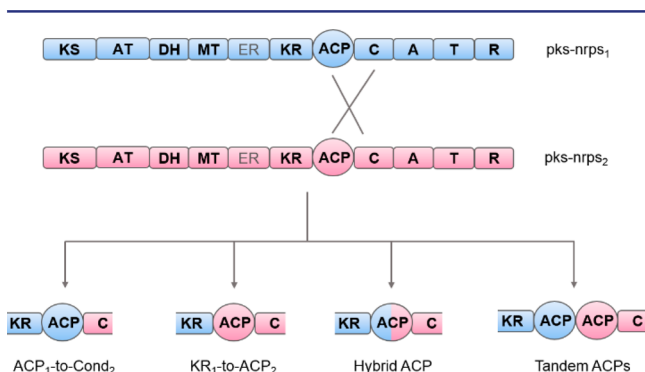
terminal region of the ACP natively fused to the acceptor NRPS. Previous studies with PKS and NRPS proteins have shown that different regions of the ACP are responsible for interacting with donor and acceptor modules.[22,23] While this has not been studied in fungi, we reasoned that a hybrid ACP might overcome potential problems with donor or acceptor recognition. At the least, it would provide crucial information about fungal ACP function. Finally, since multiple ACPs are functional in some bacterial pathways,[24] we placed the ACPs from the PKS and NRPS in tandem.

Fusions were performed by yeast recombination to generate expression vectors. The fused genes were then expressed in a model platform recently developed in *Fusarium heterosporum*.[25] The strength of this platform is that it synthesizes heterologous polyketides and polyketide-peptide hybrids in yields of ~100−1000 mg kg^{-1}. We expected that some fusions might lead to decreased yields of products, and therefore, this high yielding starting point would enable even relatively poorly functioning hybrids to be accurately analyzed. This platform uses the native promoters and regulatory elements that normally produce equisetin, but instead redirects them to the production of heterologous products. We used the recently characterized equisetin synthetase EqxS and fusaridione synthetase FsdS proteins as the NRPS modules in all experiments.[11,25] These were then fused to PKS modules from *eqxS* and *fsdS*, as well as from cyclopiazonic acid synthesis (*cpaS*),[26] pseurotin synthesis (*psoA*),[15] and lovastastin synthesis (*lovB*).[12] Where auxiliary ER proteins were required by the PKS, these were coexpressed in *F. heterosporum*. A total of eight PKS-NRPS fusions were created, each with four different ACP connections, for a total of 32 recombinant clones expressed in *F. heterosporum*. Of particular importance, the function of the nonhybridized (wild-type) PKS-NRPS proteins was first assessed in *F. heterosporum*, showing that all of the proteins were functional prior to hybridization.[11,25]

The equisetin and fusaridione synthetases *eqxS/eqxC* and *fsdS* were cloned from *F. heterosporum*, the same strain used as the heterologous production line in this study. *cpaS*, *psoA*, and *lovB* and *lovC* (a trans-ER) were cloned from *Aspergillus flavus*, *A. fumigatus*, and *A. terreus* respectively. Three of the fusion types (comprising 12 total recombinants) failed to function, including *cpaS-eqxS*, *cpaS-fsdS*, and *lovB-fsdS*. An absence of products in these experiments could occur for several different reasons. For example, the polyketide product of *cpaS* is simply acetoacetate, which may not be detectable in the experimental conditions, or protein folding problems may occur. Because of these and other reasons, it is not possible to definitively conclude that absence of product is meaningful in terms of protein interactions or substrate acceptability. Therefore, these failed fusions will not be further analyzed. For the remaining five fusion types (20 total recombinants), including *eqxS-fsdS*, *fsdS-eqxS*, *psoA-fsdS*, *psoA-eqxS*, and *lovB-eqxS*, products were detected (Figure 4).

In addition to these 32 mutants, we also constructed other hybrids to answer specific questions resulting from these initial studies, as described below.

ACP-PKS Interactions. All PKSs were active in making polyketides when fused to their own ACP domains, but only a subset of fusions between a PKS and a noncognate ACP led to polyketide products (Figure 4). All four ACP combinations effectively led to the native PKS products in the reciprocal crosses, *eqxS-fsdS* and *fsdS-eqxS*. This indicated that these ACPs recognized all domains of each other's PKS module. Modest
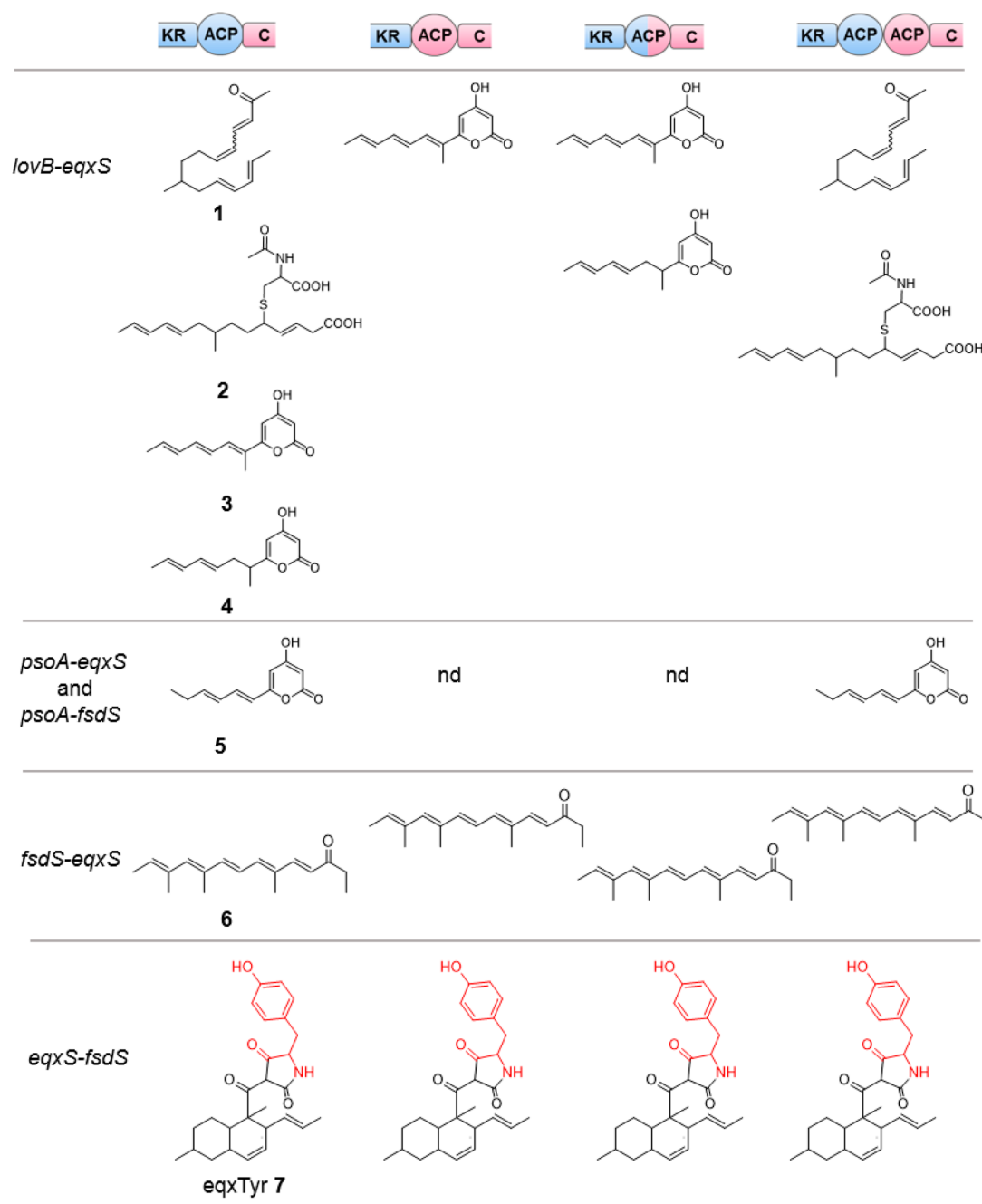

Figure 3. Design of gene fusions representing the different arrangements of the ACPs with the C domain. pksnrps1 (blue) represents *psoA*, *cpaS*, *lovB*, *eqxS* or *fsdS*; and pksnrps2 (red) represents *eqxS* or *fsdS*.

Figure 4. Summary of compounds identified from the different PKS/NRPS fusions. **1, 2, 5, 6** and **7** are new compounds (nd = no new metabolites detected in comparison to *gfp*-expressing control). Note: Compound **1**, all double bonds are trans.

and potentially insignificant yield differences were the only observed effects of swapping ACPs between these systems.

By contrast, in the case of pseurotin fusions *psoA-eqxS* and *psoA-fsdS*, only the cognate ACPs were functional, and neither the fusaridione ACP nor the equisetin ACP could substitute for the pseurotin ACP (Supporting Information Figure S1). Fusions containing either the pseurotin ACP alone, or the tandem ACP system containing one ACP from pseurotin and one from fusaridione or equisetin, both led to PKS product **5** (Figure 4). By contrast, no products were detected in the hybrid ACP or noncognate ACP fusions. Compound **5** contains the complete PKS chain in the correct reduction state as found in prepseurotin. It was likely spontaneously cleaved from the

enzyme as the pyrone, as has been proposed for similar products from fungal hr-PKS.[12]

lovB-eqxS fusions also led to PKS products, which differed depending upon which ACP combination was used. Previously, when *lovB* was expressed intact in *F. heterosporum*,[25] in addition to a large amount of the expected product, dihydromonacolin L (Figure 2), we obtained a small amount of pyrone **4** (Figure 4). The cognate *lovB* ACP led to a PKS product **1** that was most similar to the native monacolin product, but lacking Diels–Alder cyclization (Figure 4; Supporting Information Figure S2A). In addition, two pyrone shunt products **3** and **4** were identified, as found with the wild-type enzyme when expressed in *F. heterosporum*. Only these shunt products were observed when noncognate ACPs were used, including from the hybrid

ACP and equisetin-derived ACP fusions. Most interestingly, when *eqxS* provided the ACP, the product was not reduced by the trans-ER *lovC*. The best explanation for this observation is that the equisetin ACP cannot interact with LovC. An exhaustive search for new nitrogen-containing metabolites led to the isolation of a minor product **2**, a linear lovastatin-like polyketide precursor modified by *N*-acetyl cysteine (biochemically analogous to the very similar leukotriene modification[27]). This same product was produced when the ACPs were placed in tandem.

Transfer of PKS Products to the NRPS. Although 16 out of the 32 gene fusions led to functional PKS proteins with detectable products, only 4 of these led to products that were clearly passed along to the NRPS module. These were the *eqxS-fsdS* fusions, in which the expected product was clearly obtained (Figure 4; Supporting Information Figure S3). The sole recombinant natural product, eqxTyr **7**, isolated from this system included the decalin made by the equisetin PKS and the tyrosine tetramic acid made by the fusaridione NRPS. All four *eqxS-fsdS* fusions led to the same product. This indicated that the *fsdS* C domain could accept the equisetin polyketide product and could make functional contacts with the equisetin ACP.

In the reciprocal fusion, *fsdS-eqxS*, the native fusaridione polyketide was isolated, but it was not fused to serine from *eqxS* and was instead decarboxylated. Interestingly, all four ACP fusions led to the same product **6** (Figure 4; Supporting Information Figure S4). This indicated that the *eqxS* C domain could not accept polyketide products synthesized by *fsdS*, since the native *eqxS* ACP was present in some fusions, precluding a role for protein−protein interactions in governing C-domain selectivity. Both the *fsdS* and *eqxS* ACPs could act in concert with the PKS domains, leading to synthesis of the normal *fsdS* PKS product. Decarboxylation of β-ketoacids is a spontaneous reaction, so that **6** represents exactly what one would expect from a functional PKS.

Similarly, pseurotin and lovastatin fusions led only to PKS products, with no evident transfer to the *fsdS* or *eqxS* NRPS modules. In the case of the lovastatin fusions, products were still obtained with the construct containing the *eqxS* ACP alone, implicating substrate selectivity of the C domain as the major factor limiting transfer.

Diels−Alder Reaction in Lovastatin Biosynthesis. As noted above, the *lovB-eqxS* fusion led to a polyketide product **1** resembling a pre-Diels−Alder product. Natively, *lovB* includes not just the PKS module, but also an intact C domain and a short piece of the adenylation (A) domain that activates amino acids.[10,12] Previously, it has been speculated that the *lovB* C domain was responsible for Diels−Alderase activity.[6,10,28] This previous idea is strongly supported by the chemistry discovered here, the first time in which the pre-Diels−Alder product has been observed as a *lovB* PKS product. To further probe this issue, we fused *lovB* with *fsdS* at the position in which the *lovB* A domain is fragmented. This construct was expressed in tandem with *lovC*, and it yielded the lovastatin precursor, dihydromonacolin L (Supporting Information Figure S5). Thus, fusions with an intact *lovB* C domain yield the Diels−Alder product, while fusions to heterologous C domains do not and in some cases lead to pre-Diels−Alder products. This provides strong evidence supporting the C domain as the region responsible for the Diels−Alder reaction.

Of note, in the case of equisetin and other tetramic acids, the tetramic acid motif remains adjacent to the decalin ring system,

in the perfect position to accelerate a Diels−Alder reaction. By contrast, in lovastatin the final PKS intermediate is chain extended so that a methylene group is adjacent to the decalin ring. Thus, chemically, it is highly likely that the decalin ring in lovastatin must be formed at a chain length in which the enzyme-bound thioester is immediately adjacent to the nascent decalin; it is formed while still bound to the PKS. By contrast, in principle the equisetin decalin ring could still be formed after cleavage from the NRPS. Indeed, in a previous study, we found evidence that the decalin rings in the tetramic acid pyrrolocin may be formed post-PKS.[25] Based upon these results and the ideas of Hutchinson, Vederas, Tang, and others,[6,10,12,28] we speculate that the substrate selectivity and chemical recognition inherent to C domains has been redirected to template the Diels−Alder reaction in lovastatin synthesis.

C Domain Selectivity and KR Domain Activity. In previous studies, the ketoreductase (KR) domain was shown to control chain length in some fungal PKS-NRPS proteins.[7] This provided us with an opportunity to more finely investigate the role of C domain in substrate selectivity, removing protein−protein recognition from consideration and focusing solely on the chain length presented to C domain. We swapped the *eqxS* KR domain for the *lovB* KR domain, in the context of the full-length *eqxS* PKS-NRPS (Figure 5). Since LovB and EqxS natively synthesize nonaketides and octaketides respectively, it was envisioned that slightly different products would be presented to C domain. When this chimera was coexpressed with the trans-ER *eqxC*, several new metabolites were detected. The major product was characterized to be the polyketide **8**
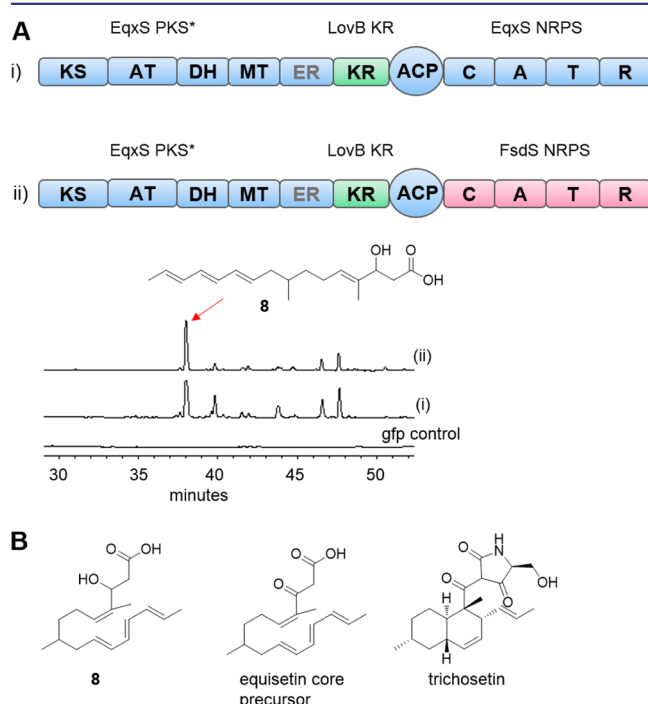

Figure 5. Analysis of chimeric *eqxS* PKS with *lovB* KR. (A) The *eqxS* KR was swapped for the LovB KR domain (green) in *eqxS* (i), and in *eqxS-fsdS* (ACP1-to-Cond2; (ii)). HPLC-DAD analysis of crude extracts of CGA cultures of FusΔeqx5 cotransformed with eqxC and either (i) or (ii) shows the production of **8**. Trichosetin and **7** are synthesized in minor amounts by (i) and (ii), respectively. (B) Comparison of **8** with the equisetin polyketide core and trichosetin showing the position of the β-hydroxyl group.
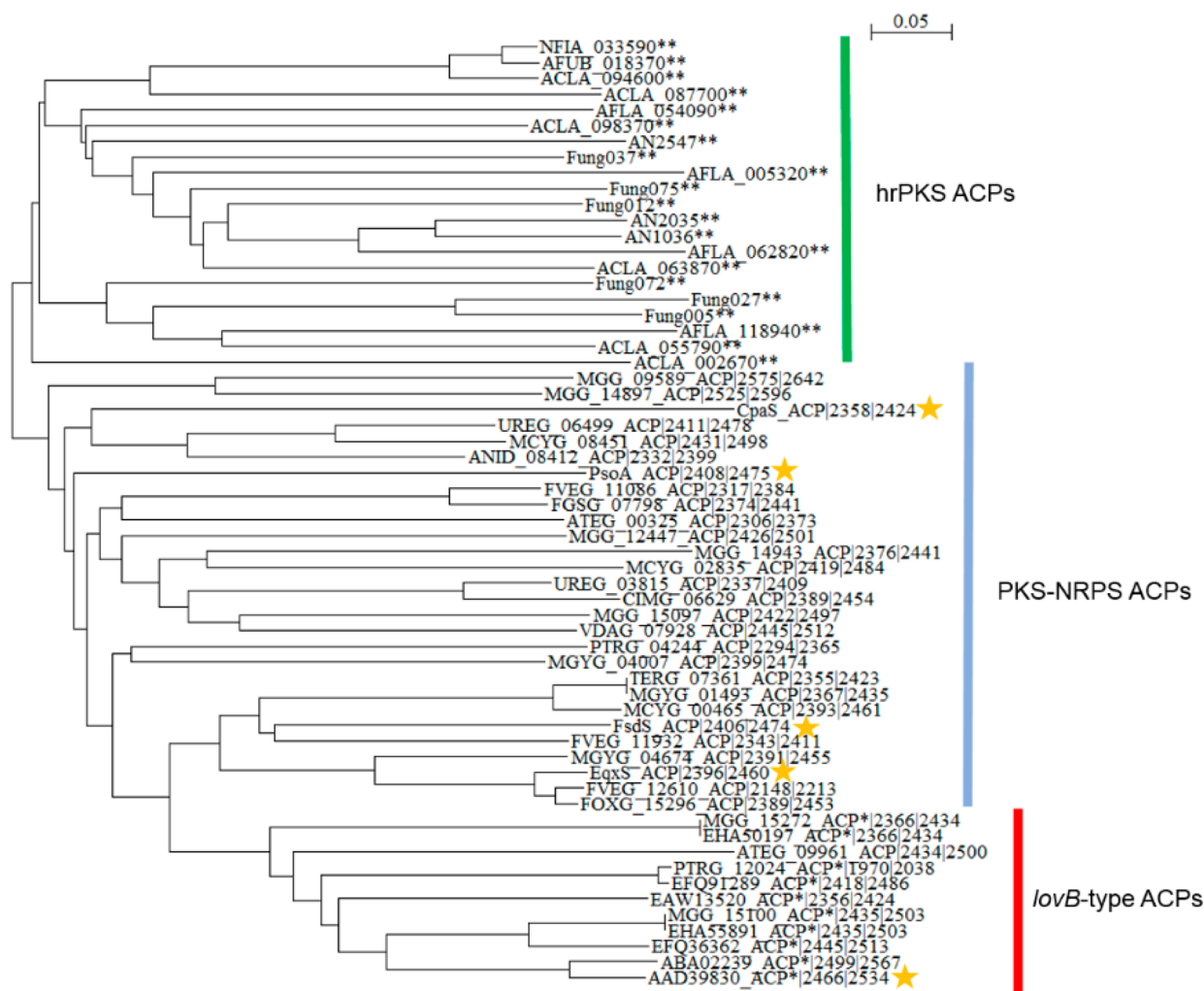
Figure 6. A phylogenetic analysis of ACPs from hrPKSs, PKS-NRPSs, and *lovB*-type PKSs (with truncated NRPS). Starred ACPs were investigated in this study. The *eqxS* ACP and *fsdS* ACP are more closely related to each other than they are to *psoA*, *cpaS*, and *lovB* ACPs.

which is not conjugated to serine (Figure 5A). To our surprise, **8** was the same length as the equisetin pre-Diels–Alder polyketide chain, but with a β-hydroxyl group in place of the keto moiety (Figure 5B). A very minor amount of trichosetin, normally the major product of *eqxS* + *eqxC*, was observed. In contrast to expectation, *eqxC* cooperated with the chimeric protein to perform the two normal reductions observed in equisetin synthesis.

Similarly, a fusion was made in which the *lovB* KR was swapped into the *eqxS-fsdS* chimera. The same major pre-Diels–Alder compound **8** was again observed, with a very minor amount of **7**, the tyrosine analog of equisetin (Figure 5). Both KR swaps provided the same result: simple reduction of the β-keto group absolutely abolished transfer to the NRPS module. Only a small amount of product, escaping the kinetic reduction of the β-keto group, was captured by the equisetin and fusaridione NRPS domain and added to an amino acid. These results serve to strongly reinforce the role of C domain in selecting the precise substrate for chain elongation.

Potential Impact of R Domain on Selectivity. The R domains at the C-terminus of *eqxS* and *fsdS* belong to a subset of proteins that catalyze Dieckmann cyclization to afford tetramic acids; they are not competent reductases and do not bind to NAD(P)H.[29,30] Previously, biochemical experiments with FsdS R and ATR domains showed that, in the case of

FsdS, R has broad substrate selectivity that would not be expected to impact the experiments described here.[30] For example, acetoacetyl-alanine was readily accepted by the domains in vitro. To further probe this issue, we examined the products of *lovB-eqxS* and *cpaS-eqxS* fusions, in which the terminal R domain was covalently linked to green fluorescent protein (sGFP). In previous work with full length *eqxS*, we showed that introduction of sGFP in this position blocked activity of R, so that hydrolytic products were obtained instead of Dieckmann cyclase products.[25] The fusion products of *lovB-eqxS-sGFP* were identical to those from *lovB-eqxS*, and as in the previous experiments no products were detected from *cpaS-eqxS-sGFP*. This experiment provides further evidence that the primary chemical selectivity for the *eqxS* NRPS is likely to be at the C domain.

■ DISCUSSION

Here, we answer some of the key unsolved questions about the function and engineering of fungal PKS-NRPS hybrids. In total, we synthesized and tested 57 fusion products, including 32 PKS-NRPS module swaps, 2 KR swaps, 1 *lovB* C-A fusion, and 22 *GFP* fusions to the PKS-NRPS hybrids. By analyzing the chemical products resulting from these fusions, we tested the hypothesis of ACP domain specificity, revealing specificity

elements that will be essential in engineering. We showed that the *eqxS* and *fsdS* C domains are highly substrate selective. Our secondary goal was to better understand fungal hr-PKS function. We provide new insights about the role of specific domains in determination of Diels–Alderase activity and chain-length determination. More fundamentally, we show that the PKS module alone is sufficient to provide the expected polyketide structure, and that C domain does not contribute beyond the Diels–Alder reaction. We demonstrate for the first time a new product resulting from a functional in cis fusion between complete PKS and NRPS modules. The first pre-Diels–Alder lovastatin enzymatic products were discovered, which in tandem with other evidence provide strong support for the role of C domain in lovastatin Diels–Alder reaction.

The biochemical basis of the Diels–Alder reaction has been the source of great interest for the past 15 years, but it has not yet been completely resolved. Elegant work by Vederas and co-workers showed that LovB catalyzes the on-enzyme, stereospecific Diels–Alder reaction to form the decalin ring found in the lovastatin precursor, dihydromonacolin L.[28] Later studies reported the inability of a truncated LovB, lacking the C domain, to form monacolins when coexpressed with LovC, but instead truncated pyrones are formed.[10] The LovB C domain was however able to restore monacolin production when added in trans to the truncated LovB in vitro. Another study demonstrated that LovB fused post-C domain with the chaetoglobosin A synthetase (CheA) was capable of synthesizing dihydromonacolin L, whereas the direct fusion of the lovB PKS with CheA C domain was nonproductive.[6]

In some ways, our results here are quite similar to the latter two studies. We found that, in the absence of the native C domain, LovB synthesizes polyunsaturated pyrones. When the intact LovB C and a small part of LovB A are added back, rather than tetramate formation, dihydromonacolin L is found. The major difference is that, for the first time, we have also identified a pre-Diels–Alder product **1** from fermentations with LovB lacking its native C domain. The isolation of **1** provides direct evidence showing concretely that all of the domains in LovB and LovC are functional, and that they synthesize an intermediate that would be cyclized if a functional Diels–Alderase were present. When that Diels–Alderase (C domain) is added back, the Diels–Alder reaction takes place. This evidence strongly favors the C domain hypothesis forwarded previously by other researchers. We were able to isolate and characterize this product because we designed our study to use an expression platform that natively produces a high yield of recombinant products.[25]

We thoroughly define the role of the ACP in governing PKS function. In pairing PKS domains with noncognate ACPs, we were first faced with the question of whether PKS domains, including the trans ER, would interact with the noncognate ACP to form the native PKS polyketide intermediate to present to the C domain. This question has been explored in several other types of PKS proteins,[22,31,32] but never before within the fungal hr-PKSs. Within the fungal hr-PKS and PKS-NRPS proteins, the role of ACP has been previously indirectly tested in a series of domain fusions within the closely related bassianin and tenellin PKSs.[7] In these fusions, the tenellin ACP was always in place, and it could interact with several other domains from the bassianin PKS. Our work explicitly and thoroughly explores the role of ACP in PKS interactions. The *eqxS* ACP and *fsdS* ACP are equivalent and interchangeable. The *lovB-eqxS* fusions reveal that the *eqxS* ACP can interact with the

LovB PKS, but does not interact with LovC. Interactions with LovC are restored by reintroducing a portion of the *lovB* ACP in the hybrid ACP fusion. By contrast, the *eqxS* and *fsdS* ACPs are incapable of interacting with the *psoA* PKS domains.

A phylogenetic analysis of the ACPs (Figure 6) illuminates the mixed results from these ACP-PKS interaction studies. Strikingly, more closely related ACPs are interchangeable, while more distantly related ACPs are not. The intermediate case, involving the *lovB* swaps, provides an intermediate level of transferability. Therefore, for swaps involving the use of noncognate ACPs, source modules with closely related ACPs may result in production of predicted PKS intermediates.

Transfer of polyketide intermediates to NRPS modules has also been investigated in several cases. Within the fungal enzymes, only three other studies have investigated the compatibility of PKSs with noncognate NRPS modules. (1) The aspyridone synthetase (ApdA) PKS module and the CpaS NRPS module were expressed as separate proteins in *Saccharomyces cerevisiae*.[8] This in trans experiment led to successful production of a small amount of the predicted tetraketide-derived tetramate, whereas natively a diketide intermediate is accepted by the CpaS NRPS module. The ApdA PKS was shown to have the potential to synthesize longer polyketide intermediates, but only the tetraketide was captured by both the ApdA and CpaS C domains. (2) Highly similar proteins desmethylbassianin synthetase DmbS and tenellin synthetase TenS (>85% protein sequence identity) were fused, leading to a successful synthesis of similar products.[7,20] These proteins synthesized hexaketide and pentaketide tetramates with slightly different methylation patterns. (3) The *lovB* PKS was fused to the CheA NRPS, but no products were detected.[6] The incompatibility of LovB PKS with CheA NRPS was attributed to an early evolutionary divergence between the LovB-type PKSs and other PKS-NRPSs as observed from phylogenetic analyses.

A complicating factor in the above studies was that they did not differentiate between the impact of protein–protein and substrate selectivity interactions between modules, and they used only single swaps rather than thoroughly investigating multiple combinations. Here, we tested the effect of C domain selectivity independent of potential unfavorable ACP interactions for a larger set of PKS/NRPS pairings that differ significantly in sequence similarity. In the context of previous work by others, our thorough data here demonstrates that C domains of PKS-NRPS enzymes are often highly chemoselective. The *lovB-eqxS* fusions synthesize several nonamidated polyketide chains. Even in the presence of the *eqxS* ACP which forms interactions with *eqxS* C domain, the polyketides are not transferred to the NRPS. These cases imply an underlying strict *eqxS* C domain substrate selectivity beyond the presence of productive ACP interactions. This is further supported by the inability of similar *fsdS-eqxS* fusions to synthesize tetramate products, but instead form the polyketide **6**. In fact, the experiment with the *lovB* KR domain swap suggests that the *eqxS* C domain possesses a high selectivity for only the equisetin polyketide core because even when it is the disfavored/aberrant polyketide intermediate among others, trichosetin is the only amidated product detected. On the other hand, the *fsdS* C domain displays more relaxed substrate selectivity, accepting the equisetin polyketide core in addition to the more rigid polyene polyketide chain found in fusaridione. In these *eqxS-fsdS* expressions, many other nontetramate products are detected; it is therefore intriguing that eqxTyr **7**

is the only amidated product observed. One explanation for this is that the *fsdS* C domain is not truly accepting of broad substrates, but rather selects for particular attributes of the substrate which may include length, degree of unsaturation and/or methylation pattern. This becomes more apparent when the octaketide intermediate of equisetin is compared to the heptaketide polyketide product of *fsdS* PKS. It is possible that *fsdS* C accepts the linear pre-Diels–Alderase intermediate. The pseudorelaxed selectivity of the *fsdS* C domain for the equisetin core is further demonstrated in the *lovB* KR swap experiment, where trace amounts of eqxTyr were detected, when the equisetin polyketide core is produced as an aberrant product among others. Tenellin synthetase (TenS) produces several tetramates in the absence of its trans ER.[20] Upon the basis of this result, C domains have been speculated to be broad substrate, but our data shows that this is not universal. Even for TenS, the polyketide chains did not differ significantly among the tetramates synthesized; the lengths of these chains are within one ketide unit of the native substrate of the C domain. This difference is exactly what we found for the successful chain transfer in *eqxS-fsdS* fusions.

Our studies also enabled other interesting observations into PKS domain function when the *eqxS* KR was swapped for the *lovB* KR. The synthesis of the same equisetin core polyketide chain albeit with a β-hydroxyl group instead of the anticipated nonaketide chain, in contrast with previous work,[7] shows that iterative PKS programming is much more complex and each system should be evaluated individually until more universal rules can be discovered. The chain length factor for *eqxS* probably lies with the KS domain. This result also shows that, for the case of *eqxS* biosynthesis, capture of the polyketide intermediate by the C domain occurs after polyketide elongation is complete. KR and C domain compete with each other for the last intermediate in chain elongation, with the KR reaction dominating. A caveat in interpreting this result is that the β-hydroxyamide would not be a substrate for Dieckmann cyclization, so that later hydrolysis of the amide bond would provide the detected product even though successful chain transfer had in fact occurred. Arguing against this possibility is the fact that we have previously isolated non-Dieckmann products from experiments with both FsdS and EqxS, and they were perfectly metabolically stable.[11,25]

The results of this study, adding to previous work, will be useful in the engineering of fungal polyketides and polyketide-peptide hybrids. First, the work shows that the expected polyketide products themselves can be directly formed via expression of these fusions; merely swapping out C domains can provide valuable products. Second, it shows that successful engineering strategies will require appropriate selection of ACPs. Currently, it is clear from this work and previous studies that phylogenetically closely related ACPs are empirically likely to function in a heterologous context. However, more research is required to delineate exactly the recognition factors between ACPs and these complex enzymes that may facilitate rational swaps between PKSs. Finally, the substrate selectivity of C domains is a serious hurdle in the generation of recombinant products. The best way forward at this point will involve using C domains that accept somewhat similar PKS products, but that append different amino acids to the compound. The value of this approach is clearly shown by the very conservative previous successes in synthesizing bassianin and cyclopiazonic acid derivatives and in our case of switching the amino acid on equisetin from serine to tyrosine. Fortunately, fungal PKS-

NRPS genes are ubiquitous,[33] so that an enormous biodiversity is available for such engineering efforts.

■ METHODS

Cloning of Fungal Expression Plasmids. Generally, genes were amplified from genomic DNA or subcloned coding sequences using the high-fidelity Phusion polymerase (NEB), and cloned into the expression vectors by yeast recombination.[34,35] Plasmid selection and amplification was then done in *Escherichia coli*. Details of plasmid construction are contained in Supporting Information. Alignments and phylogenetic analysis were generated with ClustalX.

Fungal Transformation. *Fusarium heterosporum* FusΔeqx5 and FusΔeqx5pyrG10 were used as the expression hosts.[25] The plasmids were linearized with PacI/AscI prior to transformation into FusΔeqx5 as previously described.[11,25] Selection was done with hygromycin 150 μg mL^{-1}, and isolated transformants cultivated in potato dextrose broth (PDB) 250 mL for 7 d at 30 °C and corn grit agar (CGA) 50 g for 21 d at rt. *psoA-eqxS* and *psoA-fsdS* mutants were cultured in PDB supplemented with sodium propionate (20 mM) at the 24-h point of the culture because propionyl CoA is the first unit loaded by PsoA.[15] Absent addition of propionate, no new compounds were observed in the *psoA* expressions.

General Procedures for Extraction, Purification, and Characterization of Compounds. PDB cultures were extracted with an equal volume of ethyl acetate, and solvent removed under a vacuum. The crude extracts were analyzed by HPLC using a Hitachi LaChrom Elite instrument equipped with a diode array detector over a Phenomenex Luna C18 column (4.6 × 250 mm, 5 μm). Preparative HPLC was done using the Discovery HS C18 column (25 cm × 10 mm, 5 μm). Extracts were further characterized by LC/MS using the Agilent ZQ to screen for new metabolites. Cultures on CGA were extracted with acetone and the crude extracts treated similarly. Compounds were generally purified by first performing flash chromatography on end-plugged C18 using a methanol/water gradient. The fractions containing the target compounds were then purified by preparative HPLC using an acetonitrile/water solvent system (with or without 0.05% TFA). The collected fractions from several rounds of HPLC were pooled and the solvent removed under a vacuum. HPLC-grade solvents were used as purchased. HRESIMS data for the purified compounds was obtained from a Waters Micromass Q-TOF Micro mass spectrometer. All NMR data was acquired on a Varian INOVA 500, except the ^{1}H, HSQC and HMBC–^{15}N data for **2** which was acquired on a Varian INOVA 600 fitted with a cryoprobe.

Purification of 1. Compound **1** was purified by HPLC using 80% acetonitrile/water mobile phase containing 0.05% TFA. Solvent was removed under a vacuum to afford **1** (1.3 mg); HRESIMS *m/z* 233.1883 [M + H]$^+$ (Calcd for C$_{16}$H$_{25}$O 233.1900; Δ −7.3 ppm); 1D and 2D NMR data (Supporting Information Table S5; Figure S6).

Purification of 2. Compound **2** was purified by HPLC using 40% ACN/20% water mobile phase containing 0.15% TFA. Removal of solvent under a vacuum yielded **2** (0.4 mg); HRESIMS *m/z* 398.1994 [M + H]$^+$ (Calcd for C$_{20}$H$_{32}$NO$_5$S 398.1996; Δ −0.5 ppm); 1D and 2D NMR data (Supporting Information Table S2; Figure S7).

Purification of 6. The crude extract from a PDB culture (500 mL) was subjected to flash chromatography over end-plugged C18. The fraction containing **6** was dried under a vacuum and the residue further purified by HPLC using an 85% ACN/15% water mobile phase. Solvent was removed under a vacuum to afford **6** (1.1 mg); λ_{max} = 373 nm; HRESIMS *m/z* 245.1875 [M + H]$^+$ (Calcd for C$_{17}$H$_{25}$O 245.1900; Δ −10.2 ppm); 1D and 2D NMR data (Supporting Information Table S4; Figure S8).

Purification of eqxTyr 7. A portion of the semipurified fraction from flash chromatography over end-plugged C18 was further purified by HPLC using a 75% ACN/25% water mobile phase containing 0.05% TFA. The collected fractions were pooled and solvent was removed by vacuum to afford an off-white powder (0.9 mg); HRESIMS *m/z* 436.2492 [M + H]$^+$ (Calcd for C$_{27}$H$_{34}$NO$_4$

436.2482; Δ 2.3 ppm); 1D and 2D NMR data (Supporting Information Table S3; Figure S9).

Purification of 8. Compound **8** was purified by HPLC using a 67% ACN/33% water mobile phase containing 0.05% TFA. Solvent was removed to afford **8** (1.8 mg); HRESIMS *m/z* 291.1963 [M-H]⁻ (Calcd for $C_{18}H_{27}O_3$ 291.1966; Δ 1.0 ppm); 1D and 2D NMR data (Supporting Information Table S7; Figure S10).

Characterization of 3, 4, and Dihydromonacolin L. These were characterized by comparing UV absorbance data, *m/z* values, and MS fragmentation to previously reported data.[10,25]

Characterization of 5. Compound **5** was purified by preparative HPLC with a gradient running from 30% to 70% acetonitrile/water in 20 min. Pooled fractions from several rounds of HPLC were dried under a vacuum to afford **5** (1.1 mg); HRESIMS *m/z* 193.0863 [M + H]⁺ (Calcd for $C_{11}H_{13}O_3$ 193.0859; Δ 2.1 ppm). ¹H NMR spectra displayed broad signals in both CDCl₃ and CD₃OD, and 2D NMR showed weak signals. On the basis of this NMR data, we predicted **5** to be an unsaturated pyrone.[12] After purification, **5** was kept in the dark, under argon. ¹H and COSY NMR data indicated an unsaturated side chain as shown in the structure diagram, but like other compounds in this class[12] the compound was not sufficiently well behaved for NMR characterization. Therefore, characterization of **5** was done by comparative gas chromatography-electron impact-MS (GC-EI-MS) with a sample of **4**. To dried samples, methoxyamine hydrochloride (40 µL, 40 mg mL⁻¹ in pyridine) was added and the mixture heated at 40 °C for 1 h. After cooling, MSTFA (40 µL) was added and the reaction mixture was left at rt for 12 h. GC/MS analyses were conducted using an HP6890 instrument interfaced with an MSD-HP5973 detector and equipped with a Zebron ZB-5MSi Guardian (30 m × 0.25 mm ID, 0.25 µm film thickness; Phenomenex) column and an HP7682 injector. Helium was used as a carrier gas with a 100:1 split ratio at an injection volume of 1 µL. The injector temperature was set to 250 °C. The oven temperature gradient was programmed as follows: 95 °C held for 1.5 min increased at a rate of 40 °C/min to 118 °C, held for 1 min, increased at a rate of 5 °C/min to 250 °C, increased at a rate of 25 °C/min to 330 °C and held for 12.3 min. MS spectra were obtained in EI mode within a range of *m/z* 50−500. Other parameters: MS quad and source temperatures were set to 150 and 230 °C, respectively; solvent cut time was 4 min; and scanning was done at 16 scans/sec. MS fragmentation clearly supported the proposed structure **5**; see Supporting Information Table S8.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

Supporting tables and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

ews1@utah.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Boettger, D.; Hertweck, C. *ChemBioChem* **2013**, *14*, 28.

(2) Fisch, K. M. *RSC Adv.* **2013**, *3*, 18228.

(3) Carter, S. B. *Nature* **1967**, *213*, 261.

(4) Graupner, P. R.; Carr, A.; Clancy, E.; Gilbert, J.; Bailey, K. L.; Derby, J.-A.; Gerwick, B. C. *J. Nat. Prod.* **2003**, *66*, 1558.

(5) Song, Z.; Cox, R. J.; Lazarus, C. M.; Simpson, T. T. *ChemBioChem* **2004**, *5*, 1196.

(6) Boettger, D.; Bergmann, H.; Kuehn, B.; Shelest, E.; Hertweck, C. *ChemBioChem* **2012**, *13*, 2363.

(7) 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.

(8) Xu, W.; Cai, X.; Jung, M. E.; Tang, Y. *J. Am. Chem. Soc.* **2010**, *132*, 13604.

(9) Chooi, Y. H.; Tang, Y. *J. Org. Chem.* **2012**, *77*, 9933.

(10) Ma, S. M.; Li, J. W. H.; Choi, J. W.; Zhou, H.; Lee, K. K. M.; Moorthie, V. A.; Xie, X.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C.; Tang, Y. *Science* **2009**, *326*, 589.

(11) Kakule, T. B.; Sardar, D.; Lin, Z.; Schmidt, E. W. *ACS Chem. Biol.* **2013**, *8*, 1549.

(12) Kennedy, J.; Auclair, K.; Kendrew, S. G.; Park, C.; Vederas, J. C.; Richard Hutchinson, C. *Science* **1999**, *284*, 1368.

(13) Ames, B. D.; Chi, N.; Bruegger, J.; Smith, P.; Xu, W.; Ma, S.; Wong, E.; Wong, S.; Xie, X.; Li, J. W. H.; Vederas, J. C.; Tang, Y.; Tsai, S.-C. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 11144.

(14) Seshime, Y.; Juvvadi, P. R.; Tokuoka, M.; Koyama, Y.; Kitamoto, K.; Ebizuka, Y.; Fujii, I. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 3288.

(15) Maiya, S.; Grundmann, A.; Li, X.; Li, S.-M.; Turner, G. *ChemBioChem* **2007**, *8*, 1736.

(16) Tsunematsu, Y.; Fukutomi, M.; Saruwatari, T.; Noguchi, H.; Hotta, K.; Tang, Y.; Watanabe, K. *Angew. Chem., Int. Ed.* **2014**, *53*, 8475.

(17) Phillips, N. J.; Goodwin, J. T.; Fraiman, A.; Cole, R. J.; Lynn, D. G. *J. Am. Chem. Soc.* **1989**, *111*, 8223.

(18) 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.

(19) 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.

(20) 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.

(21) Schobert, R.; Schlenk, A. *Bioorg. Med. Chem.* **2008**, *16*, 4203.

(22) Kapur, S.; Chen, A. Y.; Cane, D. E.; Khosla, C. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 22066.

(23) Mofid, M. R.; Finking, R.; Marahiel, M. A. *J. Biol. Chem.* **2002**, *277*, 17023.

(24) Gu, L.; Eisman, E. B.; Dutta, S.; Franzmann, T. M.; Walter, S.; Gerwick, W. H.; Skiniotis, G.; Sherman, D. H. *Angew. Chem., Int. Ed.* **2011**, *50*, 2795.

(25) Kakule, T. B.; Jadulco, R. C.; Koch, M.; Janso, J. E.; Barrows, L. R.; Schmidt, E. W. *ACS Synth. Biol.* **2014**, DOI: 10.1021/sb500296p.

(26) Tokuoka, M.; Seshime, Y.; Fujii, I.; Kitamoto, K.; Takahashi, T.; Koyama, Y. *Fungal Genet. Biol.* **2008**, *45*, 1608.

(27) Wang, W.; Ballatori, N. *Pharmacol. Rev.* **1998**, *50*, 335.

(28) 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.

(29) Liu, X.; Walsh, C. T. *Biochemistry* **2009**, *48*, 8746.

(30) Sims, J. W.; Schmidt, E. W. *J. Am. Chem. Soc.* **2008**, *130*, 11149.

(31) Bruegger, J.; Haushalter, B.; Vagstad, A.; Shakya, G.; Mih, N.; Townsend, Craig A.; Burkart, M. D.; Tsai, S.-C. *Chem. Biol.* **2013**, *20*, 1135.

(32) Ma, Y.; Smith, L. H.; Cox, R. J.; Beltran-Alvarez, P.; Arthur, C. J.; Simpson, T. J. *ChemBioChem* **2006**, *7*, 1951.

(33) von Döhren, H. *Fungal Genet. Biol.* **2009**, *46*, S45.

(34) Gietz, R. D.; Schiestl, R. H.; Willems, A. R.; Woods, R. A. *Yeast* **1995**, *11*, 355.

(35) Ma, H.; Kunes, S.; Schatz, P. J.; Botstein, D. *Gene* **1987**, *58*, 201.