Thioesterase Domains of Fungal Nonreducing Polyketide Synthases Act as Decision Gates during Combinatorial Biosynthesis

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · JULY 2013

Impact Factor: 12.11 · DOI: 10.1021/ja4041362 · Source: PubMed

6 AUTHORS, INCLUDING:

Xu Yuquan
Chinese Academy of Agricultural Sciences
31 PUBLICATIONS **747** CITATIONS
SEE PROFILE

Tong Zhou
University of Minnesota Twin Cities
14 PUBLICATIONS **88** CITATIONS
SEE PROFILE

István Molnár
The University of Arizona
57 PUBLICATIONS **1,910** CITATIONS
SEE PROFILE

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Thioesterase Domains of Fungal Nonreducing Polyketide Synthases Act as Decision Gates during Combinatorial Biosynthesis

Yuquan Xu,[†,∥] Tong Zhou,[‡,∥] Shuwei Zhang,[‡,§] Li-Jiang Xuan,[§] Jixun Zhan,*[,‡] and István Molnár*[,†,∥]

[†]Natural Products Center, School of Natural Resources and the Environment, The University of Arizona, 250 East Valencia Road, Tucson, Arizona 85706, United States

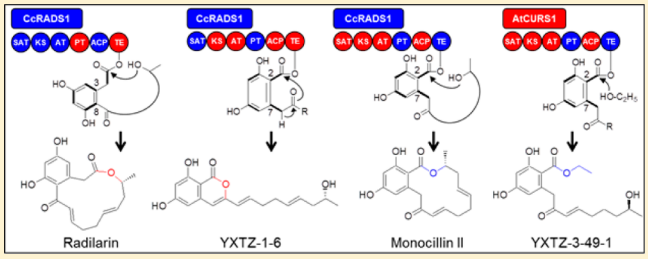
[‡]Department of Biological Engineering, Utah State University, 4105 Old Main Hill, Logan, Utah 84322, United States

[§]State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Shanghai Institute for Biological Sciences, Chinese Academy of Sciences, 501 Haike Road, Zhangjiang Hi-Tech Park, Shanghai 201203, China

[∥]Bio5 Institute, The University of Arizona, 1657 East Helen Street, Tucson, Arizona 85721, United States

Supporting Information

ABSTRACT: A crucial step during the programmed biosynthesis of fungal polyketide natural products is the release of the final polyketide intermediate from the iterative polyketide synthases (iPKSs), most frequently by a thioesterase (TE) domain. Realization of combinatorial biosynthesis with iPKSs requires TE domains that can accept altered polyketide intermediates generated by hybrid synthase enzymes and successfully release "unnatural products" with the desired structure. Achieving precise control over product release is of paramount importance with O−C bond-forming TE domains capable of macrocyclization, hydrolysis, transesterification, and pyrone formation that channel reactive, pluripotent polyketide intermediates to defined structural classes of bioactive secondary metabolites. By exploiting chimeric iPKS enzymes to offer substrates with controlled structural variety to two orthologous O−C bond-forming TE domains in situ, we show that these enzymes act as nonequivalent decision gates, determining context-dependent release mechanisms and overall product flux. Inappropriate choice of a TE could eradicate product formation in an otherwise highly productive chassis. Conversely, a judicious choice of a TE may allow the production of a desired hybrid metabolite. Finally, a serendipitous choice of a TE may reveal the unexpected productivity of some chassis. The ultimate decision gating role of TE domains influences the observable outcome of combinatorial domain swaps, emphasizing that the deduced programming rules are context dependent. These factors may complicate engineering the biosynthesis of a desired "unnatural product" but may also open additional avenues to create biosynthetic novelty based on fungal nonreduced polyketides.



INTRODUCTION

Fungi biosynthesize a large variety of structurally diverse polyketide natural products that mediate various ecological interactions as mycotoxins, virulence factors, and signaling molecules. These compounds are also of a prime interest to the pharmaceutical industry due to their various antibiotic, cancer cell antiproliferative, immunosuppressive, and enzyme inhibitory activities, and have provided the scaffolds for highly successful drugs as well as inspiration for chemical synthesis.[1,2] Among fungal polyketides, the 1,3-benzenediol lactone family (encompassing the resorcylic acid lactones (RALs) and the dihydroxyphenylacetic acid lactones (DALs)) offer particularly interesting pharmacophores for bioprospecting. For instance, the DAL 10,11-dehydrocurvularin (**1**, Figure 1) shows anti-inflammatory and immune system modulatory activities due to its inhibition of the inducible nitric oxide synthase.[3,4] Both **1** and the RAL monocillin II (**2**) are also potent inhibitors of the heat shock response, an evolutionarily conserved coping mechanism of eukaryotic cells. By disturbing protein homeostasis, both compounds display promising broad spectrum cancer cell antiproliferative activities in vitro.[5−8] Other RALs such as zearalenone and hypothemycin exhibit estrogen agonist and selective mitogen-activated protein kinase inhibitory activities, respectively.[9,10]

The biosynthesis of fungal polyketides is catalyzed by iterative polyketide synthases (iPKSs).[11] Although the architecture of iPKSs is similar to a single module of bacterial type I modular PKSs,[12] the iterative nature of these enzymes is nevertheless analogous to dissociated bacterial type II PKSs.[22] While the assembly of most fungal polyketides requires a single iPKS enzyme, the biosynthesis of the polyketide scaffold of both RALs and DALs involves a pair of collaborating iPKS enzymes (Figure 1).[13,18,19,23−26] These iPKSs each harbor a single set of ketoacyl synthase (KS), acyl transferase (AT), and acyl carrier protein (ACP) domains and use these domains

Received: April 25, 2013
Published: July 3, 2013

   dx.doi.org/10.1021/ja4041362 | J. Am. Chem. Soc. 2013, 135, 10783−10791

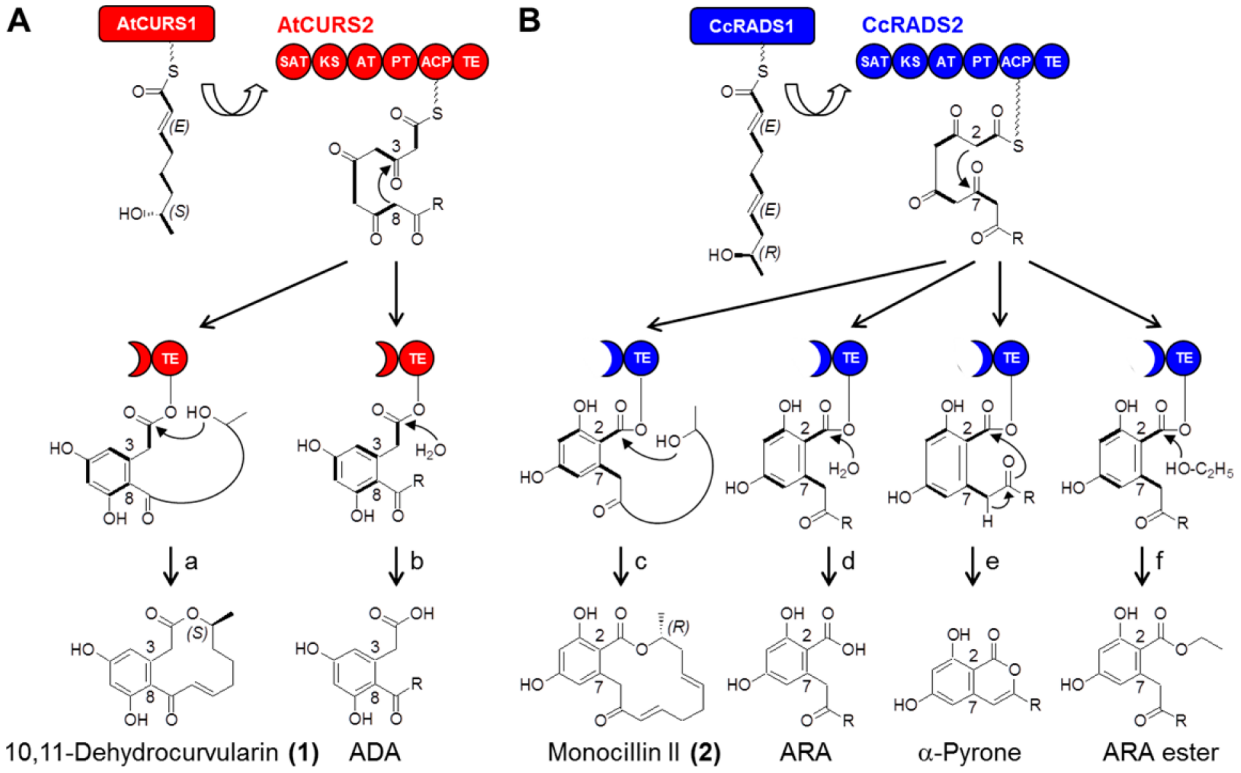


Figure 1. Role of TE domains during polyketide formation by collaborating iPKSs. (A) Biosynthesis of the DAL 10,11-dehydrocurvularin (**1**) involves the hrPKS AtCURS1 producing a reduced tetraketide starter unit that primes the nrPKS AtCURS2.[13,14] Following four additional condensation events with malonyl-CoA, the first ring is closed by the PT domain[15] in a C8−C3 aldol condensation event in the S-type folding mode.[13,16,17] Polyketide products may be released by intramolecular macrolactone formation to yield a DAL like **1** (route a) or by hydrolysis to form an acyl dihydroxyphenylacetic acid (ADA) (route b). (B) Biosynthesis of monocillin II (**2**) also involves a sequentially acting, collaborating nrPKS pair.[18] The pentaketide[19] produced by the hrPKS CcRADS1 is further elaborated by the nrPKS CcRADS2, including C2−C7 condensation in the F-type folding mode.[16,20] Product release by macrolactone formation yields a RAL like **2** (route c); hydrolysis provides acyl resorcylic acid (ARA) products (route d); α-pyrones are formed by the attack of the C9 enol on the oxoester (route e); and ARA ethyl esters are produced by utilizing ethanol as the nucleophile (route f).[18,21] C−C bonds in bold indicate intact acetate equivalents (malonate-derived C_2 units) incorporated into the polyketide chain by the iPKSs.

iteratively to conduct recursive thio-Claisen condensations of malonyl-CoA extender units. First, a highly reducing iPKS (hrPKS) assembles a reduced linear polyketide chain (Figure 1). The hrPKS harbors ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) domains to reduce the nascent β-ketoacyl intermediates in a context-dependent manner to execute a cryptic biosynthetic program.[2,27] Next, the reduced polyketide chain is directly transferred from the hrPKS onto a nonreducing iPKS (nrPKS) by the starter unit: ACP transacylase (SAT) domain of the nrPKS.[14] After a programmed number of further chain extensions starting with this advanced priming unit, the nrPKS directs ring closure by regiospecific aldol condensation to yield the 1,3-benzenediol moiety, catalyzed by the product template (PT) domain.[15,17,20,28] Finally, the thioesterase (TE) domain is responsible for the release of the RAL or DAL product from the nrPKS by closure of the bridging macrolactone ring.[29]

A crucial step of the programmed biosynthesis of fungal polyketide natural products is the release of the finished polyketide intermediate from the iPKS enzyme, most frequently by a TE domain.[2,23,30,31] iPKS TEs feature an α/β-hydrolase catalytic core with a Ser/His/Asp catalytic triad and a flexible lid loop that closes the substrate binding chamber.[12,32] Most TEs from fungal iPKSs catalyze product release by intramolecular C−C bond formation through Claisen/Dieckmann cyclization (TE/CLC domains),[28,30,33] in some cases coupled to product truncation by deacylation.[32] Nevertheless, a smaller number of fungal iPKSs, including those involved in RAL or DAL biosynthesis,[13,18,24,25,34] feature TE domains that catalyze O−C bond formation through macrolactone closure, hydrolysis, and ester or pyrone formation.[23,29,31] These O−C bond-forming iPKS TEs are divergent from the TE/CLC domains, with identities <25%. They also share little sequence identity with the O−C bond-forming TEs of the prokaryotic type I modular PKSs and NRPSs, in spite of their functional similarities.[18,23,29,35−37]

Combinatorial polyketide biosynthesis requires TE domains that can successfully release "unnatural products" by accepting and processing altered intermediates generated by hybrid synthase enzymes.[12] Achieving precise control over the mode of product release (macrocyclization, hydrolysis or other mechanisms) is also of paramount importance, considering that this process channels pluripotent, unstable intermediates toward varied polyketide structural classes with defined biological activities. The present work investigated the surprisingly different product release specificities of two closely related O−C bond-forming macrolactone synthase TE domains, one from the nrPKS CcRADS2 involved in the biosynthesis of the RAL **2**,[18] and the other from AtCURS2 yielding the DAL **1**.[13] By exploiting a large array of chimeric nrPKS enzymes, we show that these TE domains act as

dx.doi.org/10.1021/ja4041362 | J. Am. Chem. Soc. 2013, 135, 10783−10791

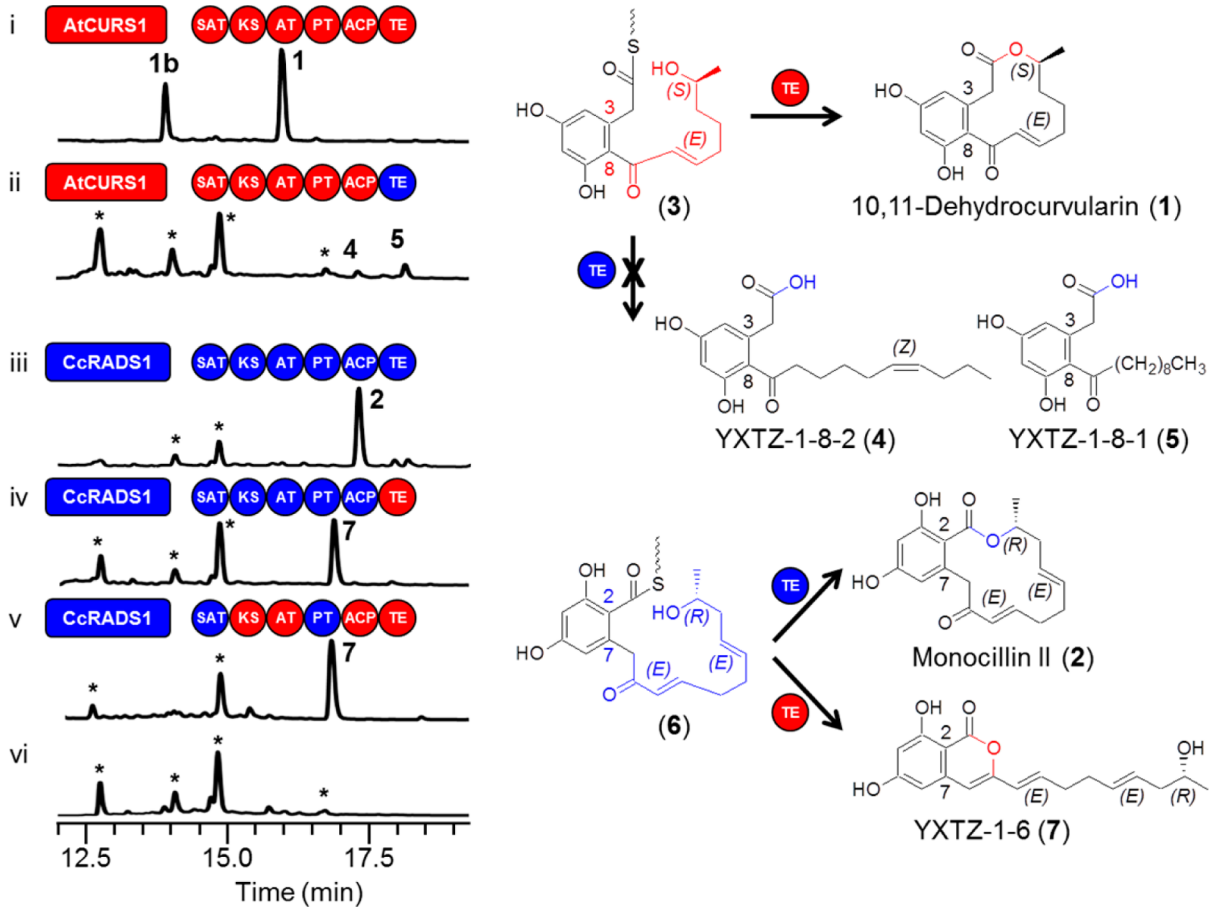


Figure 2. $TE_{AtCURS2}$ and $TE_{CcRADS2}$ are not interchangeable. Product profiles (HPLC traces recorded at 300 nm) of *S. cerevisiae* BJ5464-NpgA[19,41] cotransformed with engineered or native hrPKS-nrPKS pairs: (i) YEpAtCURS1 and YEpAtCURS2; (ii) YEpATCURS1 and YEpYX8; (iii) YEpCcRADS1 and YEpCcRADS2; (iv) YEpCcRADS1 and YEpYX6; (v) YEpCcRADS1 and YEpYX65; (vi) no-PKS control. The hrPKS-generated portions of the proposed ACP-bound thioester intermediates are color-coordinated with the synthase. The peak in trace (i) labeled with **1b** corresponds to 11-hydroxycurvularin, a spontaneous hydration product of **1**.[13] Peaks labeled with a star represent yeast metabolites unrelated to the iPKS products. Domains drawn as red circles are derived from AtCURS2, while those in blue are from CcRADS2.

nonequivalent decision gates determining both the shape and the yield of polyketide products.

RESULTS

Macrolactone-Forming TE Domains in AtCURS2 and CcRADS2. The O−C bond-forming TE domains of AtCURS2 and CcRADS2 share 52% identity and 64% similarity,[13,18] and comparable levels of similarities are also evident with other RAL nrPKSs.[24−26] However, sequence identities do not exceed 25% with the C−C bond-forming TE/CLC domains, such as the noranthrone (aflatoxin) synthase nrPKS whose structure has been experimentally determined.[30] All RAL/DAL TE domains harbor an invariable Ser/His/Asp catalytic triad (AtCURS2: $S^{1880}/H^{2055}/D^{1907}$; CcRADS2: $S^{1930}/H^{2109}/D^{1957}$). In the first half-reaction, the polyketide intermediate undergoes transacylation to the Ser nucleophile which is activated by the His catalytic base with the help of the Asp. In the second half-reaction, the resulting acyl-O-TE oxoester is attacked by a secondary alcohol of the polyketide chain to afford product release by macrolactone formation. Product release by hydrolysis, ester formation, or α-pyrone (isocoumarin) formation has also been observed.[13,17,18,21]

The thioester intermediates that serve as substrates for the O−C bond-forming TE domains are structurally complex and may be susceptible to spontaneous rearrangements. To gain insight into the programming of these TE domains, we decided to generate such substrates in situ and to present them to the DAL-forming $TE_{AtCURS2}$ and the RAL-forming $TE_{CcRADS2}$. Since *in trans* complementation of dissected domains is known to incur a penalty in product yield and fidelity,[1,17,20,38] the TE domains were grafted onto various nrPKSs, created from AtCURS2 and CcRADS2 by domain swaps. These hybrid nrPKSs were then paired with the AtCURS1 or the CcRADS1 hPKS to create RAL or DAL synthase iPKS pairs. Noncognate iPKS partners were coupled by interchanging the nrPKS SAT domains.[14,39,40] Polyketide production was reconstituted in vivo by expressing these recombinant hrPKS + nrPKS pairs from compatible plasmids in the host *Saccharomyces cerevisiae* BJ5464-NpgA.[13,17,19,41]

Deletion or inactivation of Claisen cyclase TE domains (TE/CLC) of nrPKSs has been shown to yield variable amounts of α-pyrone shunt metabolites by spontaneous O−C cyclization.[1,28,30,32,42] To evaluate the extent of spontaneous product release in the absence of the macrocycle-forming TE domains, we constructed truncated AtCURS2 and CcRADS2 versions. Deletion of $TE_{AtCURS2}$ completely eliminated polyketide production, while a TE-less CcRADS2 produced only trace amounts of pyrone **7** (<0.2 mg/L, Figure S1, traces i and ii, Supporting Information). This result indicates that spontaneous product release may not efficiently compensate for the

dx.doi.org/10.1021/ja4041362 | J. Am. Chem. Soc. 2013, 135, 10783−10791

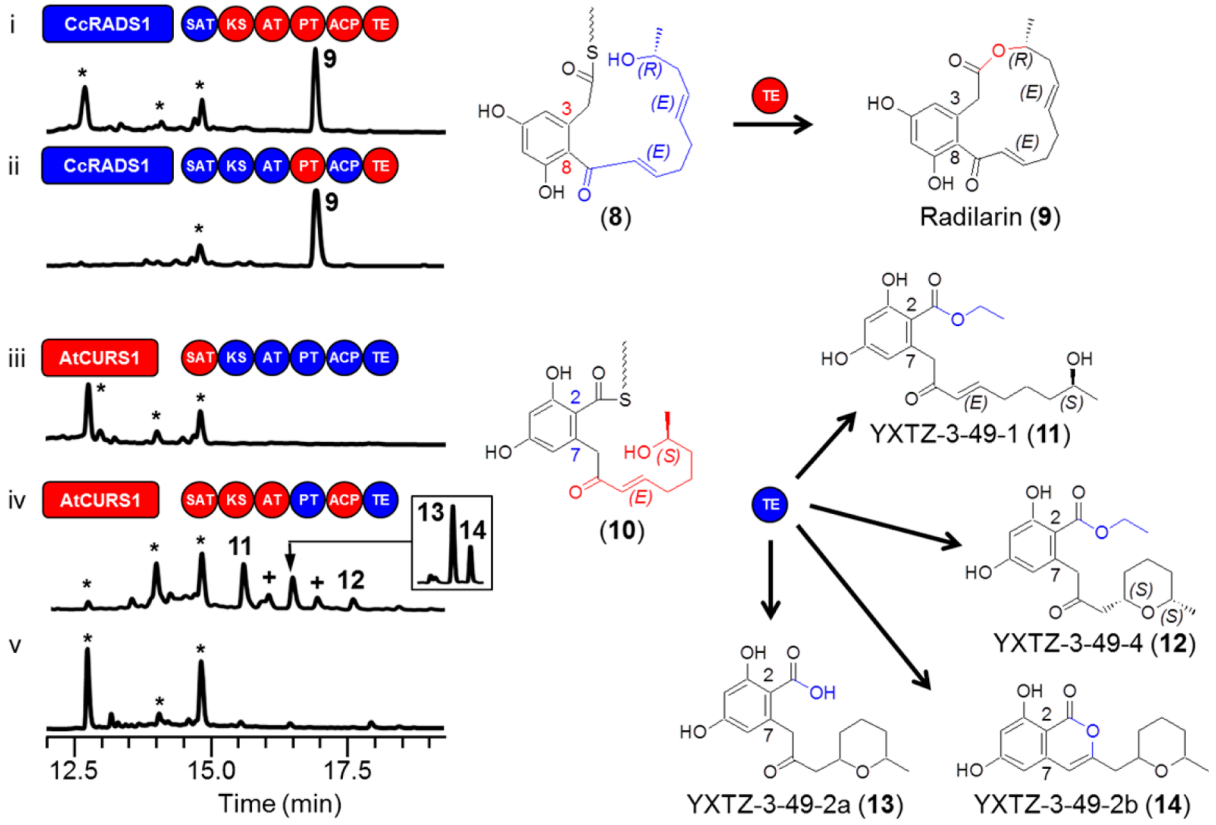


Figure 3. Influence of the reduced acyl priming chain on TE-catalyzed product formation. Product profiles of *S. cerevisiae* BJ5464-NpgA[19,41] cotransformed with hrPKS−nrPKS pairs: (i) YEpCcRADS1 and YEpYX24; (ii) YEpCcRADS1 and YEpYX67; (iii) YEpAtCURS1 and YEpYX12; (iv) YEpAtCURS1 and YEpYX49; (v) no-PKS control. The hrPKS-generated portions of the proposed ACP-bound thioester intermediates are color-coordinated with the synthase. Domains drawn as red circles are derived from AtCURS2, while those in blue are from CcRADS2. Inset: The mixture of **13** and **14** may be separated by HPLC using a linear gradient of 5% to 95% CH_3CN in H_2O over 40 min. *Yeast metabolites unrelated to the iPKS products. [+]Compounds decomposed during isolation.

absence of a functional TE with these synthases, similar to what was seen with the CTB1 nor-toralactone synthase or the Pks1 melanin synthase in the absence of their TE/CLC domains.[32,43] Thus, the emergence of polyketides in significant amounts during fermentations with recombinant yeasts may be attributed to TE-catalyzed product release in the case of the RAL and DAL synthases, and the polyketide scaffolds of these products could be used to deduce the TE function and specificity.

$TE_{AtCURS2}$ and $TE_{CcRADS2}$ Are Not Equivalent. Surprisingly, replacement of the TE domain of AtCURS2 with $TE_{CcRADS2}$ eliminated the formation of cognate products derived from the expected thioester intermediate **3** (Figure 2, trace ii). This was not the consequence of an incapacitated nrPKS enzyme as small amounts of the acyl dihydroxyphenylacetic acids (ADA) **4** and **5** were still produced by the strain (0.2 and 0.5 mg/L for **4** and **5**, respectively). These products presumably originate from C_{10} carboxylic acid priming units from fatty acid biosynthesis or degradation in the yeast host. Formation of these and similar ADA in similar yields by hydrolysis (Figure 1, route b) have already been observed with intact AtCURS2 in the absence of its hrPKS partner,[13] similar to the formation of acyl resorcylic acid (ARA) products with the unaccompanied zearalenone nrPKS.[23] The lack of the synthesis of cognate products was not due to unproductive interactions between the $TE_{CcRADS2}$ and the ACP or the KS-AT chassis of AtCURS2: replacement of these domains with their CcRADS2-derived equivalents failed to restore the production of **1** (Figure S1, traces iii and iv, Supporting Information).

The converse experiment, replacement of the TE domain of CcRADS2 with $TE_{AtCURS2}$, eliminated the formation of **2** but afforded the isocoumarin **7** in a high yield (Figure 2 trace iv, 3 mg/L). Thus, although $TE_{AtCURS2}$ was able to process intermediate **6** by pyrone formation using the C9 enol as a nucleophile (Figure 1 route e),[42] macrocycle formation was apparently inhibited. Replacement of the KS, AT and ACP chassis of the enzyme played no role in determining the nature or the amount of the product (Figure 2 trace v, yield of **7** approximately 3.5 mg/L). Taken together, these experiments show that in spite of the high similarity and phylogenetic proximity of $TE_{AtCURS2}$ and $TE_{CcRADS2}$, these enzymes are not freely interchangeable during combinatorial biosynthesis for creating RAL or DAL products.

Role of the hrPKS-Derived Reduced Acyl Chain. In the previous set of experiments, the TE domains were challenged with the putative ACP-bound acyl thioesters **3** and **6**. These thioesters differ in the length and structure of the acyl chain assembled by the hrPKS (a tetraketide for **3** and a pentaketide for **6**, Figure 1), as well as in the aldol condensation register of the 1,3-benzenediol moiety generated by the nrPKS.[16,17,20] To disentangle the roles of these two variables in TE substrate recognition and processing, we have created hybrid iPKS pairs where the priming unit for RAL/DAL biosynthesis is assembled by the hrPKS from the other biosynthetic system, while the aldol condensation register is maintained for the selected TE.

dx.doi.org/10.1021/ja4041362 | J. Am. Chem. Soc. 2013, 135, 10783−10791

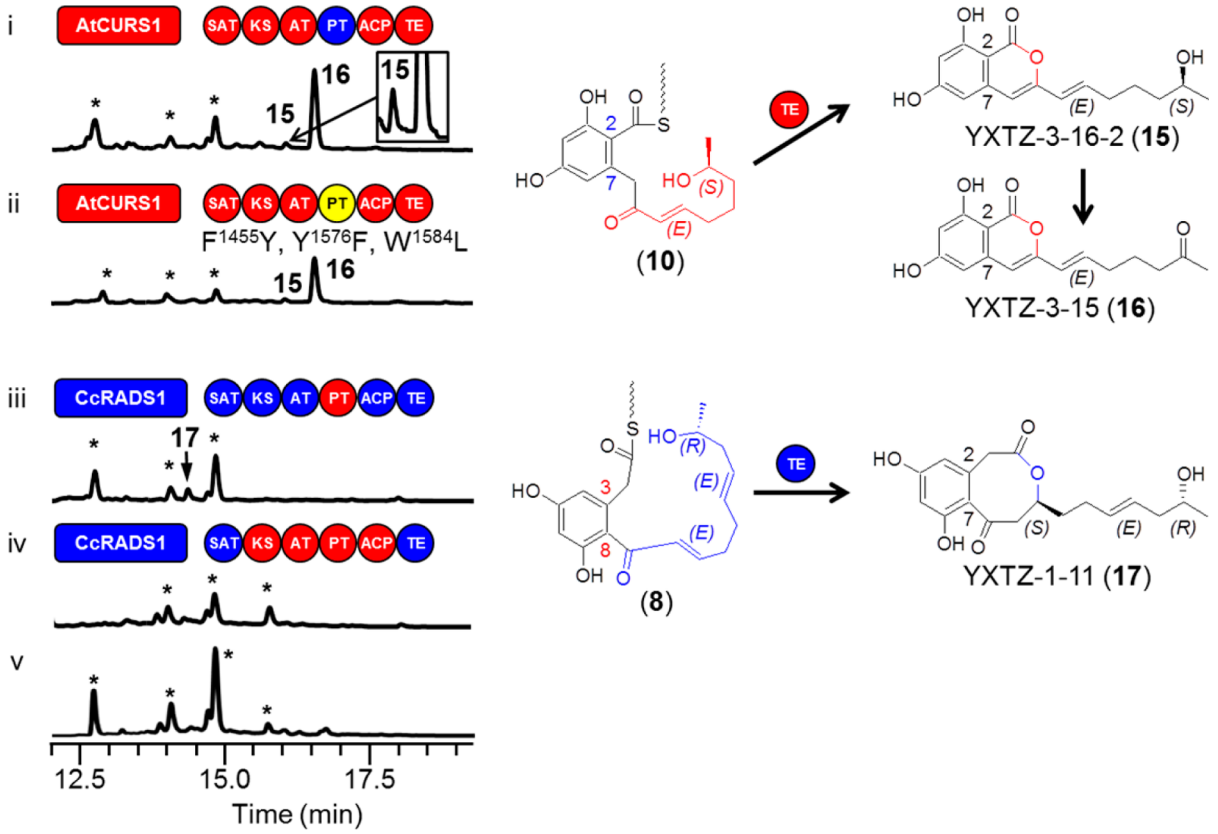


Figure 4. Influence of the aldol register of the first ring on TE-catalyzed product formation. Product profiles of *S. cerevisiae* BJ5464-NpgA[19,41] cotransformed with hrPKS–nrPKS pairs: (i) YEpAtCURS1 and YEpAtCURS2-PTCcRADS2; (ii) YEpAtCURS1 and YEpAtCURS2-($F^{1459}Y$,$Y^{1576}F$,$W^{1584}L$); (iii) YEpCcRADS1 and YEpYX11; (iv) YEpCcRADS1 and YEpYX33; (v) no-PKS control. *Yeast metabolites unrelated to the iPKS products. The hrPKS-generated portions of the proposed ACP-bound thioester intermediates are color-coordinated with the synthase. Domains drawn as red circles are derived from AtCURS2, while those in blue are from CcRADS2. The PT domain in (ii) drawn as a yellow circle is the $F^{1459}Y/Y^{1576}F/W^{1584}L$ mutant of $PT_{AtCURS2}$.

Thus, the presumed thioester **8** (Figure 3) is formed from a reduced pentaketide chain as in radicicol, but the register of aldol condensation is C8–C3 (S-type folding) as in **1**.[16,17] $TE_{AtCURS2}$ had no difficulty in processing thioester **8** to the novel DAL radilarin (Figure 3 trace i, 9 mg/L) by macrocycle formation (Figure 1 route a). As before, replacing the KS/AT/ACP chassis did not influence product yield (Figure 3 trace ii). Biosynthesis of **9** is remarkable as no 14-member DAL is known from natural sources; thus, **9** represents a unique class of truly unnatural products.

Conversely, challenging TECcRADS2 with an intermediate that features a shorter acyl chain on a benzenediol formed in the C2–C7 register (thioester **10** with F-type folding)[16,17] seemed initially unproductive when using the CcRADS2 chassis (Figure 3, trace iii). Fusing this chassis with $TE_{AtCURS2}$ or the ACP-TE didomain from AtCURS2 was similarly unproductive (Figure S1, traces v and vi, Supporting Information), hinting at a suboptimal interaction of the $KS_{CcRADS2}$ domain with the incoming tetraketide presented by the $SAT_{AtCURS2}$.[38] However, replacing the KS/AT/ACP chassis with that of AtCURS2 (Figure 3, trace iv) led to the production of a variety of ARA products by hydrolysis (**13**, 0.4 mg/L), or nucleophilic attacks by ethanol (**11** and **12**, 1.5 and 0.5 mg/L, respectively) or by the C9 enol (**14**, 0.2 mg/L). Product release by transesterification with alcohols (mainly ethanol) present in the media has been documented for AtCURS2 and the zearalenone nrPKS.[13,23,29] Thus, $TE_{CcRADS2}$ is capable of downloading a C2–C7 intermediate with a shorter acyl chain, but macrolactone formation with the foreign substrate is apparently inhibited. Taken together, these experiments indicate that product release by these TE domains in a combinatorial biosynthetic context is permissive to variation of the length and identity of the priming acyl chain, as has been shown with the isolated zearalenone synthase nrPKS[23] and with dissected nrPKS hybrids in vitro.[43] However, macrocycle formation may be limited by the correct positioning of the distal alcohol of the polyketide chain for nucleophilic attack on the oxoester within the TE catalytic chamber.

Role of S- or F-Type Folding of the First Ring. Next, we investigated the role of the configuration of the substituted 1,3-benzenediol moiety on product formation by the TE domains. Thus, we created hybrid biosynthetic systems where each TE is challenged with a thioester intermediate featuring a hrPKS-derived acyl chain cognate to that TE, but with a first ring that had formed in the opposite aldol register.[13] When intermediate **10** with a C2–C7 geometry was offered to $TE_{AtCURS2}$, isocoumarins **15** and **16** were formed in good yields (0.3 and 3 mg/mL, respectively, Figure 4, trace i), as already observed in our recent work on the rational engineering of aldol cyclization by PT swaps.[17] **16** is formed by a chance oxidation of the C15-OH of **15**, catalyzed by an endogenous host enzyme.[17] Conversely, offering the C8–C3-cyclized intermediate **8** to $TE_{CcRADS2}$ afforded small amounts of **17** (Figure 4, trace iii, 0.3 mg/L), where the C8–C3 dihydroxyphenylacetic acid moiety is

dx.doi.org/10.1021/ja4041362 | J. Am. Chem. Soc. 2013, 135, 10783–10791

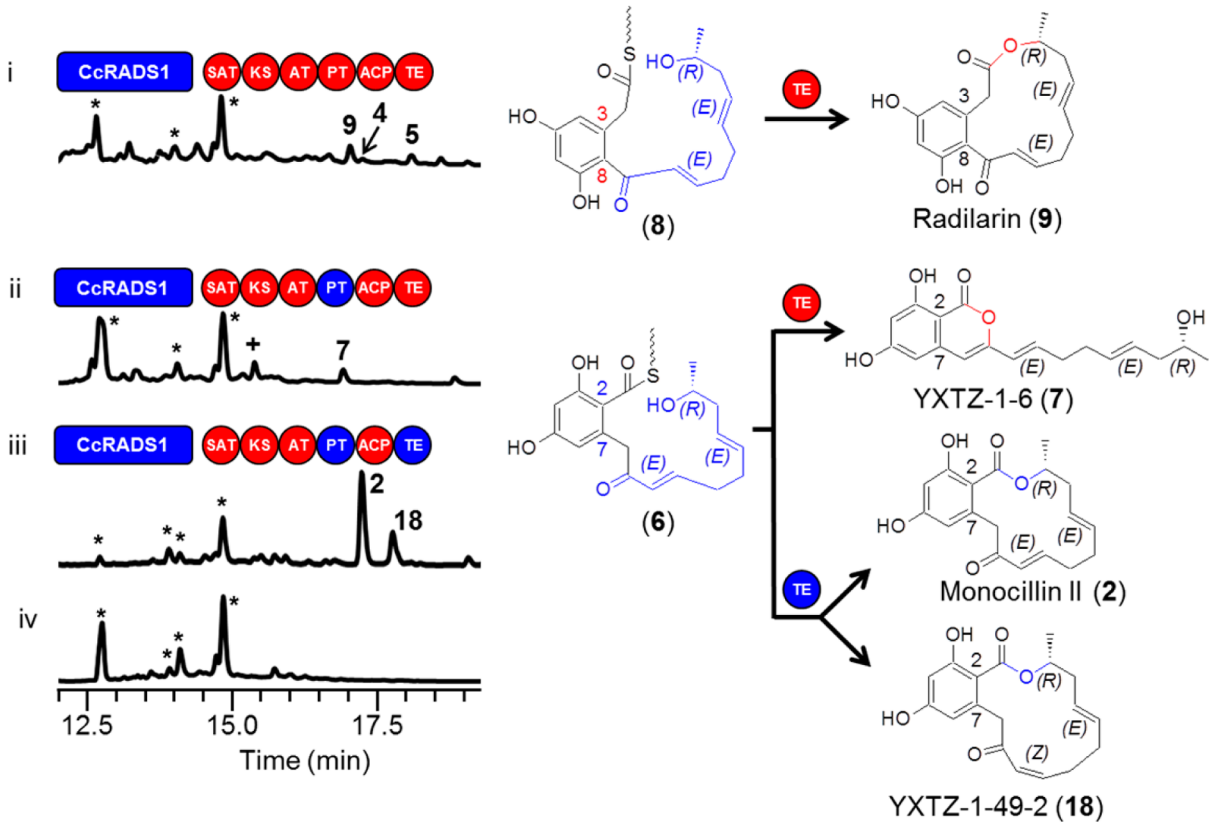


Figure 5. TEs may facilitate unexpected biosynthetic interactions. Product profiles of *S. cerevisiae* BJ5464-NpgA[19,41] cotransformed with engineered hrPKS–nrPKS pairs as indicated: (i) YEpCcRADS1 and YEpAtCURS2; (ii) YEpCcRADS1 and YEpAtCURS2-$PT_{CcRADS2}$; (iii) YEpCcRADS1 and YEpYX49; (iv) no-PKS control. The hrPKS-generated portions of the proposed ACP-bound thioester intermediates are color-coordinated with the synthase. Domains drawn as red circles are derived from AtCURS2, while those in blue are from CcRADS2. *Yeast metabolites unrelated to the iPKS products. ⁺Compound decomposed during isolation.

bridged by an 8-membered lactone.[17] Whether the formation of the 4-oxo-2-oxacyclooctane ring of **17** by the facile attack of the C1 carboxyl on the C11 enone is spontaneous or involves $TE_{CcRADS2}$ is unclear at this point. Replacing the KS/AT/ACP chassis with the one from AtCURS2 failed to increase product formation (Figure 4, trace iv). In any case, low level production of **17**, or absence of product formation is not due to an enfeebled chassis as shown by vigorous product formation when these same chassis are paired with $TE_{AtCURS2}$ (Figure 3, traces i and ii). Collectively, these experiments indicate that the register of the PT- catalyzed aldol ring formation is an important determinant for the productivity of a given TE in a combinatorial context. Thus, $TE_{CcRADS2}$ is incompetent in forming the desired DAL (or even an ADA) with the S-type substrate **8** even if the priming acyl chain is appropriate for this TE. Similarly, although $TE_{AtCURS2}$ can effectively download C2–C7-cyclized intermediates like **6** and **10** to form pyrone products like **7** and **15** (compare Figure 2, traces iv and v, with Figure 4, trace i), macrolactone formation is inhibited. Future studies would need to determine whether the improper configuration of the substrate itself, or the absence of proper protein–protein contacts between the PT and TE domains,[12] is more important to hinder macrocycle formation. However, it is remarkable that macrocyclization of thioester intermediate **10** is still inhibited with an AtCURS2 derivative harboring only three point mutations in its PT domain (Figure 4, trace ii). These point mutations in the active site pocket of the PT domain are sufficient to bring about an F-type folding and C2–C7 cyclization, instead of the native S-type folding and C8–3 first ring closure.[17] However, they are unlikely to disturb the native domain–domain interactions between $PT_{AtCURS2}$ and $TE_{AtCURS2}$, implying an essential role for direct substrate recognition to determine the mode of product release by the TE.

Combinations of Altered Acyl Chains and Aldol Registers. As shown above, combination of a longer (pentaketide) acyl primer chain and a noncognate first ring geometry (as in **6**) is not an impediment to efficient product release by pyrone formation with $TE_{AtCURS2}$ (Figure 2, traces iv and v). Conversely, both macrolactone formation and product downloading is eliminated when an intermediate with an S-type first ring and a shorter starter chain is presented to $TE_{CcRADS2}$ (Figure 2, trace ii, intermediate **3**). This indicates that the effects of acyl chain alterations are additive with the impacts of the modification of the first ring register. Thus, $TE_{AtCURS2}$-catalyzed product downloading is efficient for both octa- and nonaketide intermediates regardless of aldol register (**1**, **7**, **9**, and **15**), but macrocycle formation is only detected with S-type (C8–C3) thioester intermediates (**1** and **9**). $TE_{CcRADS2}$ is a more stringent catalyst that prefers a nonaketide intermediate. It may still download an octaketide but is apparently unable to form a macrolactone with such a shorter intermediate (**11** to **14**, Figure 3, trace iv). S-Type (C8–C3) aldol intermediates are not preferred substrates either, and products may only be released if the intermediate is a nonaketide. Even then, products emerge only in small amounts by hydrolysis (**4** and **5**, Figure 2,

dx.doi.org/10.1021/ja4041362 | J. Am. Chem. Soc. 2013, 135, 10783–10791

trace ii) or by formation of an 8-membered ring (**17**, Figure 4, trace iii).

TE Domains May Reveal Unexpected Biosynthetic Interactions. While investigating the substrate preferences and programming rules of these TEs in our in vivo reconstituted systems, we also noticed that the presence of a heterologous TE domain may reveal unexpected product formation in combinatorial contexts. Combining the CcRADS1 hPKS with the AtCURS2 nrPKS provided small amounts of the 14-membered DAL radilarin (**9**) and the fatty acyl-derived ADA products **4** and **5** (Figure 5 trace i, 0.5, 0.1, 0.3 mg/L for **9**, **4** and **5**, respectively). Replacing the PT domain of AtCURS2 with $PT_{CcRADS2}$ similarly yielded minor amounts of the isocoumarin **7** (Figure 5 trace ii, 0.5 mg/L), featuring the engineered C2−C7 first ring and a pentaketide acyl chain. These two experiments indicated that the SAT domain of AtCURS2 is somewhat promiscuous and is able to recognize and transfer the pentaketide product of CcRADS1 onto AtCURS2. In both cases, $TE_{AtCURS2}$ appeared permissive to the formation of these minor products. Surprisingly, when $TE_{CcRADS2}$ was introduced into the latter construct, **2** was formed in large amounts (Figure 5, trace iii, 6 mg/L), with **18** as the minor product (1 mg/mL). Compound **18** features a *cis* double bound between C10 and C11, likely formed by an endogenous enzyme of the host. Thus, a TE domain with an inbuilt preference for the acyl thioester presented to it by the nrPKS can overrule the expected impediment to product formation caused by the imperfect pairing of the hrPKS and the nrPKS by a mismatched SAT domain.

■ DISCUSSION

Very recently, Yeh et al. replaced the TE domain of an nrPKS responsible for the production of 2,4-dihydroxy-3,5,6-trimethylbenzaldehyde with TE domains from several other nrPKSs and concluded that the phylogenetic position of the nrPKSs (and by extension, their TE domains) is a good predictor for the success of product formation.[44] They posited that close phylogenetic relationship translates to better domain−domain interactions and, further, to successful product release in engineered synthases. In contrast, our results emphasize that the shape and size of the polyketide substrate offered to the TE by the rest of the synthase is the crucial determinant for product release. This interpretation is also in agreement with the results of Yeh et al.[44] Thus, the relatedness of the product structures is a better predictor of combinatorial success than phylogenetic relationship of the TEs/nrPKSs. This view emphasizes that TE domains are discriminative catalysts, and even more importantly, that they act as decision gates that determine both the final shape of the product and the extent of turnover by nrPKS enzymes. This control role is different from an already-recognized housekeeping role responsible for the restoration of the biosynthetic flux by the removal of stalled, aberrant acyl thioesters, as seen with the noranthrone synthase TE.[38]

In another very recent set of elegant experiments, Vagstad et al. investigated the effect of altered chain length on the ability of TE/CLC domains to release products derived from completely unreduced poly-β-ketoacyl intermediates.[43] They offered a heptaketide chain produced by the purified SAT-KS-AT chassis of the CTB1 nor-toralactone synthase to isolated PT, ACP, and TE domain sets from other nrPKSs in vitro. Product release by Claisen/Dieckman cyclization, the native mode of the investigated TE domains, was detected to varying degrees with three PT-ACP-TE sets. Not surprisingly, the most efficient combination turned out to be the one where the native substrate of the incoming PT-ACP-TE set was identical to that offered by the heterologous SAT-KS-AT chassis. A PT-ACP-TE set that "expected" a nonaketide was barely functional with the heptaketide substrate. Only TE-independent spontaneous product release was observed with another PT-ACP-TE set that generates a tricyclic product in its native context. Thus, this in vitro reconstituted system supports the requirement of TEs to be presented with thioester intermediates similar to their native substrates, while showed some permissiveness in terms of substrate chain length. Alteration of cyclization modes was not investigated independent of chain length in this study as a factor in TE-catalyzed product release, nor was the production of novel unnatural polyketides detected with these in vitro domain combinations.

The decision gating role of the TE domains, revealed in our study, has significant consequences over and above the perhaps expected lower product yields due to incomplete processing of a foreign substrate. Inappropriate choice of a TE could completely eradicate product formation in an otherwise highly productive chassis, as seen when $TE_{CcRADS2}$ failed to process the carrier protein thioester **3** (compare Figure 2 trace ii and trace i). Conversely, a judicious choice of a TE may allow the production of a desired hybrid metabolite, as seen when $TE_{AtCURS2}$, but not $TE_{CcRADS2}$ successfully processed thioester **8** leading to the effective biosynthesis of the novel unnatural product radilarin (**9**) (Figure 3, trace ii, vs Figure 4, trace iii). Finally, a serendipitous choice of a TE may reveal the unexpected productivity of some chassis, as seen when $TE_{CcRADS2}$, but not $TE_{AtCURS2}$, permitted the high level production of **2** in an "uncoupled" hrPKS-nrPKS pair (Figure 5, trace iii vs trace ii). These observations point to the complexity of the decision gating role of the nrPKS TE domains. First, overall product flux is determined by these TEs, similar to the situation in modular PKSs of bacteria where the turnover rate of TE-catalyzed product release regulates extension cycles in "congested" upstream modules.[45] Next, the TE domains of the nrPKSs investigated in this study also display context-dependent release mechanisms, yielding macrolactones, carboxylic acids and their esters, and pyrones. Thus, the TE decision gate joins previously identified control points for polyketide assembly on nrPKSs: selection of an appropriate primer unit by the SAT; acyl chain length monitoring and kinetic control of chain extension vs cyclization by the KS; and regiospecific cyclization by the PT domain.[17,38] This ultimate decision gate influences the observable outcome of combinatorial domain swaps, making interpretation of such experiments more complex, and emphasizing that the deduced programming rules for "rational domain heterocombinations"[43] are, in fact, context dependent.

■ CONCLUSIONS

Understanding the inbuilt programming differences of product release from engineered fungal iPKSs is fundamental to guide efforts to generate novel chemical diversity from natural fungal polyketides. By exploiting an array of chimeric iPKS enzyme pairs during the reprogrammed biosynthesis of unnatural benzenediol lactone products, we show that even closely related, orthologous O−C bond-forming TE domains may yield different product release outcomes. Influenced by both the chain lengths and the geometries of the first rings of the polyketide intermediates presented by the chassis of the

synthase, these TE domains determine both the shape and the yield of polyketide products, and may obstruct, or conversely, facilitate product formation in an idiosyncratic, context-dependent manner. Thus, TE domains of fungal iPKSs act as nonequivalent decision gates that direct the malleable, reactive intermediates toward defined structural and functional classes of the polyketide space.

Importantly, workers of combinatorial biosynthesis need to be cognizant of the decision gating role of the TE domains of collaborating iPKS pairs, and by extension, other nrPKSs. Thus, it will be necessary to consider both the intrinsic substrate repertoire of the given TE, and the modulation of the product release mode by the incoming polyketide intermediate among hydrolysis, pyrone formation, transesterification, and macrolactone formation as demonstrated in this study, and also by Claisen/Dieckmann cyclization as shown by others.[43] These factors may complicate engineering approaches for the biosynthesis of a desired "unnatural product" but may also open additional avenues for the creation of biosynthetic novelty based on fungal nonreduced polyketides.

MATERIALS AND METHODS

Strains and Culture Conditions. *E. coli* DH10B and plasmid pJET1.2 (Fermentas) were used for routine cloning and sequencing. The yeast−*E. coli* shuttle vectors YEpADH2p−FLAG-URA and YEpADH2p−FLAG-TRP with the URA3 or with the TRP1 selectable markers[13,19] were used for the expression of hrPKS−nrPKS pairs in *Saccharomyces cerevisiae* BJ5464−NpgA (MATα ura3-52 his3-Δ200 leu2-Δ1 trp1 pep4::HIS3 prb1 Δ1.6R can1 GAL).[41,46] Primers used in this study and details on the construction of gene variants are described in the Materials and Methods in the Supporting Information. Polyketide production was analyzed in three to five independent transformants for each recombinant yeast strain by small scale fermentation. Fermentations with representative isolates were repeated at least three times to confirm results and scaled up to 1−5 L to isolate products as described.[13,17]

Chemical Characterization of Polyketide Products. Optical rotations were recorded on a Rudolph Autopol IV polarimeter using a 10-cm microcell. Circular dichroism (CD) spectra were acquired with a JASCO J-810 instrument using a path length of 1 cm. ^{1}H, ^{13}C, and 2D NMR (COSY, HSQC, HMBC, ROESY) spectra were obtained in CD_3OD or C_5D_5N on a JEOL ECX-300 spectrometer. ESI-MS data were collected on an Agilent 6130 Single Quad LC-MS. Accurate mass measurements were performed with matrix assisted laser desorption/ionization (MALDI) on a Bruker Ultraflex III MALDI TOF-TOF instrument, or with electrospray ionization (ESI) on a Bruker 9.4 T Fourier transform ion-cyclotron resonance (FT-ICR) instrument. Low-resolution tandem mass spectra were obtained on a Thermoelectron LCQ instrument, and high-resolution MS/MS was performed by FT-ICR using ESI. See the Supporting Information for details.

ASSOCIATED CONTENT

Supporting Information

Additional methods, results, Figures S1−S9, primer list, and spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

imolnar@email.arizona.edu; jixun.zhan@usu.edu

Author Contributions

‖These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by grants from the National Science Foundation (MCB-0948751 to I.M.), the National Institutes of Health (AI065357 to J.Z.), and Utah State University (Seed Program to Advance Research Collaborations, to J.Z.). We are grateful to Professors Nancy A. Da Silva (University of California, Irvine) and Yi Tang (University of California, Los Angeles) for providing the yeast host strain and expression vectors and to Árpád Somogyi (University of Arizona, Tucson) for the MS/MS spectra and accurate mass measurements.

REFERENCES

(1) Crawford, J. M.; Townsend, C. A. *Nat. Rev. Microbiol.* **2010**, *8*, 879.

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

(3) Schmidt, N.; Pautz, A.; Art, J.; Rauschkolb, P.; Jung, M.; Erkel, G.; Goldring, M. B.; Kleinert, H. *Biochem. Pharmacol.* **2010**, *79*, 722.

(4) Elzner, S.; Schmidt, D.; Schollmeyer, D.; Erkel, G.; Anke, T.; Kleinert, H.; Förstermann, U.; Kunz, H. *ChemMedChem* **2008**, *3*, 924.

(5) Santagata, S.; Xu, Y. M.; Wijeratne, E. M.; Kontnik, R.; Rooney, C.; Perley, C. C.; Kwon, H.; Clardy, J.; Kesari, S.; Whitesell, L.; Lindquist, S.; Gunatilaka, A. A. L. *ACS Chem. Biol.* **2012**, *7*, 340.

(6) McLellan, C. A.; Turbyville, T. J.; Wijeratne, E. M. K.; Kerschen, A.; Vierling, E.; Queitsch, C.; Whitesell, L.; Gunatilaka, A. A. L. *Plant Physiol.* **2007**, *145*, 174.

(7) Workman, P.; Burrows, F.; Neckers, L.; Rosen, N. *Ann. N.Y. Acad. Sci.* **2007**, *1113*, 202.

(8) McDonald, E.; Workman, P.; Jones, K. *Curr. Top. Med. Chem.* **2006**, *6*, 1091.

(9) Winssinger, N.; Fontaine, J. G.; Barluenga, S. *Curr. Top. Med. Chem.* **2009**, *9*, 1419.

(10) Winssinger, N.; Barluenga, S. *Chem. Commun. (Camb).* **2007**, *2007*, 22.

(11) Cox, R. J. *Org. Biomol. Chem.* **2007**, *5*, 2010.

(12) Keatinge-Clay, A. T. *Nat. Prod. Rep.* **2012**, *29*, 1050.

(13) Xu, Y.; Espinosa-Artiles, P.; Schubert, V.; Xu, Y. M.; Zhang, W.; Lin, M.; Gunatilaka, A. A. L.; Süssmuth, R.; Molnár, I. *Appl. Environ. Microbiol.* **2013**, *79*, 2038.

(14) Foulke-Abel, J.; Townsend, C. A. *ChemBioChem* **2012**, *13*, 1880.

(15) 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.

(16) Thomas, R. *ChemBioChem* **2001**, *2*, 612.

(17) Xu, Y.; Zhou, T.; Zhou, Z.; Su, S.; Roberts, S. A.; Montfort, W. R.; Zeng, J.; Chen, M.; Zhang, W.; Zhan, J.; Molnár, I. *Proc. Natl. Acad. Sci. USA.* **2013**, *110*, 5398.

(18) Wang, S.; Xu, Y.; Maine, E. A.; Wijeratne, E. M. K.; Espinosa-Artiles, P.; Gunatilaka, A. A. L.; Molnár, I. *Chem. Biol.* **2008**, *15*, 1328.

(19) Zhou, H.; Qiao, K.; Gao, Z.; Vederas, J. C.; Tang, Y. *J. Biol. Chem.* **2010**, *285*, 41412.

(20) Li, Y.; Xu, W.; Tang, Y. *J. Biol. Chem.* **2010**, *285*, 22764.

(21) Wijeratne, E. M. K.; Paranagama, P. A.; Gunatilaka, A. A. L. *Tetrahedron* **2006**, *62*, 8439.

(22) Zhan, J. *Curr. Top. Med. Chem.* **2009**, *9*, 1598.

(23) Zhou, H.; Zhan, J.; Watanabe, K.; Xie, X.; Tang, Y. *Proc. Natl. Acad. Sci. USA.* **2008**, *105*, 6249.

(24) Reeves, C. D.; Hu, Z.; Reid, R.; Kealey, J. T. *Appl. Environ. Microbiol.* **2008**, *74*, 5121.

(25) 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.

(26) Gaffoor, I.; Brown, D. W.; Plattner, R.; Proctor, R. H.; Qi, W.; Trail, F. *Eukaryot. Cell* **2005**, *4*, 1926.

(27) Zhou, H.; Gao, Z.; Qiao, K.; Wang, J.; Vederas, J. C.; Tang, Y. *Nat. Chem. Biol.* **2012**, *8*, 331.

(28) Crawford, J. M.; Thomas, P. M.; Scheerer, J. R.; Vagstad, A. L.; Kelleher, N. L.; Townsend, C. A. *Science* **2008**, *320*, 243.

(29) Wang, M.; Zhou, H.; Wirz, M.; Tang, Y.; Boddy, C. N. *Biochem.* **2009**, *48*, 6288.
(30) 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.
(31) Du, L.; Lou, L. *Nat. Prod. Rep.* **2010**, *27*, 255.
(32) Vagstad, A. L.; Hill, E. A.; Labonte, J. W.; Townsend, C. A. *Chem. Biol.* **2012**, *19*, 1525.
(33) Fujii, I.; Watanabe, A.; Sankawa, U.; Ebizuka, Y. *Chem. Biol.* **2001**, *8*, 189.
(34) Gaffoor, I.; Trail, F. *Appl. Environ. Microbiol.* **2006**, *72*, 1793.
(35) Beck, Z. Q.; Aldrich, C. C.; Magarvey, N. A.; Georg, G. I.; Sherman, D. H. *Biochemistry* **2005**, *44*, 13457.
(36) Kohli, R. M.; Walsh, C. T.; Burkart, M. D. *Nature* **2002**, *418*, 658.
(37) Kopp, F.; Marahiel, M. A. *Curr. Opin. Biotechnol.* **2007**, *18*, 513.
(38) Vagstad, A. L.; Bumpus, S. B.; Belecki, K.; Kelleher, N. L.; Townsend, C. A. *J. Am. Chem. Soc.* **2012**, *134*, 6865.
(39) Liu, T.; Chiang, Y. M.; Somoza, A. D.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2011**, *133*, 13314.
(40) Crawford, J. M.; Vagstad, A. L.; Ehrlich, K. C.; Townsend, C. A. *Bioorg. Chem.* **2008**, *36*, 16.
(41) 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.
(42) Newman, A. G.; Vagstad, A. L.; Belecki, K.; Scheerer, J. R.; Townsend, C. A. *Chem. Commun.* **2012**, *48*, 11772.
(43) Vagstad, A. L.; Newman, A. G.; Storm, P. A.; Belecki, K.; Crawford, J. M.; Townsend, C. A. *Angew. Chem., Int. Ed.* **2013**, *52*, 1718.
(44) Yeh, H. H.; Chang, S. L.; Chiang, Y. M.; Bruno, K. S.; Oakley, B. R.; Wu, T. K.; Wang, C. C. *Org. Lett.* **2013**, *15*, 756.
(45) Hong, H.; Leadlay, P. F.; Staunton, J. *FEBS J.* **2009**, *276*, 7057.
(46) Lee, K. K.; Da Silva, N. A.; Kealey, J. T. *Anal. Biochem.* **2009**, *394*, 75.

SUPPLEMENTARY INFORMATION FOR

Thioesterase Domains of Fungal Nonreducing Polyketide Synthases Act as Decision Gates during Combinatorial Biosynthesis

Yuquan Xu[†,||], Tong Zhou[‡,||], Shuwei Zhang[‡,#], Li-Jiang Xuan[#], Jixun Zhan[‡,*], and István Molnár[†,§,*]

[†] Natural Products Center, School of Natural Resources and the Environment, The University of Arizona, 250 E. Valencia Rd., Tucson, AZ 85706, USA

[‡] Department of Biological Engineering, Utah State University, 4105 Old Main Hill, Logan, UT 84322, USA

[#] State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Shanghai Institute for Biological Sciences, Chinese Academy of Sciences, 501 Haike Road, Zhangjiang Hi-Tech Park, Shanghai 201203, China

[§] Bio5 Institute, The University of Arizona, 1657 E. Helen St., Tucson, AZ 85721, USA

[||] These authors contributed equally to this work.

* Corresponding authors: imolnar@email.arizona.edu (I. Molnár); jixun.zhan@usu.edu (J. Zhan)

Table of Contents

SI Materials and Methods

1. Molecular cloning

The construction of the YEpCcRADS1, YEpCcRADS2, YEpAtCURS1, YEpAtCURS2, YEpAtCURS2-PT$_{CcRADS2}$, YEpAtCURS2(F^{1459}Y,Y^{1576}F,W^{1584}L), and the YEpCcRADS2-PT$_{AtCURS2}$ expression vectors has been described earlier.[1,2] The Udwary-Merski algorithm was used to predict domain boundaries in nrPKSs.[3] The primers used in this study are listed in Table S1.

The plasmid YEpYX5 was constructed to express a CcRADS2 derivative where the TE domain (defined as P^{1768}-G^{2138}) was deleted. A 1,823-bp gene fragment encoding the *C*-terminal end of the truncated CcRADS2 was amplified with primers Cc_PT_Del_Up_F (Bsu36I) and Cc_TE_Del_R (EcoRV), and cloned into pJET1.2. After sequence verification, the *Bsu*36I-*Eco*RV fragment was inserted into the *Bsu*36I-*Pme*I cloning sites of YEpCcRADS2[1] to yield plasmid YEpYX5.

The plasmid YEpYX6 was constructed to express a CcRADS2 derivative where the TE domain was replaced with TE$_{AtCURS2}$. A 1,831-bp upstream fragment encoding L^{1164}-T^{1767} of CcRADS2 was amplified with primers Cc_ PT_Del_Up_F (Bsu36I) and At_TE_To_Cc_Up_R. A 1,073-bp fragment encoding S^{1730}-T^{2083} of AtCURS2 was amplified with primers At_TE_To_Cc_Dn_F and At_TE_To_Cc_Dn_R (EcoRV). The two PCR fragments were then fused during PCR using primers Cc_ PT_Del_Up_F (Bsu36I) and At_TE_To_Cc_Dn_R (EcoRV), and cloned into pJET1.2. After sequence verification, the *Bsu*36I -*Eco*RV fragment was inserted into the *Bsu*36I-*Pme*I cloning sites of YEpCcRADS2[1] to yield plasmid YEpYX6.

The plasmid YEpYX7 was constructed to express an AtCURS2 derivative where the TE domain (defined as S^{1730}-T^{2083}) was deleted. A 1,841-bp gene fragment encoding the *C*-terminal end of the truncated AtCURS2 was amplified with primers At_PT_Del_Up_F (AgeI) and At_TE_Del_R (EcoRV) and cloned into pJET1.2. After sequence verification, the *Age*I-*Eco*RV fragment was inserted into the *Age*I-*Pme*I cloning sites of YEpAtCURS2[2] to yield plasmid YEpYX7.

The plasmid YEpYX8 was constructed to express an AtCURS2 derivative where the TE domain was replaced with $TE_{CcRADS2}$. A 1,850-bp upstream fragment encoding P^{1120}-S^{1729} of AtCURS2 was amplified with primers At_PT_Del_Up_F (AgeI) and Cc_TE_To_At_Up_R. A 1,127-bp fragment encoding P^{1768}-G^{2135} of CcRADS2 was amplified with primers Cc_TE_To_At_Dn_F and Cc_TE_To_At_Dn_R (EcoRV). The two PCR fragments were then fused during PCR using primers At_PT_Del_Up_F (AgeI) and Cc_TE_To_At_Dn_R (EcoRV), and cloned into pJET1.2. After sequence verification, the *Age*I-*Eco*RV fragment was inserted into the *Age*I-*Pme*I cloning sites of YEpATCURS2[2] to yield plasmid YEpYX8.

The plasmid YEpYX12 was constructed to express a CcRADS2 derivative where the SAT domain was replaced with $SAT_{CcRADS2}$. A 1,110-bp gene fragment encoding M^{1}-T^{362} of AtCURS2 was amplified with primers At_SAT_to_Cc_F (NdeI) and At_SAT_to_Cc_R. A 1,516-bp downstream fragment encoding V^{366}-A^{870} of CcRADS2 was amplified with primers Cc_SAT_Dn_F and Cc_SAT_Dn_R (AgeI). The two PCR fragments were then fused during PCR using primers At_SAT_to_Cc_F (NdeI) and Cc_SAT_Dn_R (AgeI), and cloned into pJET1.2. After sequence verification, the *Nde*I-*Age*I fragment was inserted into the same cloning sites of YEpCcRADS2[1] to yield plasmid YEpYX12.

The plasmid YEpYX24 was constructed to express an AtCURS2 derivative where the SAT domain was replaced with $SAT_{CcRADS2}$. A 1,125-bp fragment encoding M^{1}-S^{365} of CcRADS2 was amplified with primers Cc_SAT_to_At_F (NdeI) and Cc_SAT_to_At_R. A 1,058-bp downstream fragment encoding P^{363}-L^{715} of AtCURS2 was amplified with primers At_SAT_Dn_F and At_SAT_Dn_R (Bsu36I). The two PCR fragments were then fused during PCR using primers Cc_SAT_to_At_F (NdeI) and At_SAT_Dn_R (Bsu36I) and cloned into pJET1.2. After sequence verification, the *Nde*I-*Bsu*36I fragment was inserted into the same cloning sites of YEpAtCURS2[2] to yield plasmid YEpYX24.

The plasmid YEpYX32 was constructed to express a CcRADS2 derivative in which the SAT and the TE domains were replaced by $SAT_{AtCURS2}$ and $TE_{AtCURS2}$, respectively. A 2,592-bp *Nde*I-*Age*I fragment of plasmid YEpYX12 was inserted the same cloning sites of YEpYX6 to yield plasmid YEpYX32.

The plasmid YEpYX33 was constructed to express an AtCURS2 derivative in which the SAT and TE domains were replaced with $SAT_{CcRADS2}$ and $TE_{CcRADS2}$, respectively. A 2,019-bp *Nde*I-*Bsu*36I fragment of plasmid YEpYX24 was inserted the same cloning sites of YEpYX8 to yield plasmid YEpYX33.

The plasmid YEpYX49 was constructed to express an AtCURS2 derivative in which the PT and the TE domains were replaced with $PT_{CcRADS2}$ and $TE_{CcRADS2}$, respectively. A 1,542-bp fragment was amplified from plasmid YEpAtCURS2-$PT_{CcRADS2}$[1] with primers Cc_PT_TE_to_At_F (AgeI) and Cc_PT_to_At_R2. A 1,517-bp fragment was amplified from plasmid YEpYX8 with primers At_PT_Del_Dn_F and Cc_PT_TE_to_At_R (SwaI). These two PCR fragments were then fused during PCR using primers Cc_PT_TE_to_At_F (AgeI) and Cc_PT_TE_to_At_R (SwaI), and

cloned into pJET1.2. After sequence verification, the *Age*I-*Swa*I fragment was inserted into the *Age*I-*Pme*I cloning sites of YEpAtCURS2[2] to yield plasmid YEpYX49.

The plasmid YEpYX50 was constructed to express an AtCURS2 derivative in which the ACP-TE didomain was replaced with ACP-$TE_{CcRADS2}$. A 1,455-bp upstream fragment encoding I^{1129}-V^{1613} of AtCURS2 was amplified with primers Cc ACP-TE_UP_F (AgeI) and Cc ACP-TE_UP_R. A 1,578-bp fragment encoding P^{1641}-D^{2138} of CcRADS2 was amplified with primers Cc ACP-TE F and At ACP-TE_R1 (SwaI). These two PCR fragments were then fused during PCR using primers Cc ACP-TE_UP_F (AgeI) and At ACP-TE_R2 (SwaI), and cloned into pJET1.2. After sequence verification, the *Age*I-*Swa*I fragment was inserted into the *Age*I-*Pme*I cloning sites of YEpAtCURS2[2] to yield plasmid YEpYX50.

The plasmid YEpYX54 was constructed to express a CcRADS2 derivative in which the SAT domain and the ACP-TE didomain were replaced with $SAT_{AtCURS2}$ and ACP-$TE_{AtCURS2}$, respectively. A 1,069-bp upstream fragment encoding M^{1}-G^{347} of AtCURS2 was amplified with primers Atnr-SAT-JET_F (NdeI) and Atnr-SAT_R. A 3,869-bp fragment encoding R^{351}-P^{1633} of CcRADS2 was amplified with primers Ccnr-KSATPT_F and Ccnr-KSATPT_R. A 1,519-bp downstream fragment encoding K^{1609}-T^{2083} of AtCURS2 was amplified with primers Atnr-rest_F and At ACP-TE_R (SwaI). These three PCR fragments were then fused during PCR using primers Atnr-SAT-JET_F (NdeI) and At ACP-TE_R (SwaI), and cloned into pJET1.2. After sequence verification, the *Nde*I-*Swa*I fragment was inserted into the *Nde*I-*Pme*I cloning sites of YEpATCURS2[2] to yield plasmid YEpYX54.

The plasmid YEpYX62 was constructed to express a CcRADS2 derivative in which the SAT and PT domains were replaced with $SAT_{AtCURS2}$ and $PT_{AtCURS2}$, respectively. A 3,513-bp *Nde*I-

*Bsu*36I fragment of plasmid YEpYX12 was inserted into the same cloning sites of YEpCcRADS2-$PT_{AtCURS2}$[1] to yield plasmid YEpYX62.

The plasmid YEpYX65 was constructed to express an AtCURS2 derivative in which the SAT and the PT domains were replaced with $SAT_{CcRADS2}$ and $PT_{CcRADS2}$, respectively. A 2,019-bp *Nde*I-*Bsu*36I fragment of plasmid YEpYX24 was inserted into the same cloning sites of plasmid YEpAtCURS2-$PT_{CcRADS2}$[1] to yield plasmid YEpYX65.

The plasmid YEpYX67 was constructed to express a CcRADS2 derivative in which the PT and the TE domains were replaced with $PT_{AtCURS2}$ and $TE_{AtCURS2}$, respectively. A 1,405-bp fragment was amplified from plasmid YEpCcRADS2-$PT_{AtCURS2}$[1] with primers At_PT_TE_to_Cc_F (Bsu36I) and At_PT_to_Cc R2. A 1,505-bp fragment was amplified from plasmid YEpYX6 with primers Cc_PT_Del_Dn_F2 and At_PT_TE_to_Cc_R (SwaI). These two PCR fragments were then fused during PCR using primers At_PT_TE_to_Cc_F (Bsu36I) and At_PT_TE_to_Cc_R (SwaI), and cloned into pJET1.2. After sequence verification, the *Bsu*36I-*Swa*I fragment was inserted into the *Bsu*36I-*Pme*I cloning sites of YEpCcRADS2[1] to yield plasmid YEpYX67.

2. Chemical characterization of polyketide products

2.1 General methods

Metabolites were routinely analyzed with a reversed-phase C18 column (Kromasil 5 µm, 4.6 mm × 250 mm) on an Agilent 1050 HPLC instrument, and the compounds were isolated with an Eclipse XDB-C18 column (5 µm, 4.6 mm × 150 mm) on an Agilent 1200 HPLC instrument. Optical rotations were measured on a Rudolph Autopol IV polarimeter using a 10 cm microcell. CD spectra were acquired with a JASCO J-810 instrument using a path length of 1 cm. ^{1}H NMR, ^{13}C NMR, and 2D NMR (^{1}H-^{1}H COSY, HSQC, HMBC, ROESY) spectra were obtained on a JEOL ECX-300 spectrometer. Chemical shifts (δ) are reported in ppm and were referenced to the residual solvent peaks of CD_3OD or C_5D_5N. Low resolution mass measurements were done on an Agilent 6130 Single Quad LC-MS. Low resolution tandem mass spectrometry was performed by electrospray ionization (ESI) on a Thermoelectron LCQ instrument. Accurate mass measurements were carried out by matrix assisted laser desorption/ionization (MALDI) on a Bruker Ultraflex III MALDI TOF-TOF instrument, or by electrospray ionization (ESI) on a Bruker 9.4 T Fourier transform ion-cyclotron resonance (FT-ICR) instrument. The samples were dissolved in ACN:H2O 1:1 (containing 0.1% of formic acid) at a of ca. 1 micromolar concentration. ESI was used to ionize the samples in both the positive (protonation) and negative (deprotonation) ion modes with standard operating conditions. Well established standard internal calibrants have been used to internally recalibrate the spectra.

2.2 Small scale fermentation and product analysis

Typically, a yeast strain was cultured in 50 ml of SC medium (0.67% yeast nitrogen base, 2% glucose, 0.72 g/l Trp/Ura DO supplement) at 30°C with shaking at 250 rpm. When the OD_{600} reached 0.6, an equal volume of YP medium (1% yeast extract, 2% peptone) was added to the culture, and the fermentation was continued under the same conditions for an additional 2 days. The culture was then adjusted to pH 5.0, and extracted with ethyl acetate three times. The extract was evaporated to dryness and analyzed by reverse phase HPLC (Kromasil C18 column, 5 µm, 4.6 mm × 250 mm; eluted with 5% acetonitrile in water for 5 min followed by a linear gradient of 5-95% of acetonitrile-water over 10 min, and 95% acetonitrile-water for 10 min at a flow rate of 0.8 ml/min; detection at 300 nm).

2.3 Large scale fermentation and compounds isolation

For a typical large scale culture, twenty 250-ml flasks, each containing 50 ml of SC medium, were inoculated with a yeast strain and cultured at 30°C with shaking at 300 rpm. After the OD_{600} reached 1.0, 50 ml of YP medium was added to each flask. The culture was harvested 3 days later by centrifugation at 3,000 rpm for 5 min to remove yeast cells. The supernatant was concentrated to 200 ml. The solution was adjusted to pH 5.0, and extracted with 600 ml of ethyl acetate three times. After evaporation of the organic solvent, the sample was subjected to a Diaion HP-20 column (4 × 30 cm) and successively eluted with water, 25% aqueous methanol, and acetone. The acetone fraction was collected and further separated by HPLC.

Specifically, **4** (2.7 mg) and **5** (5.5 mg) were isolated from *S. cerevisiae* BJ5464-NpgA [YEpATCURS1, YEpYX8] through further HPLC isolation on a C18 column (40% acetonitrile-water, 1 ml/min). **7** (2.3 mg) was derived from *S. cerevisiae* BJ5464-NpgA [YEpCcRADS1,

YEpYX6] (45% acetonitrile-water, 1 ml/min). Production of **7** in *S. cerevisiae* BJ5464-NpgA [YEpCcRADS1, YEpYX65] was confirmed by LC-DAD-ESI-MS analysis. **9** (12.8 mg) was purified from *S. cerevisiae* BJ5464-NpgA [YEpCcRADS1, YEpYX24] (45% acetonitrile-water, 1 ml/min). **11** (1.2 mg) was isolated from *S. cerevisiae* BJ5464-NpgA [YEpAtCURS1, YEpYX49] (37% acetonitrile-water, 1 ml/min). **12** (1.9 mg) was also acquired from *S. cerevisiae* BJ5464-NpgA [YEpAtCURS1, YEpYX49] (46% acetonitrile-water, 1 ml/min). **2** (6.8 mg) and **18** (3.2 mg) were isolated from *S. cerevisiae* BJ5464-NpgA [YEpCcRADS1, YEpYX49] (51-56% acetonitrile-water over 15 min, 1ml/min). Trace amounts of **13** and **14** were collected from the extract of *S. cerevisiae* BJ5464-NpgA [YEpAtCURS1, YEpYX49] using a linear gradient of 5% to 95% acetonitrile-water over 40 min on a C18 column (5 μm, 4.6 mm × 250 mm).

Compound **2** (Figure S3): white amorphous solid; UV: see Figure S5; $[\alpha]_D$ = +31.4° (*c* = 0.40, MeOH); ^{1}H NMR (300 MHz, CD_3OD): δ 6.73 (m, 1H), 6.27 (d, *J* = 2.4 Hz, 1H), 6.21 (d, *J* = 2.4 Hz, 1H), 5.85 (d, *J* = 15.5 Hz, 1H), 5.40-5.15 (m, 3H), 3.99 (d, *J* = 16.8 Hz, 1H), 3.90 (d, *J* = 16.8 Hz, 1H), 2.62 (m, 1H), 2.45-2.02 (m, 5H), 1.30 (d, *J* = 6.1 Hz, 3H); ^{13}C NMR (75 MHz, CD_3OD): δ 196.7, 171.5, 166.1, 162.5.2, 148.0, 139.3, 133.7, 131.6, 127.6, 112.9, 107.8, 102.9, 71.3, 48.5, 39.0, 30.2, 29.8, 19.0; ESI-MS (*m/z*): 315.1 $[M-H]^-$, 317.0 $[M+H]^+$, 339.3 $[M+Na]^+$.

Compound **4** (Figure S2): white amorphous solid; UV: see Figure S5; ^{1}H and ^{13}C NMR (CD_3OD): see Table S2; ESI-MS (*m/z*): 319.2 $[M-H]^-$, 321.2 $[M+H]^+$.

Compound **5** (Figure S2): white amorphous solid; UV: see Figure S5; ^{1}H and ^{13}C NMR (CD_3OD): see Table S2; ESI-MS (*m/z*): 321.2 $[M-H]^-$, 323.2 $[M+H]^+$.

Compound **7** (Figure S3): white amorphous solid; $[\alpha]_D$ = +27.8° (*c* = 0.12, MeOH); ^{1}H NMR (300 MHz, C_5D_5N): δ 6.83 (d, *J* = 2.4 Hz, 1H), 6.70 (d, *J* = 2.4 Hz, 1H), 6.57 (dt, *J* = 15.9, 7.1 Hz, 1H), 6.39 (s, 1H), 6.11 (d, *J* = 15.9 Hz, 1H), 5.75 (dt, *J* = 15.5, 7.3 Hz, 1H), 5.57 (dt, *J* = 15.5, 7.2 Hz, 1H), 4.08 (m, 1H), 2.51 (m, 1H), 2.22 (m, 1H), 1.86 (m, 1H), 1.37 (d, *J* = 6.0 Hz, 3H); ^{13}C NMR (75 MHz, CD_3OD): δ 165.7, 163.2, 161.8, 152.2, 139.3, 135.4, 131.9, 128.6, 123.6, 107.2, 104.6, 102.7, 97.5, 66.7, 43.0, 32.7, 29.3, 23.5; ESI-MS (*m/z*): 315.1 $[M-H]^-$, 317.1 $[M+H]^+$, 339.1 $[M+Na]^+$.

Compound **9** (Figure S2): yellow amorphous solid; $[\alpha]_D$ = +89.2° (*c* = 0.64, $CHCl_3$); UV: see Figure S5; ^{1}H and ^{13}C NMR (CD_3OD): see Table S2; ^{1}H NMR (300 MHz, C_5D_5N): δ 7.00 (d, *J* = 1.8 Hz, 1H), 6.88 (d, *J* = 1.8 Hz, 1H), 6.84 (dt, *J* = 15.9, 7.8 Hz, 1H), 6.71 (d, *J* = 15.9 Hz, 1H), 5.29 (ddd, *J* = 16.1, 8.9, 3.1 Hz, 1H), 5.22 (m, 1H), 5.18 (ddd, *J* = 16.1, 6.8, 2.7 Hz, 1H), 4.07 (d, *J* = 17.2 Hz, 1H), 3.86 (d, *J* = 17.2 Hz, 1H), 2.25 (m, 1H), 2.21 (m, 1H), 1.12 (d, *J* = 6.2 Hz, 1H); CD $\delta\varepsilon_{221}$ +2.25, $\delta\varepsilon_{239}$ -0.53, $\delta\varepsilon_{273}$ -0.32, $\delta\varepsilon_{301}$ +0.39, $\delta\varepsilon_{338}$ -0.15 (4.1 mM in MeOH); ESI-MS (*m/z*): 315.1 $[M-H]^-$, 317.2 $[M+H]^+$, 339.2 $[M+Na]^+$.

Compound **11** (Figure S2): yellow amorphous solid; UV: see Figure S5; HR-ESI-MS: *m/z* 359.1484 $[M+Na]^+$ (calcd. for $C_{18}H_{24}O_6Na^+$: 359.1465); LR-ESI-MS-MS: see Figure S7 and Table S4.

Compound **12** (Figure S2): white amorphous solid; $[\alpha]_D$ = +61.8° (*c* = 0.06, MeOH); UV: see Figure S5; ^{1}H and ^{13}C NMR (CD_3OD): see Table S3; ESI-MS (*m/z*): 335.0 $[M-H]^-$, 337.0 $[M+H]^+$, 359.0 $[M+Na]^+$.

Compound **13** (Figure S2): white amorphous solid; UV: see Figure S5; HR-ESI-MS: *m/z* 309.13377 $[M+H]^+$ (calcd. for $C_{16}H_{21}O_6^+$: 309.13381), 331.11532 $[M+Na]^+$ (calcd. for $C_{16}H_{20}O_6Na^+$: 331.11521), 307.11885 $[M-H]^-$ (calcd. for $C_{16}H_{19}O_6^-$: 307.11871), and *m/z* 263.12891 $[M-COOH]^-$ (calcd. for $C_{15}H_{19}O_4^-$: 263.12888); HR-ESI-MS-MS: see Figure S8 and Table S4.

Compound **14** (Figure S2): white amorphous solid; UV: see Figure S5; HR-ESI-MS: *m/z* 291.12322 $[M+H]^+$ (calcd. for $C_{16}H_{19}O_5^+$: 291.12270); HR-ESI-MS-MS: see Figure S9 and Table S4.

Compound **18** (Figure S2): white amorphous solid; $[\alpha]_D$ = -14.3° (*c* = 0.28, MeOH); UV: see Figure S5; ^1H and ^{13}C NMR (CD_3OD): see Table S3; ESI-MS (*m/z*): 315.1 $[M-H]^-$, 317.0 $[M+H]^+$.

SI Results

Structure elucidation

4 was isolated as a white amorphous solid. Its ESI-MS spectra showed $[M-H]^-$ (negative mode) and $[M+H]^+$ (positive mode) peaks at *m/z* 319.2 and 321.2, respectively, indicating a molecular weight of 320 and a potential molecular formula of $C_{18}H_{24}O_5$. The benzene ring was suggested to be C-3, C-8 *ortho-C*-substituted and C-5, C-7 *meta-O*-substituted on the basis of the coupling pattern of H-4 (δ 6.22 (d, 1H, 2.1)) and H-6 (δ 6.24 (d, 1H, 2.1)) as well the corresponding ^{13}C NMR signals (Table S2). The typical singlet signal of H-2 (δ 3.58 (s, 2H)) and its HMBC correlations with C-1, C-3, C-4, and C-8 (Figure S4) established an acetoxy side chain located at C-3 of the benzene ring. A spin system of CH_2-CH_2-CH_2-CH-CH-CH_2-CH_2-CH_3 observed in the ^1H-^1H COSY spectrum and the HMBC correlations among these groups established a Δ^{14} unsaturated aliphatic chain substituted at C-8 of the aromatic ring. The geometry of the double bond was determined as *Z*-configuration according to the coupling constant $J_{14, 15}$ = 6.5 Hz.

Therefore, **4** was identified as (*Z*)-2-(2-(dec-6-enoyl)-3,5-dihydroxyphenyl)acetic acid (**4** in Figure S2).

Spectroscopic analysis of **5** revealed that this compound is similar to **4**. Ion peaks of $[M-H]^-$ at *m/z* 321.2 and $[M+H]^+$ at *m/z* 323.2 in the ESI-MS spectra showed that **5** is 2 mass units larger than **4**. A comparison of the ^1H and ^{13}C NMR data of **4** and **5** suggested that the olefinic protons at δ 5.01 (overlap, 1H) and δ 5.79 (td, 1H, *J* = 10.3 and 6.5 Hz) in **4** was hydrogenated into two CH_2 groups at δ 1.29 (m), as confirmed by ^1H-^1H COSY and HMBC correlations (Figure S4). Therefore, the structure of **5** was identified as 2-(2-decanoyl-3,5-dihydroxyphenyl)acetic acid (Figure S2).

9 was obtained as a yellow amorphous powder. ESI-MS analysis in both negative and positive modes confirmed that it has a molecular weight of 316. ^1H and ^{13}C NMR suggested that **9** contains a similar benzene ring as in **4** and **5**. The HMBC correlations of H-2a and H-2b to C-1 and C-3 indicated an acetoxy substitution at C-3. Moreover, the doublets of H-2a and H-2b resulted from the restrained rotation of the single bond between C-2 and C-3 by the macrolide ring, which was further proved by the HMBC correlation of H-17 to C-1. Both of the olefinic bonds in **9** were deduced as *E* configuration based on their coupling constants of $J_{10, 11}$ = 15.9 Hz and $J_{14, 15}$ = 16.1 Hz [δ 6.84 (dt, *J* = 15.9, 7.8 Hz, 1H), δ 6.71 (d, *J* = 15.9 Hz, 1H), δ 5.29 (ddd, *J* = 16.1, 8.9, 3.1 Hz, 1H), δ 5.18 (ddd, *J* = 16.1, 6.8, 2.7 Hz, 1H) in C_5D_5N]. ^1H-^1H COSY and HMBC correlations confirmed that the position of the $\Delta^{10, 14}$ unsaturated aliphatic chain is attached to the benzene ring at C-8. The absolute configuration at C-17 was subsequently elucidated as *R* by comparison of its CD spectrum with those of the structurally related compounds,[4,5] based on the comparable Cotton effect of **9** with a maximum at 221 nm (Figure S6). Thus, **9** (Figure S2)

was identified as (*R*,6*E*,10*E*)-13,15-dihydroxy-4-methyl-4,5,8,9-tetrahydro-1*H*-benzo[*d*][1]oxacyclotetradecine-2,12-dione and was given, according to its combinatorial origin and chemical structure, the name radilarin.

12 was isolated as a colorless solid. The ion peaks including $[M-H]^-$ at *m/z* 335.0, $[M+H]^+$ at *m/z* 337.0 and $[M+Na]^+$ at *m/z* 359.0 in the ESI-MS spectra suggested that the molecular weight of **12** is 336. ^1H, ^{13}C and HMBC NMR spectra indicated the presence of a benzene ring, with *C*-substitution at C-2 and C-7 as well as *O*-substitution at C-3 and C-5. A spin system of CH_2-CH-CH_2-CH_2-CH_2-CH-CH_3 was observed in the ^1H-^1H COSY spectrum. HMBC correlations of H-6 to C-8, H-8 to C-2, C-7 and C-9, as well as H-10 to C-8, C-9 and C-11 proved that the side chain is linked to the benzene ring at C-7 (Figure S4). The apparent ROESY correlation between H-11 and H-15 (Figure S4), together with the molecular weight, confirmed the formation of a tetrahydropyran ring on the side chain and the *cis* relative configuration of H-11 and H-15. Since the absolute configuration of C-15 can be deduced as *S* from the biosynthetic origin of 15-OH, which resulted from the reduction of a ketone group by the AtCURS1 KR domain,[6] C-11 was also determined as *S* configuration. **12** was thus established as ethyl 2,4-dihydroxy-6-(3-((2*S*,6*S*)-6-methyltetrahydro-2*H*-pyran-2-yl)-2-oxopropyl)benzoate (Figure S2).

11 was isolated as a white amorphous powder. The molecular formula was determined to be $C_{18}H_{24}O_6$ based on the $[M+Na]^+$ ion peak at *m/z* 359.1484 (calcd. for $C_{18}H_{24}O_6Na^+$: 359.1465) in the HR-ESI-MS spectrum. Analysis of the fragmentation pattern observed in the ESI-MS/MS spectrum (Table S4 and Figure S7) and UV profile (Figure S5) of **11** revealed that this structure is similar to **12** with the carboxyl at C-1 esterified by an ethyl group. Relevant fragment ions are

indicated in Figure S7 and presented in Table S4. The structure of **11** was deduced as (*S,E*)-ethyl 2,4-dihydroxy-6-(8-hydroxy-2-oxonon-3-en-1-yl)benzoate (Figure S2).

13 was collected in a trace amount for MS analysis. HR-ESI-MS showed ion peaks including $[M+H]^+$ at *m/z* 309.13377 (calcd. for $C_{16}H_{21}O_6^+$: 309.13381) and $[M+Na]^+$ at *m/z* 331.11532 (calcd. for $C_{16}H_{20}O_6Na^+$: 331.11521) in the positive mode, as well as $[M-H]^-$ at *m/z* 307.11885 (calcd. for $C_{16}H_{19}O_6^-$: 307.11871) and $[M-COOH]^-$ at *m/z* 263.12891 (calcd. for $C_{15}H_{19}O_4^-$: 263.12888) in the negative mode. Thus, the molecular formula of **13** was determined to be $C_{16}H_{20}O_6$. The UV profile of **13** (Figure S5) is similar to that of **12**, suggesting that they share a similar chemical structure. Analysis of the HR-ESI-MS/MS fragments (negative mode) (Table S4 and Figure S8) revealed that the structure is the acid form of **12**. The configurations of both C-11 and C-15 were deduced to be *S* according to its biosynthetic origin. Thus, the structure of **13** was proposed to be 2,4-dihydroxy-6-(3-((2*S*,6*S*)-6-methyltetrahydro-2*H*-pyran-2-yl)-2-oxopropyl)benzoic acid.

14 was obtained in a trace amount for MS analysis. HR-ESI-MS showed the $[M+H]^+$ion peak at *m/z* 291.12322 in the positive mode, indicating that **14** has a molecular formula of $C_{16}H_{18}O_5$, one H_2O unit less than **13**. UV profile of **14** (Figure S5) was similar to those of isocoumarins. Analysis of the HR-ESI-MS/MS fragments (positive mode) (Table S4 and Figure S9) revealed that this structure is the isocoumarin form of **13**, which was possibly formed through spontaneous keto-enol tautomerism and esterification by the removal of a water molecule from **13**. The configurations of C-11 and C-15 were also deduced to be *S* based on its biosynthetic origin. Thus, the structure **14** was proposed to be 6,8-dihydroxy-3-(((2*S*,6*S*)-6-methyltetrahydro-2*H*-pyran-2-yl)methyl)-1*H*-isochromen-1-one.

18 was isolated as a white amorphous solid, with a molecular weight of 316 that is the same as monocillin II (**2**). The UV absorption spectrum of **18** is similar to that of **2** (Figure S5). A comparison of the ^1H and ^{13}C NMR spectra of **18** and **2** showed that these two molecules are very similar, except that the coupling constant for 10-H is 11.3 Hz in **18**, compared to 15.5 Hz in **2**. This difference suggested that there is a *cis* double bond between C-10 and C-11 in **18**,

instead of the *trans* double bond in the same position of **2**. Further 2D NMR analysis (Figure. S4) confirmed that **18** is 10-*cis*-monocillin II.

1, **2**, **15**, **16**, and **17** (Figure S3) were isolated from *Saccharomyces cerevisiae* BJ5464-NpgA [YEpAtCURS1, YEpAtCURS2], *S. cerevisiae* BJ5464-NpgA [YEpCcRADS1, YEpCcRADS2], *S. cerevisiae* BJ5464-NpgA [YEpAtCURS1, YEpAtCURS2($F^{1455}Y$,$Y^{1576}F$,$W^{1584}L$)] and *S. cerevisiae* BJ5464-NpgA [YEpCcRADS1, YEpCcRADS2-$PT_{AtCURS2}$], respectively, as previously reported.[1,2] **7** (Figure S3) was isolated from *S. cerevisiae* BJ5464-NpgA [YEpCcRADS1, YEpYX6] and determined by a comparison of the NMR data to those reported previously.[7] Production of **7** in *S. cerevisiae* BJ5464-NpgA [YEpCcRADS1, YEpYX65] was confirmed by comparing LC-UV-MS with the standard obtained from *S. cerevisiae* BJ5464-NpgA [YEpCcRADS1, YEpYX6].

SI Tables

Table S1. PCR primers used in this study

Primer name	Primer sequence
Cc_PT_Del_Up_F (Bsu36I)	5'-CTGGCCGACCCAATCTGGCTCG-3'
Cc_TE_Del_R (EcoRV)	5'-ATGATATCTCAAGTCGCTGAACGACTGGTGC-3'
At_PT_Del_Up_F (AgeI)	5'-CCGGTGGCATCAACTCTGCTCGGAT-3'
At_TE_Del_R (EcoRV)	5'-ATGATATCTCAGCTGGAGGGGGTTGTTGAGC-3'
Cc_TE_To_At_Up_R	5'-TCCTTCATGGAGCTGGATGGGCTGGAGGGGGTTGTTGAG-3'
Cc_TE_To_At_Dn_F	5'-CCATCCAGCTCCATGAAGGAAACCG-3'
Cc_TE_To_At_Dn_R (EcoRV)	5'-ATGATATCTCAACCATCCATCAAACCGTCTC-3'
At_TE_To_Cc_Up_R	5’-ATGGGATGGAGGACTGGGAAGTCGCTGAACGACTGGTGC-3'
At_TE_To_Cc_Dn_F	5'-TCCCAGTCCTCCATCCCATCG-3'
At_TE_To_Cc_Dn_R (EcoRV)	5'-ATGATATCTCAGGTAGACTCCCCCTCAAAATATG-3'
At_SAT_to_Cc_F (NdeI)	5'-CTTCATATGGACTCCCAACCGTCCTGC-3'
At_SAT_to_Cc_R	5'-GCGATGATGATGTTGATACGGTTGGTGGCTTCGCAGTG-3'
Cc_SAT_Dn_F	5'-GTATCAACATCATCATCGCAAAGTGG-3'
Cc_SAT_Dn_R (AgeI)	5'-AGGCGAACCGGTGCGGATG-3'
Cc_SAT_to_At_F (NdeI)	5'-GCTTCATATGTCTAGGCTACCACAGAAG-3'
Cc_SAT_to_At_R	5'-ATCCCGTCTCGACTGTCTGGCGAGGCCTGAAGAAGGGAC-3'
At_SAT_Dn_F	5'-CCAGACAGTCGAGACGGGATCGC-3'
At_SAT_Dn_R (Bsu36I)	5'-AGATTGGCCTTGATAGCCCCGAC-3'
Cc_PT_TE_to_At_F (AgeI)	5'-CTCGAGTTTTTCAGCAAGATCCGGTGGCATCAACTCTGCTCGGAT-3'
Cc_PT_to_At_R2	5'-GACGGCCACAGCTTGCCGACG-3'
At_PT_Del_Dn_F	5'-CGTCGGCAAGCTGTGGCCGTC-3'
Cc_PT_TE_to_At_R (SwaI)	5'-AGGAGATCTTCTAGAAAGATATCGGTCCGCACATTTAAATCAGTG-3'
At_PT_TE_to_Cc_F (Bsu36I)	5'-CTCGAGTTTTTCAGCAAGATCTGGCCGACCCAATCTGGCTCG-3'
At_PT_to_Cc_R2	5'-TGGGCGCTGTGGGGACCTCGAC-3'
Cc_PT_Del_Dn_F2	5'- GTCGAGGTCCCCACAGCGCCC-3'
At_PT_TE_to_Cc_R (SwaI)	5'-AGGAGATCTTCTAGAAAGATATCGGTCCGCACATTTAAATCAGTG-3'
Cc ACP-TE_UP_F (AgeI)	5'-ATCGTCGGGGAACCGGGTGTCTTCGACC-3'
Cc ACP-TE_UP_R	5'-GACCTCGACGGCTTTCTTGGTCTCC-3'
Cc ACP-TE_F	5'-ACCAAGAAAGCCGTCGAGGTCCCCACAGCGCCCAAGAAATCTGC-3'
At ACP-TE_R1 (SwaI)	5'-ATCGTGAAGGCATCGGTCCGCAC-3'
Atnr-SAT-JET_F (NdeI)	5'-CTCGAGTTTTTCAGCAAGATAGCTTCATATGGACTCCAACCGTC-3'
Atnr-SAT_R	5'-TCCGGCGGTTTGCAGCACC-3'
Ccnr-KSATPT_F	5'- AGGTGCTGCAAACCGCCGGAAGACACGTATCAGTTCAAGATGC-3'
Ccnr-KSATPT_R	5'-AGGCGCCTTGAGAGGAGCCGC-3'
At ACP-TE_R (SwaI)	5'-ATCGTGAAGGCATCGGTCCGCAC-3'
Atnr-rest_F	5'-CGGCTCCTCTCAAGGCGCCTAAAGCCGTCGAGGTCAGTC-3'
At ACP-TE_R2 (SwaI)	5'-AGGAGATCTTCTAGAAAGATATCGGTCCGCACATTTAAATCAGTG-3'

Table S2. ^{1}H (300 MHz) and ^{13}C (75 MHz) NMR data for 4, 5 and 9

(CD_3OD, δ in ppm, J in Hz).

No.	4		5		9	
	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H
1	174.0		174.0		171.0	
2	39.0	3.58 (s, 2H)	39.2	3.58 (s, 2H)	38.7	3.50 (d, 1H, 17.5)
						3.71 (d, 1H, 17.5)
3	138.6		135.9		134.8	
4	111.7	6.22 (d, 1H, 2.1)	110.2	6.20 (d, 1H, 2.0)	110.5	6.27 (d, 1H, 1.8)
5	161.4		160.0		160.3	
6	102.7	6.24 (d, 1H, 2.1)	101.3	6.25 (d, 1H, 2.0)	101.2	6.24 (d, 1H, 1.8)
7	159.6		158.5		159.1	
8	121.2		120.0		118.0	
9	209.1		208.3		197.9	
10	45.2	2.92 (t, 2H, 7.9)	43.8	2.91 (t, 2H, 7.9)	132.1	6.41 (d, 1H, 15.4)
11	25.6	1.63 (m, 2H)	24.0	1.62 (m, 2H)	151.1	6.54 (dt, 1H, 15.4,
						4.8)
12	30.5	1.32 (m, 2H)	29.1	1.29 (m, 2H)	31.7	2.43 (m, 2H)
13	30.2	2.03 (m, 2H)	29.1	1.29 (m, 2H)	30.1	2.14 (m, 2H)
14	114.8	5.01 (overlap, 1H)	29.2	1.29 (m, 2H)	131.5	5.27 (m, 1H)
15	140.3	5.79 (td, 1H,	29.2	1.29 (m, 2H)	127.7	5.27 (m, 1H)
		10.3, 6.5)				
16	35.0	1.32 (m, 2H)	31.7	1.29 (m, 2H)	39.6	2.13 (m, 2H)
17	23.9	1.24 (m, 2H)	22.4	1.29 (m, 2H)	70.1	5.02 (qt, 1H, 6.2, 2.0)
18	14.5	0.93 (m, 3H)	13.1	0.91 (t, 3H, 6.9)	16.3	1.20 (d, 3H, 6.2)

Table S3. ^{1}H (300 MHz) and ^{13}C (75 MHz) NMR data for 12 and 18

(CD3OD, δ in ppm, J in Hz).

No.	12		18	
	δ_C	δ_H	δ_C	δ_H
1	169.6		172.5	
2	106.5		105.0	
3	161.8		167.3	
4	101.6	6.22 (s, 1H)	103.2	6.25 (d, 1H, 2.1)
5	162.4		164.4	
6	112.1	6.13 (s, 1H)	114.4	6.12 (d, 1H, 2.1)
7	138.6		141.0	
8	50.1	3.88 (s, 2H)	53.6	4.51 (d, 1H, 18.2)
				3.66 (d, 1H, 18.2)
9	205.1		200.9	
10	48.4	2.61 (dd, 1H, 16.2, 7.2)	128.6	6.33 (d, 1H, 11.3)
		2.47 (dd, 1H, 16.2, 5.1)		
11	73.1	3.68 (m, 1H)	148.6	6.06 (m, 1H)
12	30.7	1.60 (m, 1H)	28.2	3.43 (m, 1H)
		1.18 (m, 1H)		2.29 (m, 1H)
13	23.0	1.78 (m, 1H)	32.3	2.36 (m, 1H)
		1.59 (m, 1H)		2.02 (m, 1H)
14	32.6	1.55 (m, 1H)	132.6	5.31 (m, 1H)
		1.21 (m, 1H)		
15	73.1	3.41 (m, 1H)	126.0	5.31 (m, 1H)
16	22.0	1.04 (d, 3H, 6.2)	38.2	2.54 (m, 1H)
				2.21 (m, 1H)
17	60.8	4.20 (q, 1H, 7.2)	73.5	5.31 (m, 1H)
18	14.0	1.22 (t, 3H, 6.6)	18.5	1.33 (d, 3H, 6.5)

Table S4. MS/MS fragmentation of 11 (LR), 13 (HR) and 14 (HR).

11		**13**		**14**	
Ion peak (*m/z*)	Fragment	*m/z* measured (calculated)	fragment	*m/z* measured (calculated)	fragment
337	Parent ion $[M+H]^+$	307.11894 (307.11871)	Parent ion $[M-H]^-$	291.12420 (291.12270)	Parent ion $[M+H]^+$
319	Loss of H_2O	289.10851 (289.10760)	Loss of H_2O	273.11372 (237.11268)	Loss of H_2O
291	Loss of C_2H_5OH	263.12889 (263.12888)	Loss of COOH	255.10387 (255.10212)	Loss of two units of H_2O
273	Loss of C_2H_5OH and H_2O			193.05148 (193.05008)	Loss of $C_6H_{10}O$
223	Loss of $C_7H_{14}O$				

SI Figures

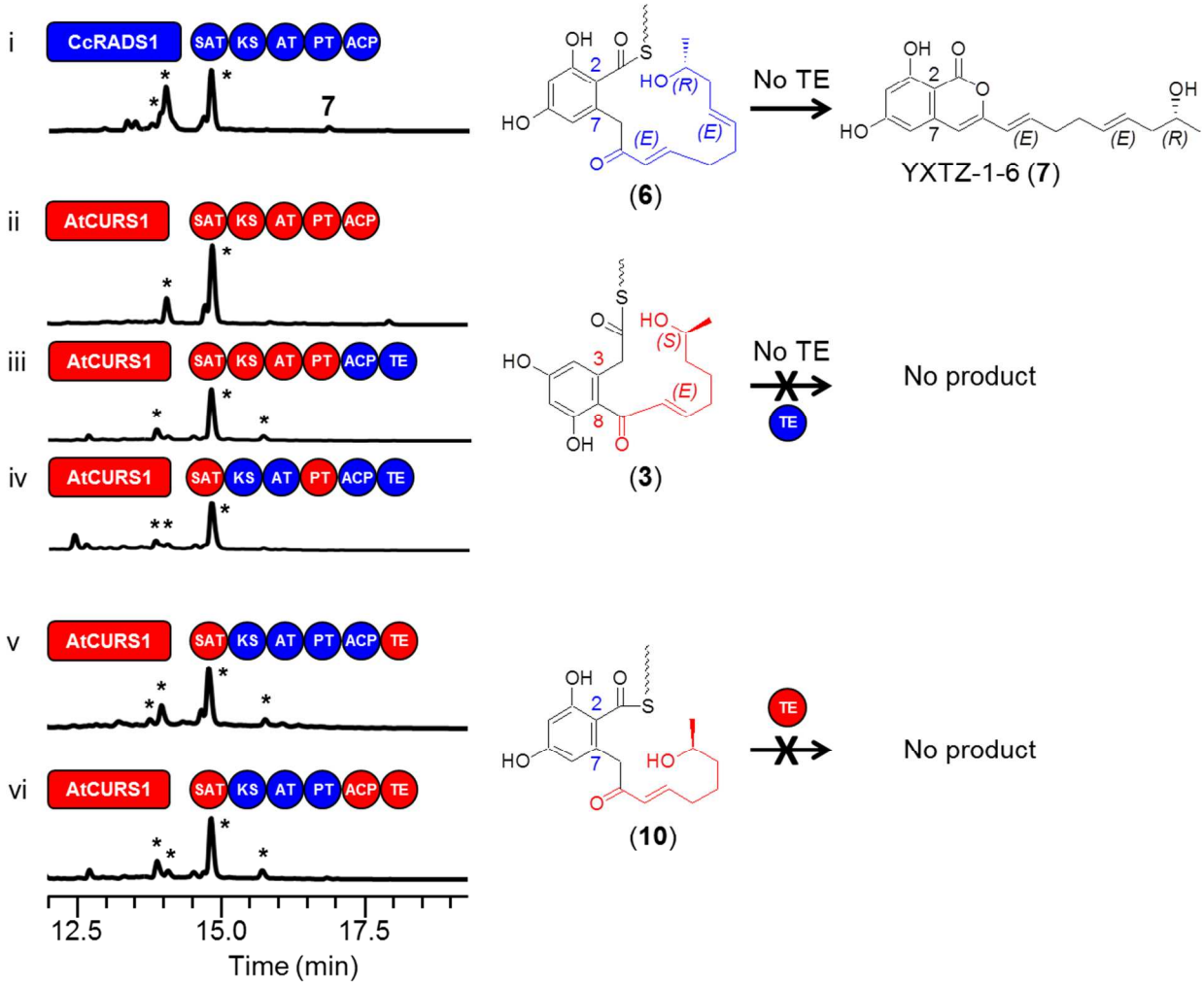


Figure S1. TE swaps during combinatorial biosynthesis

Product profiles (HPLC traces recorded at 300 nm) of *S. cerevisiae* BJ5464-NpgA[7,8] co-transformed with engineered or native hrPKS-nrPKS pairs as indicated: (i) YEpCcRADS1 and YEpXY5; (ii) YEpAtCURS1 and YEpXY7; (iii) YEpAtCURS1 and YEpXY50; (iv) YEpAtCURS1 and YEpXY62; (v) YEpAtCURS1 and YEpXY32; (vi) YEpAtCURS1 and YEpCcRADS54. The hrPKS-generated portions of the proposed ACP-bound thioester intermediates **6**, **3** and **10** are color-coordinated with the synthase. The structure of the

spontaneously released α-pyroneYXTZ-1-6 (**7**) is also shown. Domains drawn as red circles are derived from AtCURS2, while those in blue are from CcRADS2. *, Yeast metabolites unrelated to the iPKS products.

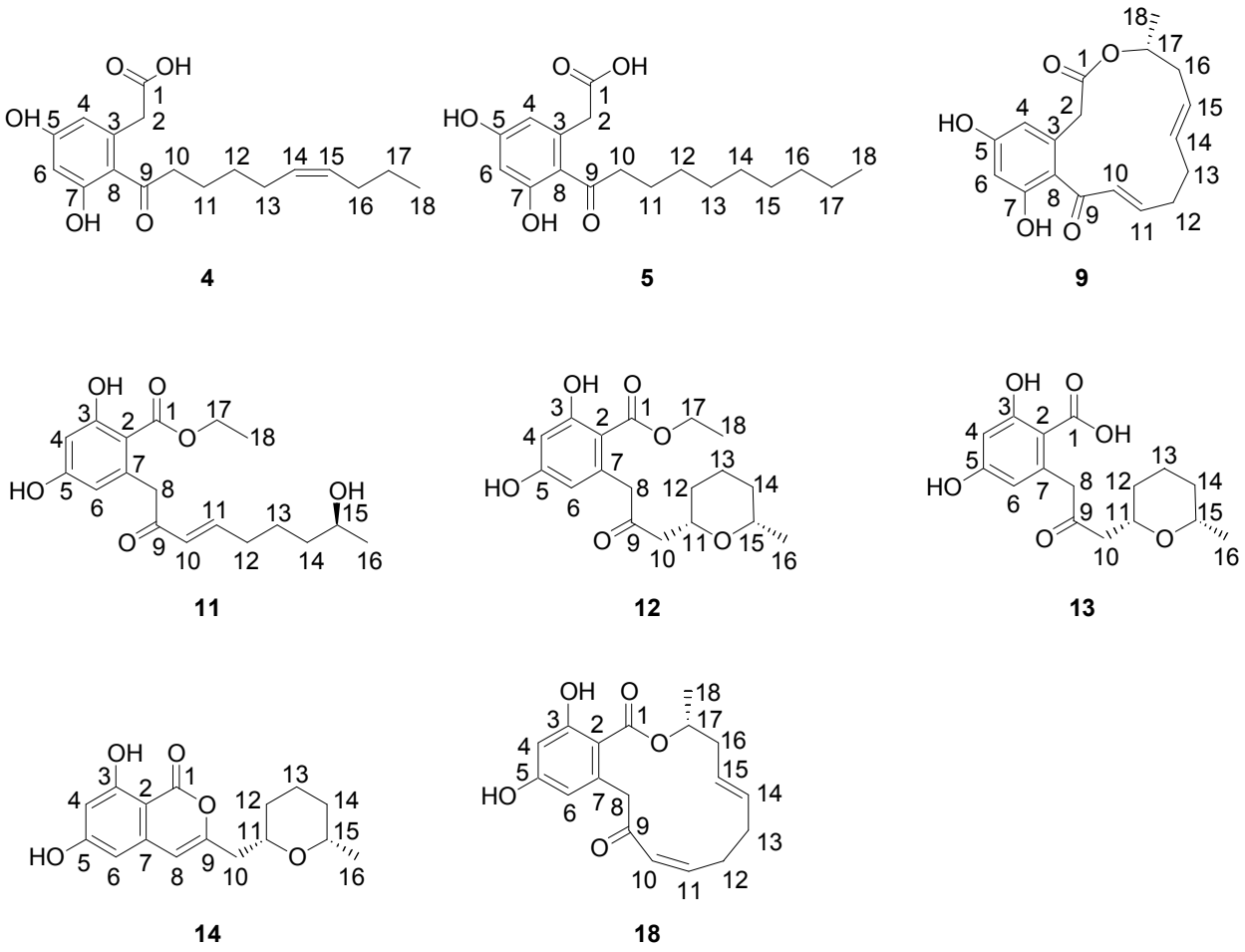


Figure S2. Chemical structures of 4, 5, 9, 11, 12, 13, 14 and 18.

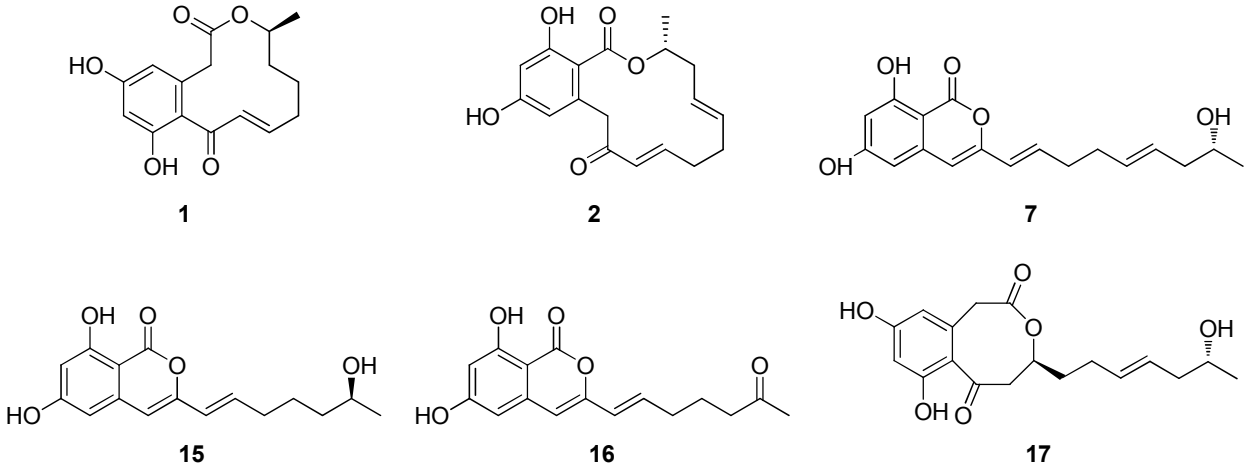


Figure S3. Chemical structures of 1, 2, 7, 15, 16 and 17.

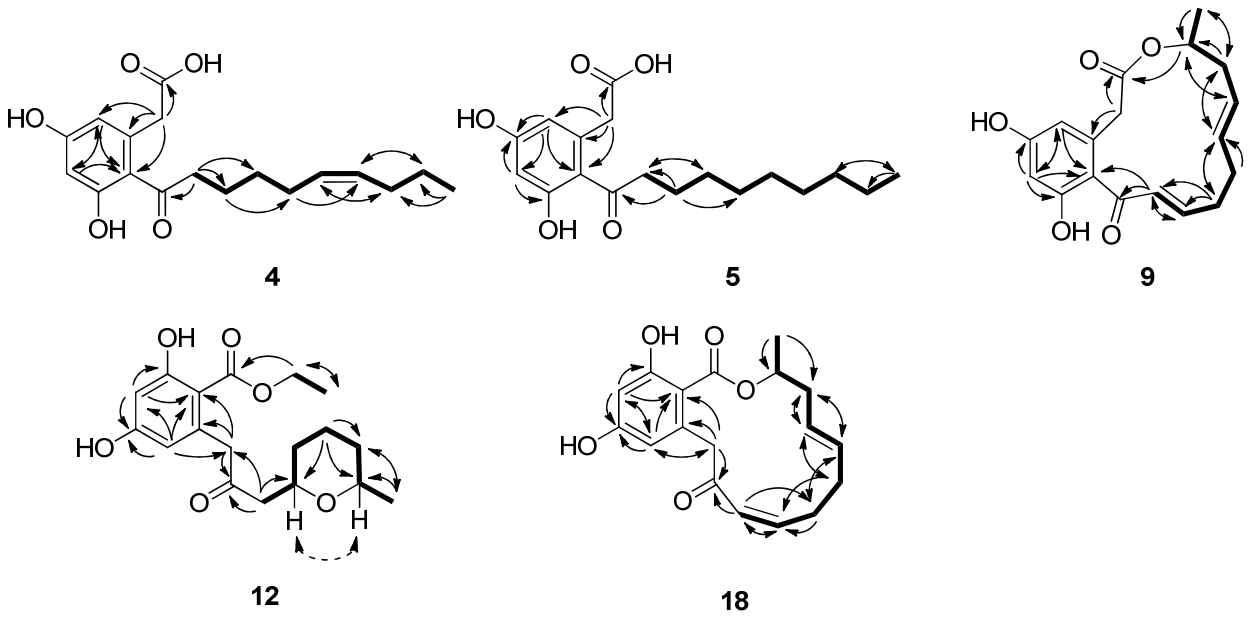


Figure S4. Key HMBC (⟶), 1H-1H-COSY (▬) and ROESY (⇢⇠) correlations for 4, 5, 9, 12 and 18.

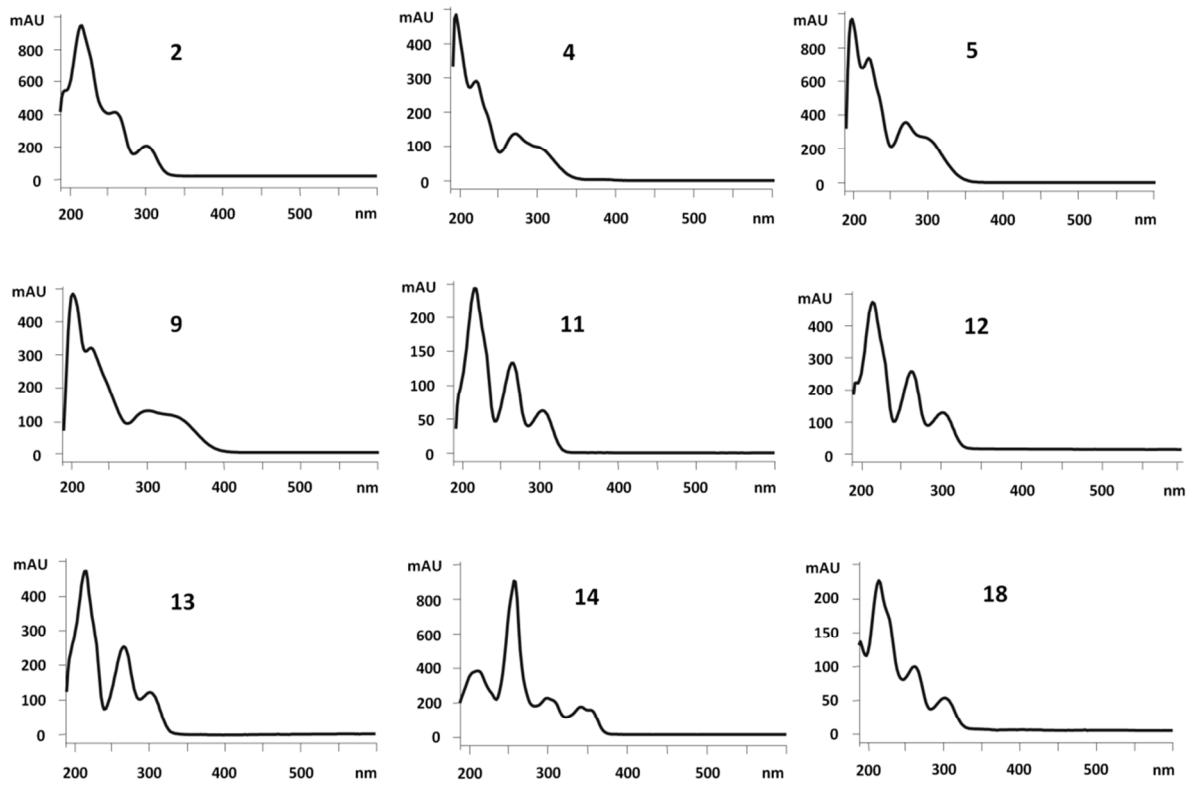


Figure S5. The UV absorption spectra of 2, 4, 5, 9, 11, 12, 13, 14 and 18.

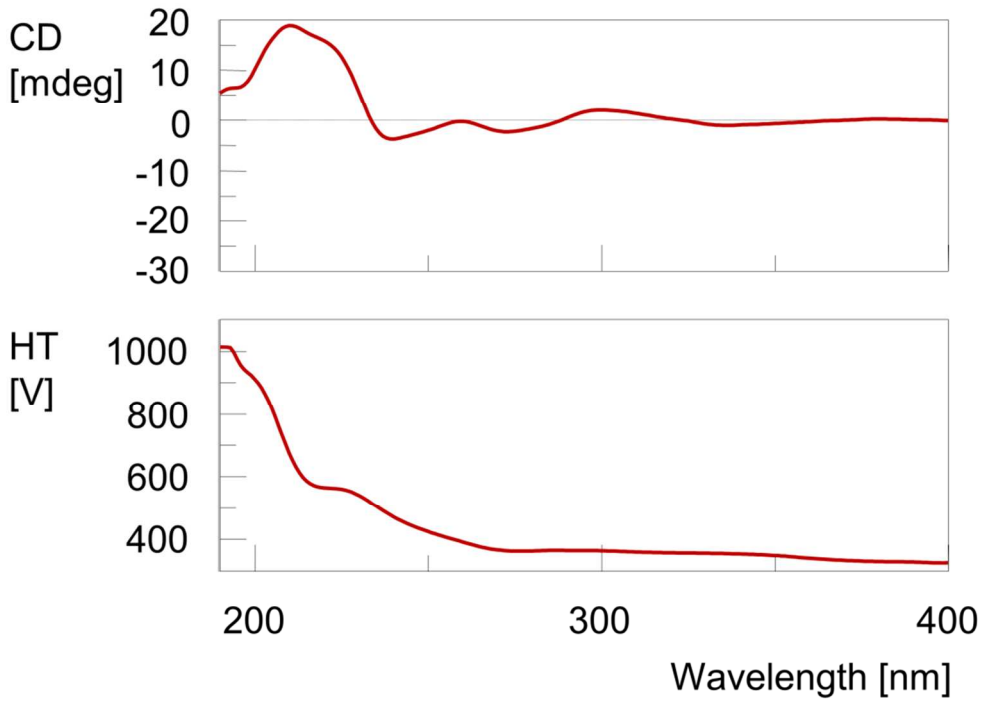


Figure S6. The CD spectra of 9.

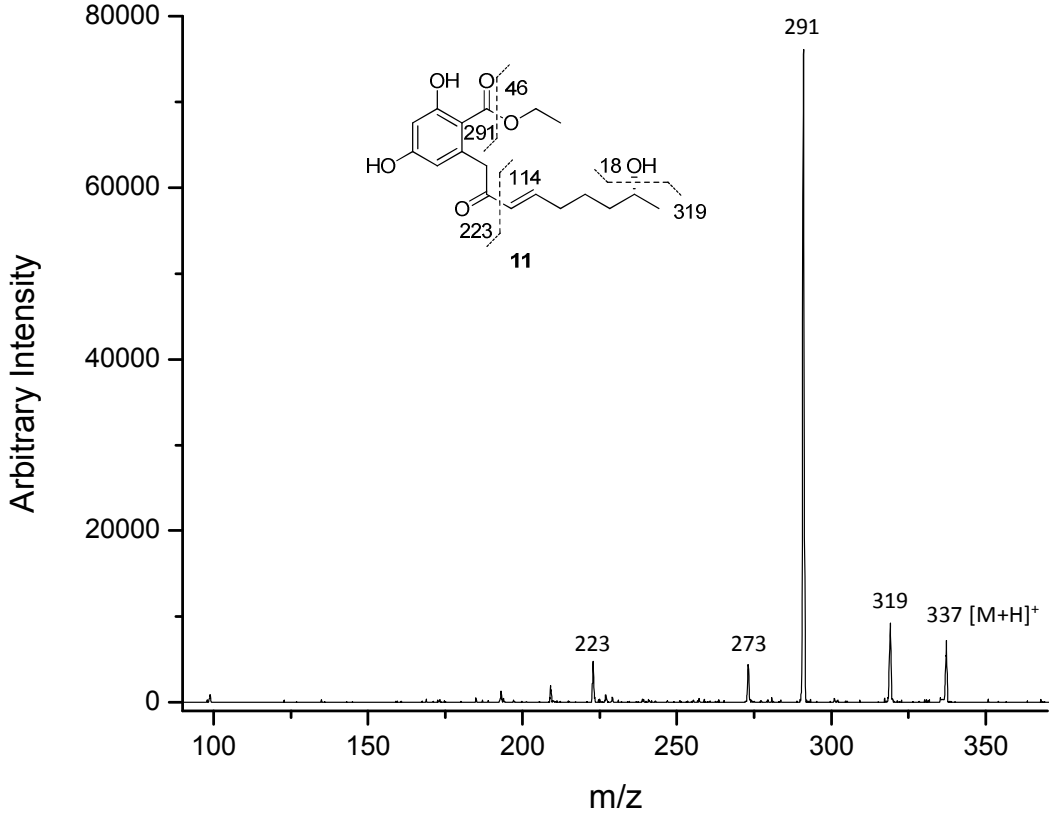


Figure S7. The tandem mass spectrum of 11.

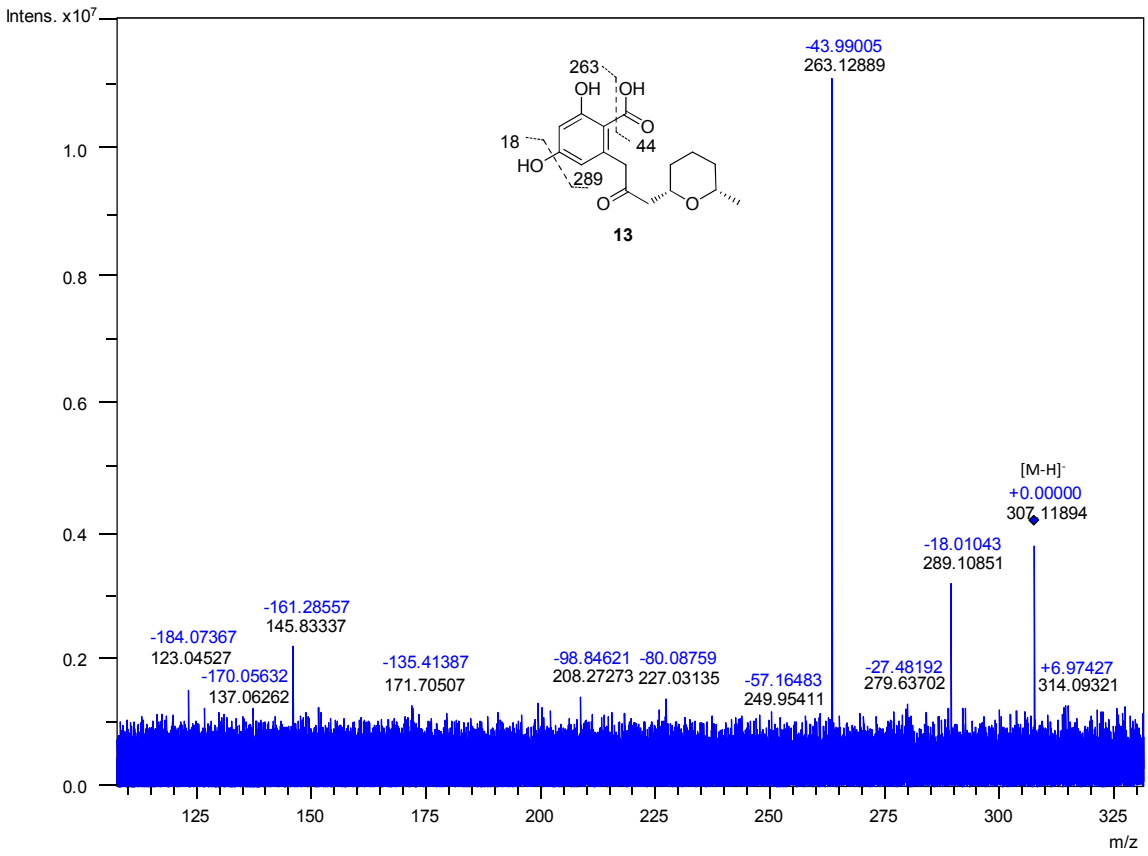


Figure S8. The tandem mass spectrum of 13.

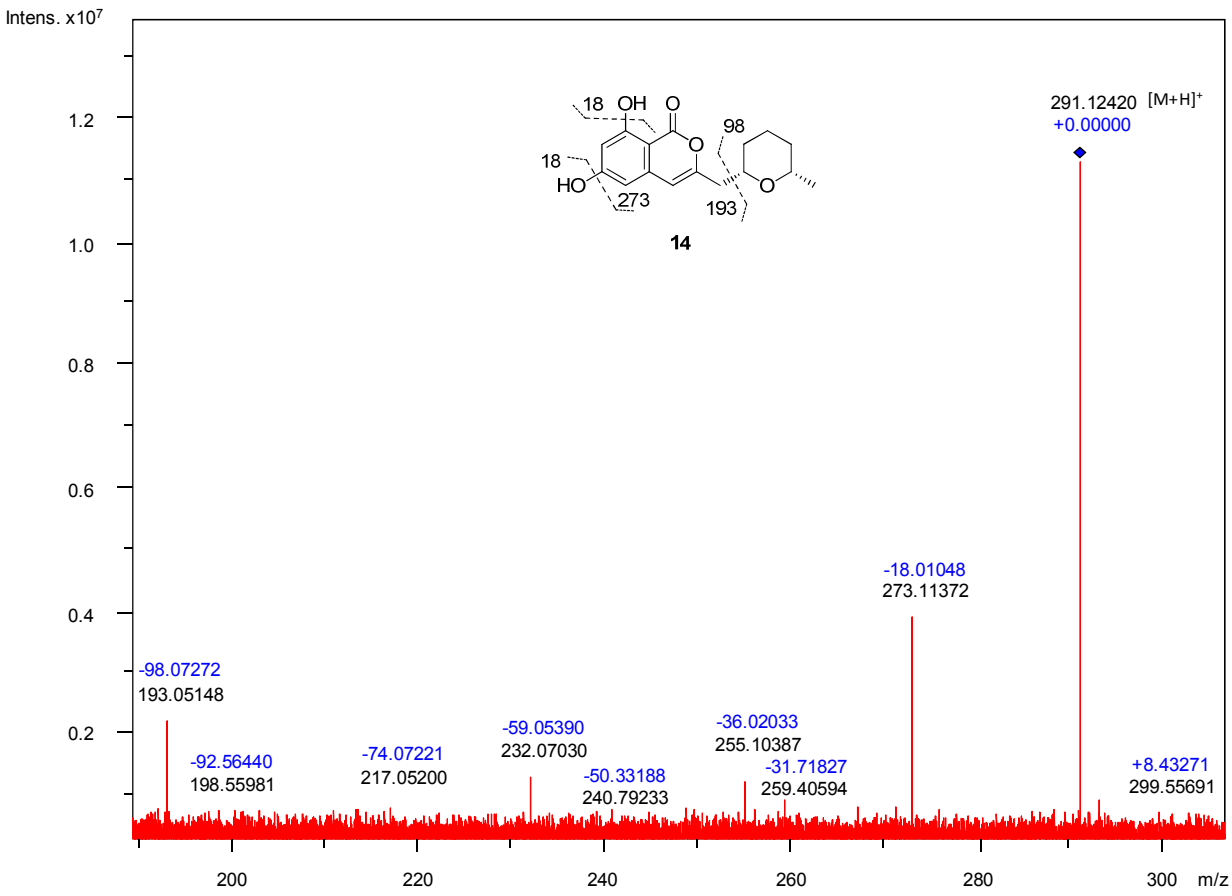


Figure S9. The tandem mass spectrum of 14.

SI References

(1) Xu, Y.; Zhou, T.; Zhou, Z.; Su, S.; Roberts, S. A.; Montfort, W. R.; Zeng, J.; Chen, M.; Zhang, W.; Zhan, J.; Molnár, I. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5398.

(2) Xu, Y.; Espinosa-Artiles, P.; Schubert, V.; Xu, Y. M.; Zhang, W.; Lin, M.; Gunatilaka, A. A. L.; Süssmuth, R.; Molnár, I. *Appl. Environ. Microbiol.* **2013**, *79*, 2038.

(3) Udwary, D. W.; Merski, M.; Townsend, C. A. *J. Mol. Biol.* **2002**, *323*, 585.

(4) Greve, H.; Schupp, P. J.; Eguereva, E.; Kehraus, S.; Kelter, G.; Maier, A.; Fiebig, H.-H.; Konig, G. M. *Eur. J. Org. Chem.* **2008**, *2008*, 5085.

(5) Dai, J.; Krohn, K.; Floerke, U.; Pescitelli, G.; Kerti, G.; Papp, T.; Koever, K. E.; Benyei, A. C.; Draeger, S.; Schulz, B.; Kurtan, T. *Eur. J. Org. Chem.* **2010**, *2010*, 6928.

(6) Zhou, H.; Gao, Z.; Qiao, K.; Wang, J.; Vederas, J. C.; Tang, Y. *Nat. Chem. Biol.* **2012**, *8*, 331.

(7) Zhou, H.; Qiao, K.; Gao, Z.; Vederas, J. C.; Tang, Y. *J. Biol. Chem.* **2010**, *285*, 41412.

(8) 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.