

Published in final edited form as:

J Am Chem Soc. 2011 October 5; 133(39): 15773–15785. doi:10.1021/ja206906d.

Comparative Characterization of Fungal Anthracenone and Naphthacenedione Biosynthetic Pathways Reveals an α -Hydroxylation-Dependent Claisen-Like Cyclization Catalyzed by a Dimanganese Thioesterase

Yanran Li[†], Yit-Heng Chooi[†], Yuewei Sheng[‡], Joan S. Valentine[‡], and Yi Tang^{*,†,‡}

[†]Department of Chemical and Biomolecular Engineering, University of California Los Angeles, 420 Westwood Plaza, Los Angeles, California 90095, USA

[‡]Department of Chemistry and Biochemistry, University of California Los Angeles, 420 Westwood Plaza, Los Angeles, California 90095, USA

Abstract

The linear tetracyclic TAN-1612 (**1**) and BMS-192548 (**2**) were isolated from different filamentous fungal strains, and have been examined as potential neuropeptide Y and neurokinin-1 receptor antagonist respectively. Although the biosynthesis of fungal aromatic polyketides has attracted much interest in recent years, the biosynthetic mechanism for such naphthacenedione-containing products has not been established. Using a targeted genome mining approach, we first located the *ada* gene cluster responsible for the biosynthesis of **1** in *Aspergillus niger* ATCC 1015. The connection between **1** and the *ada* pathway was verified through overexpression of the Zn₂Cys₆-type pathway-specific transcriptional regulator AdaR and subsequent gene expression analysis. The enzymes encoded in the *ada* gene cluster share high sequence similarities to the known *apt* pathway linked to the biosynthesis of anthraquinone asperthecin **3**. Subsequent comparative investigation of these two highly homologous gene clusters by heterologous pathway reconstitution in *Saccharomyces cerevisiae* revealed a novel α -hydroxylation-dependent Claisen cyclization cascade, which involves a flavin-dependent monooxygenase (FMO) that hydroxylates the α -carbon of an ACP-bound polyketide and a bifunctional metallo- β -lactamase-type thioesterase (M β L-TE). The bifunctional M β L-TE catalyzes the fourth ring cyclization to afford the naphthacenedione scaffold upon α -hydroxylation, while performs hydrolytic release of an anthracenone product in the absence of α -hydroxylation. Through *in vitro* biochemical characterizations and metal analyses, we verified that the *apt* M β L-TE is a dimanganese enzyme and requires both Mn²⁺ cations for the observed activities. The M β L-TE is the first example of TE in polyketide biosynthesis that catalyzes the Claisen-like condensation without an α/β hydrolase fold and forms no covalent bond with the substrate. These mechanistic features should be general to the biosynthesis of tetracyclic naphthacenedione compounds in fungi.

INTRODUCTION

Polycyclic aromatic polyketides constitute an important family of natural products and exhibit a broad range of beneficial and deleterious bioactivities, including anticancer (e.g. doxorubicin), antibacterial (e.g. tetracycline), antifungal (e.g. griseofulvin), antiparasitic

Corresponding Author, yitang@ucla.edu.

ASSOCIATED CONTENT

Supporting Information. NMR, UV and LC-MS characterization of compounds described in the main text; protein sequence alignments and additional experimental information. This material is available free of charge via the Internet at <http://pubs.acs.org>

(e.g. frenolicin) and carcinogenic (e.g. aflatoxin)¹ activities. They are synthesized by polyketide synthases (PKSs) from acetate building blocks followed by regioselective cyclizations of the highly reactive poly- β -ketone backbones into different polycyclic ring systems.² Filamentous fungi, like the actinomycetes bacteria, are prolific producers of aromatic polyketides. Fungi use non-reducing polyketide synthases (NRPKSs) to synthesize a diverse array of polycyclic compounds,³ such as the anthraquinone norsolorinic acid that is the precursor of aflatoxin,⁴ the spirobenzofuranone-containing griseofulvin,⁵ the naphtho- γ -pyrone YWA1,⁶ and the anhydrotetracycline-like naphthacenedione viridicatumtoxin (*vrt*)⁷ **4** (Scheme 1). The mechanisms of polyketide chain length control and regioselective cyclization in fungal PKSs are significantly different from the well-studied bacterial type II PKSs, are less understood, and are therefore of significant biochemical interests.⁸

Each NRPKS is an iterative type I megasynthase that contains catalytic domains arranged in a linear fashion.³ The core domains, juxtaposed from the *N*- to *C*-terminus, consist of a starter unit:ACP transacylase (SAT),⁹ a ketosynthase (KS), a malonyl-CoA:ACP transacylase (MAT), a product template (PT) domain^{4,10} and an acyl-carrier protein (ACP)¹¹. Other domains, such as methyltransferase (MeT)¹² and chain-releasing thioesterase/Claisen cyclase (TE/CLC)¹³ or reductase (R)¹⁴ can also be found in some NRPKSs. NRPKSs utilize PT domains to control the cyclization regioselectivity of poly- β -ketone backbones synthesized by the minimal PKS domains (KS, MAT and ACP).^{4,10a} The PT domains of most NRPKSs can be classified phylogenetically into five major groups with each group corresponding to a unique first-ring cyclization regioselectivity and/or product size: Group I, C2-C7 and monocyclic; Group II, C2-C7 and bicyclic; Group III, C2-C7 and multicyclic; Group IV, C4-C9 and muticyclic; and Group V, C6-C11 and multicyclic.¹⁵ In each group, the even-numbered α -carbon serves as the nucleophile and attacks the odd-numbered carbonyl via an aldol condensation to form a new 6-membered ring. Group V NRPKSs are unique in that they lack the *C*-terminal fused TE/CLC domain that catalyzes Claisenlike C-C bond formation and product release. Instead, a standalone metallo- β -lactamase-type thioesterase (M β L-TE) is typically encoded in the vicinity of the NRPKS within the gene clusters.¹⁶ Recently, two Group V NRPKSs, AptA¹⁷ from *Aspergillus nidulans* and ACAS¹⁶ from *Aspergillus terreus*, were linked to the biosynthesis of the anthraquinone asperthecin (*apt*) **3** and anthracenone atrochrysone **5** (Scheme 1), respectively. Each cluster contains a standalone M β L-TE (AptB and ACTE) gene. In the biosynthesis of **5**, atrochrysone carboxylic acid was shown to be hydrolyzed off the ACP domain of ACAS by ACTE and then spontaneously decarboxylated to yield **5**.¹⁶

The Group V PKS VrtA from *Penicillium aethiopicum* was recently linked to the biosynthesis of the tetracyclic naphthacenedione core of **4**.¹⁸ Comparing the *vrt* and *apt* gene clusters revealed that although the structures of **4** and **3** differ significantly, the protein sequences of the core PKS genes are remarkably similar. VrtA and the associated M β L-TE VrtG share high sequence similarities to AptA and AptB (60% and 62% identity, respectively). The activities of the VrtA and AptA PT domains were shown to be functionally equivalent when spliced into different NRPKSs.¹⁵ For example, when inserted into a nonaketide (C₁₈) NRPKS, the nascent polyketide was cyclized by either Group V PT domain through C6-C11 and C4-C13 aldol condensations to afford a bicyclic intermediate (Scheme 2).¹⁵ The third ring is likely spontaneously cyclized via C2-C15 aldol addition. Subsequently, in the absence of additional enzymatic modifications, spontaneous O17-C1 α -pyrone formation results in the release of the pyranoanthracenone **6a**, which is dehydrated and oxidized into **6** (Scheme 2). In the biosynthesis of **4** from a decaketide (C₂₀) backbone, the VrtA-PT is likely to function in a similar way to generate the bicyclic intermediate through the programmed cyclization steps. However, in order to form the tetracyclic naphthacenedione, an additional Claisen-like condensation between C18 and C1 must take

place regioselectively following formation of the tricyclic intermediate. The enzymatic basis of this reaction has not yet been elucidated.

Because of the similarities between the core PKS genes in the anthracenone and naphthacenedione biosynthetic pathways, a comparative investigation involving the mix and match of individual enzymes may reveal insights into the mechanism of the fourth ring cyclization. Although the biosynthetic locus of **4** has been identified,¹⁸ the requirement of the unusual malonamyl starter unit has complicated the investigation of VrtA functions. As a result, the acetate-primed TAN-1612 **1** (neuropeptide Y antagonist) from *Penicillium claviforme* FL-27337¹⁹ was identified as our naphthacenedione target in this study (Scheme 1). A related compound BMS-192548 **2** (neurokinin-1 receptor antagonist) was isolated from *Aspergillus niger* WB2346 and is likely tautomerized from **1** during purification.²⁰ For parallel biosynthetic investigation of a fungal anthracenone, the *apt* pathway that produces **3** is selected. The anthraquinone **3** is likely derived from an anthracenone precursor such as **5**. Here, we first mined the biosynthetic gene cluster of **1** (gene cluster designated as *ada*, 2-acetyl-2-decarboxamidoanthrotainin) from *A. niger* ATCC 1015 using Group V PKS PT sequence as a lead. Characterizations of the biosynthetic pathways of **1** and **3** through heterologous reconstitution in *Saccharomyces cerevisiae* and *in vitro* biochemical investigations uncovered that the C18-C1 fourthring Claisen-like condensation is catalyzed by a dimanganese M β L-TE and is dependent on C2-hydroxylation by a flavin-dependent hydroxylase. The M β L-TE is the first example of TE in polyketide biosynthesis that catalyzes the Claisen-like condensation without an α/β hydrolase fold and forms no covalent bond with the substrate.

MATERIALS AND METHODS

Strains and Culturing Conditions

A. niger ATCC 1015 was obtained from NRRL and cultured at 28°C. The *E. coli* strain XL-1 Blue (Stratagene) and TOPO10 (Invitrogen) were used for DNA manipulation, and BL21(DE3) (Stratagene) was used for protein expression. BAP1 described by Pfeifer *et al.* was obtained from Prof. Khosla.²¹ *S. cerevisiae* BJ5464-NpgA (*MAT α ura3-52 trp1 his3 Δ 200 leu2 Δ 1 pep4::HIS3 prb1 Δ 1.6R can1 GAL*) was used as the yeast expression host.²² *S. cerevisiae* transformants were selected with the appropriate complete minimal dropout agar plates, and cultured in YPD medium with 1% dextrose at 28°C.

General Techniques for DNA Manipulation

A. niger genomic DNA was prepared using the ZYMO ZR fungal/bacterial DNA kit according to supplied protocols. PCR reactions were performed with Phusion high-fidelity DNA polymerase (New England Biolabs, NEB) and Platinum Pfx DNA polymerase (Invitrogen). PCR products were subcloned into pCR-Blunt vector (Invitrogen) and confirmed by DNA sequencing. Restriction enzymes (NEB) and T4 ligase (Invitrogen) were used to digest and ligate the DNA fragments, respectively. Primers used to amplify the genes were synthesized by Integrated DNA Technologies and are listed in Table S1.

Construction of pYR311

pBARGPE1 (obtained from FGSC) is a fungal expression vector and contains the *A. nidulans gpdA* promoter and *trpC* terminator.²³ *adaR* was amplified from *A. niger* genomic DNA with primers *adaR-f* and *adaR-r*. The resulting DNA fragment was digested with *EcoRI* and *XhoI* and inserted in to pBARGPE1 (digested with *EcoRI* and *XhoI*) to yield pYR311.

Transformation of *A. niger* with pYR311

Polyethylene glycol-mediated transformation of *A. niger* was conducted essentially as described previously for *A. nidulans*, except that the protoplasts were prepared with 3 mg/mL lysing enzymes (Sigma-Aldrich) and 2 mg/mL Yatalase (Takara Bio.).^{18,24} Glufosinate used for the selection of bar transformants was prepared as previously described.¹⁸ Miniprep genomic DNA from *A. niger* transformants was prepared according to previous procedures for the screening.²⁵ Transformants were cultured in the presence of 4–8 mg/mL glufosinate for screening.¹⁸ Integration of the AdaR-overexpression cassette was confirmed by PCR using GoTaq Green Master Mix (Promega) and two pairs of primers: bar-f and bar-r, GPE-*gpdA*-f and *adaR*-r.

Expression analysis by Reverse Transcription Polymerase Chain Reaction (RT-PCR)

The total RNA of *A. niger* mutants and wild-type strains were extracted using the Qiagen RNeasy plant mini kit following the supplied protocols. The first strand cDNA was synthesized using the gene specific reverse primers (*adaA*-r, *adaB*-r, β -*tub*-r) and Improm-II reverse transcription system (Promega) according to the manufacturer's instructions. cDNA was then amplified with GoTaq Green Master Mix (Promega) using three pairs of primers: *adaA*-r and *adaA*-PT-f, *adaB*-r and *adaB*-f, β -*tub*-r and β -*tub*-f.

Purification and Characterization of Compounds

The *A. niger* transformant T4 producing **1** was cultured on 1 L GMM agar for 3 days at 28°C. The agar was extracted three times with equal volume of ethyl acetate (EtOAc). The resultant organic extracts were combined, dried over anhydrous Na₂SO₄, and concentrated *in vacuo*. The crude extract was then washed with hexane, and redissolved in 1:1 methanol (MeOH)-chloroform (CHCl₃). A Sephadex LH-20 column was used to fractionate the crude extract with 1:1 MeOH-CHCl₃ as the mobile phase. Fractions that contained **1** were combined and dried. The residue was redissolved in HPLC grade MeOH and purified by reverse-phase HPLC (Alltech Alltima 5 μ m, 250 mm \times 10 mm) on a linear gradient of 35 to 95% acetonitrile (CH₃CN, v/v in water) over 60 minutes and 95% CH₃CN (v/v) over 15 minutes at a flow rate of 2 mL/min. 1D and 2D NMR of **1** were performed on a Bruker DRX-500 spectrometer using CDCl₃ as the solvent. **7**, **8**, **9** and **10** were extracted from a 2L YPD culture of the corresponding *S. cerevisiae* transformants, purified and analyzed in deuterated NMR solvents (MeOD for **7**, acetone-d₆ for **8** and **9**, DMSO-d₆ for **10**) using similar procedures. The extraction and purification of **7** requires acidic condition.

Construction of *S. cerevisiae* Expression Plasmids

pKOS518-120A and pKOS48-172b are 2 μ -based yeast-*E. coli* shuttle plasmids containing different auxotrophic markers (TRP1 and URA3, respectively).²⁶ Primers used to amplify *adaA*, *adaB*, *adaC*, *adaD* are listed in supplemental Table S1. *adaA* was amplified by PCR from *A. niger* genomic DNA with the primers *adaA*-f and *adaA*-r and inserted into pCR-Blunt. The resulting vector was digested with *Spe*I and *Pme*I and the *adaA* gene was ligated in pKOS48-172b digested with *Spe*I and *Pml*I to yield pYR291. Similar methods were used to insert *adaB*, *adaC* and *adaD* in pKOS518-120A digested with *Nde*I and *Pme*I to yield pYR301, pYR313 and pYR315, respectively. The *ADH2p-adaC-ADH2t* cassette was obtained by digesting pYR313 with *Sal*I and *Xho*I, and inserting the fragment into the pYR301 digested with *Xho*I to yield pYR330. The *ADH2p-adaD-ADH2t* cassette obtained by digesting pYR315 with *Hind*III and *Xho*I, and inserting the fragment into pYR330 digested with *Hind*III and *Sal*I to yield pYR342. The expression plasmids for *apt* genes were constructed using similar methods, except *aptA* was inserted in the pKOS48-172b digested with *Nde*I and *Pml*I to yield pYR203; *aptB* and *aptC* were inserted in the pKOS518-120A digested with *Nde*I and *Pme*I to yield pYR204 and pYR260, respectively. The *ADH2p-aptC*-

ADH2t cassette was obtained by digesting pYR255 with *SalI* and *XhoI*, and was inserted into pYR204 digested with *SalI* to yield pYR260. The introns of *adaB*, *aptA*, *aptB*, *aptC* were removed by using splice by overlapping extension (SOE)-PCR.

Reconstitution and Investigation of the Biosynthetic Pathway of 1 in *S. cerevisiae*

Various combinations of the expression plasmids were introduced into *S. cerevisiae* strain BJ5464-NpgA²² by using the *S. c.* EasyCompTM Transformation Kit (Invitrogen). The cells were grown at 28°C in 10 mL YPD media with 1% dextrose. For product detection, 1 mL of cell culture was collected and analyzed every 12 hours after induction for 72 hours. The medium was extracted with equal volume of 99% EtOAc/1% acetic acid (AcOH), and the cell pellets were extracted with 100 μ L acetone. The organic phase was separated, evaporated to dryness, redissolved in 20 μ L MeOH and analyzed with a Shimadzu 2010 EV Liquid Chromatography Mass Spectrometer (LC-MS, phenomenex Luna 5 μ m, 2.0 \times 100 mm) using both positive and negative electrospray ionization. Compounds were separated on a linear gradient of 5% CH₃CN (v/v in water, 0.1% formic acid) to 95% CH₃CN (v/v in water, 0.1% formic acid) over 30 min with a flow rate of 0.1 mL/min.

Construction of *E. coli* Expression Plasmids

The DNA fragments encoding AdaA, AdaC, AptB and AptC were each inserted into pET24 digested with *NdeI* and *NotI* to yield pYR303, pYR316, pYR146 and pYR201, respectively. The site-directed mutants of AptB were constructed through SOE-PCR and inserted into pET24. The DNA fragments encoding AptA was inserted into pET28 digested with *NdeI* and *EcoRI* to yield pYR62.

Expression and Purification of Enzymes

Expression and purification procedures of AdaA, AdaC, AptB, AptC, AptA and MatB followed identical procedures. The expression and purification of AdaA is detailed here as an example. pYR303 was used to transform the BAP1 for protein expression.²¹ The cells were cultured at 37°C and 250 rpm in 500 mL LB medium supplemented with 35 μ g/mL kanamycin to a final OD₆₀₀ between 0.4 and 0.6. The culture was then incubated on ice for 10 min before addition of 0.1 M isopropylthio- β -d-galactoside (IPTG) to induce protein expression. The cells were further cultured at 16 °C for 12–16 h. The cells were then harvested by centrifugation (3,500 rpm, 15 min, 4°C), resuspended in ~25 mL lysis buffer (20 mM Tris-HCl, pH 7.9, 0.5 M NaCl, 10 mM imidazole) and lysed by sonication on ice. Cellular debris was removed by centrifugation (15,000 rpm, 30 min, 4°C), and Ni-NTA agarose resin was then added to the supernatant (1–2 mL/liter of culture). The suspension was swirled at 4°C for 2 h and loaded into a gravity column. The protein-bound resin was washed with one column volume of 10 mM imidazole in buffer A (50 mM Tris-HCl, pH 7.9, 2 mM EDTA, 2 mM DTT), followed by 0.5 column volume of 20 mM imidazole in buffer A. The protein was then eluted with 250 mM imidazole in buffer A. Purified AdaA was concentrated and exchanged into Buffer A+10% glycerol with the centriprep filters (Amicon), aliquoted and flash frozen using acetone/dry ice. Protein concentration was determined with the Bradford assay using bovine serum albumin as a standard.²⁷ For the expression of Mn²⁺-AptB or Fe²⁺-AptB, 200 ppm of MnSO₄ or FeSO₄ was added to the expression media. Metal contents of various AptB and mutations were determined using an inductively coupled plasma atomic emission spectrometer (ICP-AES, TJA Radial IRIS 1000).

In vitro Reconstitution of Enzymes

The reactions were performed at 100 μ L scale containing 100 mM phosphate buffer (pH 7.4) in the presence of 100 mM sodium malonate, 2 mM DTT, 7 mM MgCl₂, 20 mM ATP,

5 mM coenzyme A, 20 μ M MatB, with various combinations of NRPKSs and the tailoring enzymes.²⁸ Typically, the concentration of the NRPKS (i.e. AptA or AdaA) was 10 μ M and the tailoring enzymes (i.e. AdaC, AptB, AptC) were 50 μ M. The reactions were quenched with 1 mL of 99% EtOAc/1% AcOH. The organic phase was separated, evaporated to dryness, redissolved in 20 μ L of DMSO and analyzed with LC-MS as previously described.

RESULTS

Mining of the *ada* Gene Cluster in *A. niger*

Although *A. niger* WB2346 is the only *A. niger* species reported to produce **1**,²⁰ we postulated that the genome-sequenced *A. niger* strains may also contain gene clusters that can biosynthesize **1**. To test this idea, we scanned the sequenced genomes of *A. niger* ATCC 1015 and CBS513.88²⁹ for genes encoding Group V NRPKSs. A BLAST search of both genomes with AptA-PT¹⁵ uncovered a unique Group V NRPKS gene (An11g07310) in the CBS 513.88 strain. The corresponding ortholog (e_gw1_4.29) is also found in the ATCC 1015 strain, but the coding region is incomplete. The gaps were filled by PCR and sequencing, which revealed that the Group V NRPKSs encoded in both strains share >98% protein sequence identity (Table 1). The gene products downstream of the *A. niger* ATCC 1015 NRPKS (named AdaA) include a putative M β L-TE AdaB, a FAD-dependent monooxygenase (FMO) AdaC, an *O*-methyltransferase (*O*-MeT) AdaD and a putative Zn₂Cys₆ family transcription regulator AdaR. (Figure 1A, Table 1). The core biosynthetic enzymes AdaA, AdaB and AdaC all show very high sequence similarities to the corresponding *apt* and *vrt* enzymes (Table 1). We reason that the decaketide backbone of **1** may be synthesized and cyclized by AdaA. While the C15-OH of **1** is expected to be retained from an acetate unit during the C2-C15 aldol addition as in the biosynthesis of **5**, the C2-OH may be added by the FMO AdaC,¹⁶ which is a homolog of AptC.¹⁷ AdaD then performs the C9 *O*-methylation to complete the biosynthesis of **1**. Since the anthracenone hydrolysis activity observed for the close homolog ACTE is not applicable here, the role of the M β L-TE AdaB in the formation of **1** is not yet clear. Given that no additional enzyme from this putative gene cluster appears likely as a fourth ring Claisen cyclase, we propose that AdaB may be involved in facilitating the C18-C1 C-C bond formation and concomitant release of **1**.

Culturing of *A. niger* ATCC 1015 in different media and under various growth conditions did not lead to the production of **1** (Figure 1B, trace *i*), suggesting that the biosynthesis of **1** may be silent. Overexpression of pathway-specific transcriptional regulators has been a successful strategy for activating cryptic metabolic pathways in fungi.³⁰ AdaR is likely involved in the regulation of the *ada* gene cluster and maybe similarly overexpressed to activate the biosynthetic pathway. To do so, *adaR* was placed under the regulation of the *A. nidulans* *gpdA* promoter (*P_{gpdA}*) in pBARGPE1,²³ which carries the *bar* gene.³¹ The resulting plasmid pYR311, was transformed into *A. niger* ATCC 1015. Several glufosinate-resistant colonies exhibited yellow pigmentation on agar plates not observed in the wild type (WT) strain when grown on glucose minimal medium (Figure S1). Insertion of the *P_{gpdA}::adaR* and the *bar* resistant gene into the pigmented transformants T1–T5 were confirmed by PCR (Figure S1). Pigmented agar of one transformant (T4) was then extracted and analyzed by LC-MS (Figure 1B, trace *ii*, Figure S2). A new compound at RT (retention time) = 26.2 min, with a UV spectrum and *m/z* consistent to those reported for **1** was overproduced in T4 (Figure 1B).¹⁹ RT-PCR analysis of *adaA* and *adaB* confirmed activated transcriptions of both genes in the T4 strain but not in WT, which supports the involvement of the *ada* gene cluster in the biosynthesis of the yellow pigment in the strain (Figure 1C).

The yellow pigment was subsequently purified from T4, extensively characterized by NMR, and confirmed to be **1** (Figure S3, Table S3). Taken together, these results verified the

proposed correlation between the *ada* gene cluster and the biosynthesis of **1** in *A. niger* ATCC 1015. Overexpression of *adaR* led to an increased titer of **1** in *A. niger* ATCC 1015 from undetectable levels to ~ 400 mg/L, which represents a 100-fold increase compared to the titer reported from *A. niger* WB2346 (2.9 mg/L).^{20b}

Heterologous reconstitution of *ada* and *apt* pathways in BJ5464-NpgA

To verify if the four-enzyme ensemble (AdaA-D) is sufficient for the biosynthesis of **1**, we reconstituted the putative four-gene pathway in *S. cerevisiae* BJ5464-NpgA, which is a vacuolar protease-deficient strain carrying the *A. nidulans* phosphopantetheinyl transferase *npgA*.²² The intron-free coding sequences of the four genes were amplified from *A. niger* ATCC 1015 genomic DNA using SOE-PCR and placed under the control of the ADH2 promoter. Two resulting yeast 2 μ expression plasmids pYR291 (carrying *adaA*) and pYR342 (carrying *adaB-D*) were transformed into BJ5464-NpgA (Table S2). The two-day culture of the yeast transformant synthesized **1** as essentially the only polyketide product with a titer of ~ 10 mg/L (Figure 2A, trace *ii*, Table 2). Subsequently, AdaD was excluded from the heterologous host (Table S2) and the resulting strain produced a new compound (~8 mg/L) with UV spectrum (Figure 2A, trace *iii*, Table 2) and mass (m/z [M+H]⁺ = 401) consistent with that of **7** (Figure S4, Scheme 3). The structure of **7** was verified to be the C9-hydroxyl version of **1** by NMR analysis (Figure S5, Table S4), which confirmed that AdaD is indeed the C9-*O*-methyltransferase. The efficient production of **7** in BJ5464-NpgA shows that AdaA, AdaB and AdaC are sufficient for the biosynthesis of a naphthacenedione, and sets the stage for more systematic investigation into the roles of each enzyme in the yeast host. In subsequent studies, *adaD* was excluded from the heterologous host/plasmid combinations.

Using the same approach as above, we reconstituted the activities of AptA, AptB and AptC from the biosynthetic pathway of **3**. The two expression plasmids encoding the three *apt* genes (pYR203 and pYR260, Table S2) were introduced into BJ5464-NpgA, and the resulting strain synthesized a new compound **8** (m/z [M+H]⁺ = 333) (Figure 2B, trace *ii*, Figure S6). The molecular weight of **8** agrees with that of a nonaketide that has undergone two dehydrations, one decarboxylation and one hydroxylation. This was surprising since we originally anticipated the NRPKS AptA to be an octaketide (C₁₆) synthase based on the size of the final product **3**.¹⁷ **8** was purified from the yeast culture and the structure was established to be a multiply hydroxylated anthracenone (Scheme 2, Figure S7, Table S5). The coordinated activities of AptA, AptB and AptC in the biosynthesis of **8** are shown in Scheme 2. The conversion of **8** to **3** may require additional *A. nidulans* enzymes encoded outside of the compact *apt* cluster.

Comparative analysis of the *ada* and *apt* pathways through gene shuffling

The heterologous reconstitution of AdaA-C and AptA-C clearly demonstrate these two sets of highly similar enzymes synthesize the structurally distinct **7** and **8**, respectively. This raises the obvious question of how hydrolysis of **8** and fourth-ring cyclization of **7** may be accomplished by the corresponding M β L-TEs, which are highly similar in sequence (Table 1). To dissect the roles of each enzyme and the biochemical differences between the homologous enzymes, we constructed various shuffled pathways in BJ5464-NpgA (Table 2). When AptA was combined with the *ada* tailoring enzymes AdaB and AdaC, biosynthesis of anthracenone **8** was observed as the major product with similar yield as the intact *apt* pathway (Figure 2B, trace *iii*, Table 2). This indicates that AdaB and AdaC can efficiently function on a smaller, nonaketide substrate. Conversely, coexpression of AptB and AptC with AdaA produced **7** as the major product, albeit at a yield that is ~10-fold lower than the intact *ada* pathway (Figure 2A, trace *iv*, Table 2). The lower yield may be attributed to the less efficient function of AptC or AptB on the larger decaaketide substrate (more evidence on

this later). The biosynthesis of **7** and **8** by these shuffled pathways implicates that i) the polyketide chain length control lies in the NRPKSs AdaA and AptA; ii) both the M β L-TE and the FMO in the two pathways can process substrates of unnatural chain lengths and are therefore functionally homologous (with some differences in efficiency); and iii) most importantly, both M β L-TEs appear to be bifunctional and can catalyze either hydrolysis (towards the formation of **8**), or the Claisen-like fourth-ring cyclization (towards the formation of **7**) when the substrate is of sufficient chain length (C₂₀).

We next excluded either AptC or AdaC from the reconstituted pathways to verify the role of the FMO and the necessity of the modification to subsequent steps. When only AptA and AptB were expressed in BJ5464-NpgA, a major product **9** with a mass of 316 was recovered. (Figure 2B, trace *iv*, Figure S8, Table 2). Following large scale culture and NMR characterizations, the structure of **9** was shown to be identical to **8** except the absence of C2-OH (Figure S9, Table S6). This single structural difference between **8** and **9** confirms that AptC is the C2-hydroxylase (Scheme 2). Following thioester hydrolysis by AptB to release **9a**, spontaneous decarboxylation occurs to afford the anthracenone **9**. The C2-hydroxylation does not appear to affect the hydrolytic activities of AptB since both **8** and **9** were produced at comparable levels (~10 mg/L) from the heterologous yeast host. As expected, substitution of AptB with AdaB also resulted in hydrolysis and production of **9** (Figure 2B, trace *v*, Table 2).

On the other hand, coexpression of AdaA and AdaB in BJ5464-NpgA did not yield **11**, which is an analog of **7** that is unhydroxylated at C2 (Scheme 3). Instead, the resulting strain synthesized **10** (m/z [M+H]⁺ = 359) (Figure 2A, trace *v*, Table 2, Figure S10), the molecular weight of which is consistent with a decaketide that has undergone two dehydrations and one decarboxylation. The structure of **10** was analyzed by NMR (Figure S11, Table S7) and was confirmed to be a tricyclic, 1,3-diketone-containing anthracenone (Scheme 3). Compared to **9**, **10** contains an additional ketone at C18 that is reflective of the longer polyketide backbone synthesized by AdaA. While the absence of the C2-hydroxyl also verifies the role of AdaC, the failure of AdaA and AdaB to synthesize **11** in the absence of AdaC hints that C2-hydroxylation may be essential for the Claisen cyclization. In the absence of the C2-OH modification, AdaB acts as a hydrolytic TE to release the diketo-anthracenone carboxylic acid **10a** from AdaA, which then spontaneously decarboxylates to form **10**. This result supports the timing of C2-hydroxylation to precede cyclization as shown in Scheme 3. Therefore, the α -hydroxylation step appears to be crucial for switching the nucleophile in the chain releasing step catalyzed by AdaB, from a hydroxide ion to the C18 enolate anion. Lastly, BJ5464-NpgA transformants expressing AdaA and AptB did not produce any obvious products (Figure 2A, trace *vi*, Table 2). This indicates a subtle difference between the two M β L-TEs in the hydrolysis reaction: while AdaB can hydrolyze both C₁₈ and C₂₀ anthracenone-S-ACP substrates to yield **9** and **10**, respectively, AptB can only hydrolyze the C₁₈ substrate to yield **9**.

To probe whether offloading of **7** and **8** indeed requires the M β L-TE, we coexpressed the individual NRPKSs with their corresponding FMO only. BJ5464-NpgA expressing *aptA* and *aptC* did not produce any polyketide (Figure 2B, trace *vi*), confirming that nonenzymatic hydrolysis to yield **8a** does not occur. In contrast, BJ5464-NpgA expressing AdaA and AdaC produced trace amounts of **7** (Figure 2A, trace *vii*). Hence the C18-C1 Claisen cyclization can take place uncatalyzed with very low efficiency, and is significantly enhanced by AdaB.

In vitro* Biosynthesis of **7** and **8*

To complement the heterologous biosynthesis results, we reconstituted the synthesis of **7** and **8** using purified enzymes. His₆-tagged AptA and AdaA were expressed and purified

from *E. coli* BAP1,²¹ while the other enzymes (AptB, AptC, AdaC) were expressed and purified from BL21(DE3). AptB purified from the culture medium without supplementation of metal ions is designated as AptB*. Soluble forms of AdaB could not be obtained from *E. coli* and is therefore not included in the assays. Purified AptC and AdaC were thermally denatured and the prosthetic groups were confirmed to be FAD by LC-MS analysis as previously described (Figure S12).³² To generate malonyl-CoA *in situ* for use in polyketide turnover assays, we used the malonyl-CoA synthetase MatB from *Rhizobium trifolii* to produce the extender units from malonate, ATP and CoA.²⁸

Upon mixing AptA, AptB* and AptC in the presence of the required reagents and cofactors for 12 hours, the organic products were analyzed by LC-MS. In agreement with the heterologous reconstitution results, the *in vitro* reaction produced **8** as the dominant product, while excluding AptC from the reaction yielded **9** (Figure 3A, traces *i* and *ii*). When AptB and AptC were both absent from the reaction mixture, AptA synthesized **6** from the spontaneous lactonization of the fourth ring (Figure S13), which was previously observed with other hybrid Group V nonaketide NRPKS.¹⁵ Therefore, the activities observed in the yeast host can be completely recapitulated by using purified enzyme components. To test if the hydroxylation of C2 occurs on the free standing substrate **9** to yield **8**, AptC was assayed with **9** in the presence of reduced nicotinamide adenine dinucleotide phosphate (NADPH). However, no conversion was observed (Figure S14), which is in agreement with the proposed mechanism that AptC hydroxylates the anthracenone carboxyl thioester attached to the AptA-ACP (Scheme 2).

In vitro reactions using purified components also enabled us to prove the trace amounts of **7** synthesized by AdaA and AdaC is not an artificial result from the actions of endogenous yeast enzymes. Indeed, release of small amounts of **7** can be detected after overnight incubation of AdaC and AdaA (Figure 3A, trace *iii*). Inclusion of AptB* in the assay, however, led to a higher level synthesis of **7** (Figure 3A, trace *iv*). To quantify the role of AptB* in accelerating the cyclization, we monitored the time-course production of **7** by using an HPLC assay (Figure 3B). Following the reaction progress from 5 to 120 minutes after initiation showed that AptB* can increase the turnover of **7** up to 3-fold in two hours.

Combining the *in vivo* and *in vitro* results, the dual catalytic functions of fungal PKS M β L-TEs such as AdaB and AptB can be established as shown Schemes 2 and 3. In both cases, the M β L-TE is responsible for product turnover. First, M β L-TE catalyzes the hydrolysis of anthracenone carboxylic acids such as **9a** and **10a** from the NRPKS with certain chain-length specificities; second, when the decaketide anthracenone carboxyl-S-ACP is hydroxylated at C2 by the FMO, the M β L-TE accelerates the Claisen-like condensation between C18 and C1 to release the naphthacenedione product from the NRPKS.

Biochemical Characterization of AptB as a Bifunctional TE

As proposed in the study of ACTE, M β L-TEs belong to the metallo- β -lactamase (M β L) superfamily, and consist of the conserved divalent metal binding motif HXHXDH.¹⁶ In addition, alignment of the sequences of AptB, AdaB, VrtG and ACTE shows that these M β L-TEs are similar to the subclass B3 M β Ls (Figure S15),³³ which employ a HXHXDH(X)_iH(X)_jH motif to bind two divalent transition metal ions such as Zn²⁺ and Fe²⁺. The metal contents of AptB* was determined by ICP-AES, and showed that AptB* contains Fe²⁺, Mn²⁺ and Ni²⁺ ions and more than 55% of AptB* has no metal ions in either of the metal binding sites (Table S8). We believe that Ni²⁺ ions originated from the affinity purification step. As zinc ions are naturally abundant, the presence of Zn²⁺ in AptB* at such low concentration suggests the unlikelihood of Zn²⁺ being the physiological cofactor. Dialysis of AptB* with 10 mM EDTA led to significant precipitation, likely due to instability of the apo AptB as a result of metal loss. ICP-AES data on AptB* that remained

in solution showed a significant decrease in the metal content of divalent metal ions other than Mn^{2+} (Table S8), which suggests that the other metal ions (Fe^{2+} , Ni^{2+} , Zn^{2+}) are loosely bound in AptB*. To obtain a higher percentage of AptB in holo form, 200 ppm of MnSO_4 or FeSO_4 was supplemented to the AptB expression medium. The metal content of AptB purified from the Mn^{2+} -enriched medium (Mn^{2+} -AptB) was determined to be 1.5 Mn^{2+} cations per monomer (Table S8), while Fe^{2+} -AptB (purified from Fe^{2+} -enriched medium) contained 0.9 Fe^{2+} and 0.4 Mn^{2+} cations per monomer (Table S8). This evidence strongly suggests dinuclear Mn^{2+} as the physiologically preferred metal cofactor of AptB. Dialysis of Mn^{2+} -AptB with 10 mM EDTA did not result in notable precipitates, indicating the enzyme- Mn^{2+} interactions are strong.

The catalytic functions of AptB*, Fe^{2+} -AptB and Mn^{2+} -AptB were then compared by measuring the time courses of *in vitro* production of **7**, **8** or **9** (Figure 3B–D). As shown in Figure 3B–D, Mn^{2+} -AptB is much more efficient (up to three to four-fold higher in product turnover) in catalyzing both the hydrolysis and Claisen-like condensation than Fe^{2+} -AptB and AptB*. In the presence of Mn^{2+} -AptB, the apparent production rate of **7** is ~17 ng/min, while that of **8** or **9** is ~25-fold lower at 0.7 ng/min. Therefore, even though the natural role of AptB in the *apt* pathway is hydrolysis to yield **9a**, it can promote the C18-C1 Claisen-like condensation much more efficiently than the hydrolysis reaction under the assay conditions.

To visualize the metal binding sites, we generated a homology structural model of AptB from the I-TASSER server (Figure 4A, left panel).³⁴ Comparative analysis with the previously reported structures of MβL superfamily enzymes, such as ST1585³⁵ from *Sulfolobus tokodaii* (Figure 4A, right panel) and YcbL³⁶ from *Salmonella enteric* serovar Typhimurium, suggests that H97, H99, D101, H102, H153 and H207 of AptB are the six residues that chelate the two Mn^{2+} cations. Among them H97, H99 and H153 are predicted to bind one Mn^{2+} atom, while D101, H102 and H207 are predicted to bind the other Mn^{2+} .³⁷ In addition, the binuclear Mn^{2+} ions are normally bridged with a hydroxide ion and an aspartate side chain, which is from the conserved D172 among AptB, AdaB and VrtG (Figure 4A, Figure S15).³⁷

Site-directed mutations to alanines at these seven residues in Mn^{2+} -AptB were then performed to confirm their roles in metal binding and catalytic activity. All these mutants were expressed and purified from *E. coli* BL21(DE3) in the Mn^{2+} -enriched medium with similar yields as the wild type Mn^{2+} -AptB. Secondary structure analysis using far-UV circular dichroism confirmed that the mutants did not undergo significant structural changes when compared to Mn^{2+} -AptB or apo-AptB (Figure S16). The hydrolytic activities of these mutants were assayed along with AptA and estimated by the production of **9**; whereas the Claisen-like cyclization activities were assayed in the presence of AdaA/AdaC and estimated by the relative levels of **7** (wild type Mn^{2+} -AptB is defined to be 1). As expected, the hydrolytic activities of these seven mutants (H97A, H99A, D101A, H102A, H153A, D172A and H207A) were completely abolished, while the levels of **7** were attenuated to that of uncatalyzed levels with AdaA/AdaC alone (Figure 4B). As expected, ICP-AES analysis revealed the loss of both Mn^{2+} ions by most of the mutants, and the loss of one ion in H102A and H207A AptB (Table S8). Therefore, chelation of *both* Mn^{2+} ions are essential for the activities of AptB, which is in contrast to the well-studied dimanganese arginase in which loss of one metal binding site leads to partial loss of enzyme activity.³⁸

DISCUSSION

In this work, we have successfully utilized the PT phylogenetic classification to locate the gene cluster for the biosynthesis of the fungal naphthacenedione **1**. The connection between the *ada* gene cluster and biosynthesis of **1** was confirmed by overexpressing the pathway-

specific transcriptional regulator AdaR in *A. niger*. The gene cluster of **1**, including the NRPKS, is highly similar to those synthesize anthracenones. Subsequent parallel investigations with *ada* and *apt* core PKS enzymes in *S. cerevisiae* led to the identification of a group of bifunctional M β L-TEs that can catalyze both hydrolysis and Claisen cyclization.

AptB is a Dimanganese Enzyme that Catalyzes C-C Bond formation

The M β L enzymes belong to the broader family of hydrolytic metalloenzymes.³⁹ The hydrolytic activity of M β Ls is both dependent on the electrostatic activation of the proton on the metal-coordinated water molecule and the consequent formation of the hydroxide ion (OH⁻).³⁷ Zn²⁺ as the hardest Lewis acid among the first row of transition metals, is most utilized among the M β Ls.⁴⁰ Our investigation suggests that AptB is different from most M β Ls and does not bind Zn²⁺. Instead, AptB exhibits preferential catalytic and physiological requirements for two Mn²⁺ cations. A number of hydrolytic dimanganese enzymes have been reported,⁴¹ including the best characterized arginase,⁴² inorganic pyrophosphatase⁴³ and aminopeptidase.⁴⁴ AptB is the first example among the binuclear manganese enzymes that can catalyze C-C bond formation through Claisen-like condensation. Based on sequence alignment and homology modeling to other M β L-TEs, each Mn²⁺ cation in AptB is likely penta-coordinated in a trigonal bipyramidal geometry (Figure 4A). Mutating any of the metal-coordinating residues to alanine (H97A, H99A, D101A, H102A, H153A, D172A and H207A) led to the loss of both the Claisen cyclization and hydrolysis activities.

Proposed Mechanism of the Hydrolysis Reaction

AptB can facilitate the turnover of naphthacenedione **7** more than 20 times faster than towards anthracenone **8** or **9** (Figure 3B–D). This was surprising considering AptB is naturally recruited by the *apt* pathway to perform the hydrolytic release of **8a**. This suggests that the active site environment of AptB, including positioning of the two Mn²⁺ cations, is optimally setup to perform the Claisen-like condensation, while the hydrolysis reaction occurs when the thioester is placed in close proximity to a strong hydroxide ion. Hence we could expect the hydrolytic reaction to be nonspecific towards substrates that have similar structure and hydrophobicity. Indeed, when AptB was added to the *in vitro* reaction containing the lovastatin nonaketide synthase (LNKS) LovB and the enoylreductase LovC, AptB was also able to catalyze the hydrolytic release and turnover of the otherwise stalled dihydromonacolin (data not shown).⁴⁵ Other types of cyclization TE domains in PKSs, including those that have α/β hydrolase folds, have also been noted to display substantial substrate promiscuity in the loading and hydrolysis of nonnative acyl-thioesters.^{13a,46}

The hydrolytic mechanism of AptB and AdaB likely resembles that of other dimanganese metalloenzymes such as arginase (Scheme 4).^{41c,42} For example, the dimanganese stabilizes the highly reactive hydroxide ion under physiological pH, which serves as the nucleophile in attacking the C1 thioester of the tricyclic polyketide intermediates to yield **9a** and **10a**. The dependence of hydrolytic activity on both Mn²⁺ suggests that the Mn²⁺ ions may play additional roles in catalysis not seen with those in arginase, in which the catalytic center is the metal-ion loaded in the M₁ site and the enzyme is active with a mononuclear center.³⁸ Here, one of the Mn²⁺ ions in AptB may play a crucial role in stabilizing the oxyanion generated in the tetrahedral intermediate. Alternatively, in a model similar to that proposed by Dismukes for arginase,⁴⁷ one of the Mn²⁺ may be involved in substrate binding through coordination to the C1 (or even C3) carbonyl oxygen (as well as stabilizing the oxyanion hole), whereas the other Mn²⁺ chelates a water molecule. During catalysis, a general base such as a nearby histidine can deprotonate the water to initiate the hydrolysis reaction.

Proposed Mechanism of the α -Hydroxylation Dependent Claisen Cyclization

The comparatively weak hydrolytic activity of AptB suggests that the Claisen cyclization is its true function. Our studies here reveal that the cyclization of **7** requires C2-hydroxylation by the accompanying of FMO (AdaC or AptC). If the decaketide substrate is not C2-hydroxylated, the corresponding cyclization product **11** is not observed. Conversely, upon C2-hydroxylation, no hydrolysis product such as **12a** (and the decarboxylated product **12**) can be detected (Scheme 3). Therefore, the α -hydroxylation step essentially functions as a nucleophile switch that modulates the dual functions of AptB. Insertion of the tertiary hydroxyl group α to the C1 thioester carbonyl can have profound effects in substrate reactivity and conformation. This is evidenced in the slow but detectable spontaneous cyclization of **7** in the absence of AptB when C2 is hydroxylated. In contrast, the hydrolysis of **8a** cannot be observed without AptB. The C2-OH can disrupt the keto-enol tautomerization across C1 and C3, thereby increasing the electrophilicity of C1 for nucleophilic attack. Additionally, the C2-hydroxyl group may hydrogen bond to the C1-carbonyl and may provide stabilizing effects to the tetrahedral intermediate. However, these electronic effects should be observed in both hydrolysis and Claisen cyclization reactions, and therefore cannot be used solely to explain the “nucleophile switching” role of this modification.

Although the absolute stereochemistries of **7** and **1** are unknown, the C2-OH is mostly like in the (*R*) configuration and is *syn* to the C15-OH as seen in the crystal structure of **4**.⁴⁸ The *syn* configuration may lead to rearrangement of the flexible C17-C19 diketone moiety, such as through hydrogen bonding, and position the C18 nucleophile in a more favorable position for catalysis. Furthermore, this stereochemical configuration can lead to the possible positioning of a cluster of oxygen atoms, including those on C1, C2, C15, C17 and C19 on the same face of the substrate and are all within suitable distances for chelating both Mn^{2+} ions when bound to AptB. Such chelation may significantly increase the binding affinity of the substrate in the active site. For example, one of the Mn^{2+} ions can chelate to both C2-OH and C1 carbonyl to yield a five-membered-ring (Scheme 4). Such kind of penta-ring structure is not unprecedented, and has been seen in dihydroxyacetone phosphate chelating a Zn^{2+} cation in *E. coli* class II fructose-1,6-bisphosphate aldolase.⁴⁹ In this mode, the water molecule only chelates to the other Mn^{2+} ion and sufficient rearrangement of the active site may prevent deprotonation required for hydrolysis. The general base is instead positioned more favorably to deprotonate the C18 α -carbon.

Alternatively, in addition to the five-membered ring formed through coordination to the first Mn^{2+} , there is a possibility that the C17-C19 diketone moiety can be coordinated to the second Mn^{2+} ion following substrate conformation changes induced by C2-hydroxylation (Scheme 4). This indeed represents the most favorable model for the Claisen cyclization, and is similar to that seen with aldolases.⁵⁰ The coordination of C17 and C19 carbonyls strongly favors the formation of the enol and deprotonation of C18 without the requirement of a general base. The displacement of the hydroxide or water ligand on the second Mn^{2+} should also completely suppress the hydrolysis reaction. Additional biophysical and structure studies on this novel Claisen cyclase is needed to understand the exact modes of substrate binding and metal coordination.

The Bifunctional M β L-TEs are different from the TE/CLC domains

The α -hydroxylation-dependent Claisen cyclization strategy employed by AptB and AdaB is different from the mechanism used by TE/CLC domain found in many NRPKSs, such as those involved in the biosynthesis of bikaverin,⁵¹ naphthopyrone^{13b} and aflatoxin.^{13a} During the Claisen-like cyclization catalyzed by TE/CLC, the active site serine forms an oxyester bond with C1 of the polyketide substrate, releasing it from the

phosphopantetheinyl-arm of the ACP. Formation of the covalent intermediate was suggested to reorient the nucleophilic α -carbon with respect to the Ser-His-Asp catalytic triad and enables formation of the enol.^{13a} The reaction then proceeds via deprotonation of the α -carbon and nucleophilic attack on C1. From the vastly different reaction mechanisms, it is not surprising that fungal TE/CLC and M β L-TE share no sequence similarity. Sequence alignment of AptB, AdaB and VrtG does indicate a conserved serine residue (S191 for AptB, Figure S15), however, mutation to alanine in AptB did not affect the efficiency of the Claisen cyclization (Figure 4B). Therefore, M β L-TEs employ a completely different mechanism to enolize the α -carbon, and no covalent bond is formed between the polyketide intermediate and the M β L-TEs during the cyclization. Furthermore, in the biosynthesis of fungal aromatic polyketides requiring TE/CLC domains, the C1 carbonyl is usually directly connected to an aromatic ring due to the C2-C7 or C4-C9/C2-C11 cyclization steps catalyzed by PT, and the hyperconjugation between the aromatic ring and C1 carbonyl reduces the reactivity of C1. Therefore, the covalently bound enzyme-substrate intermediate is required to promote the cyclization reaction in TE/CLC.

In contrast, bacterial type II PKSs infrequently perform the final ring closure through Claisen cyclization, except in the biosynthesis of tetracyclines and aureolic acids. Anhydrotetracycline⁵² and premithramycinone,⁵³ the key intermediates in tetracycline and aureolic acid biosynthesis, respectively, share similar naphthacenedione scaffolds. Correspondingly, the biosynthesis steps also involve a C18-C1 Claisen-like ring closure. The electronic environment of the α -carbon is more similar to substrates of the TE/CLC cyclized reactions, in that the C1 carbonyl is directly connected to and conjugated with the aromatic portions of the molecules. Therefore, although the exact catalytic mechanism of the C18-C1 fourth ring cyclization in bacteria has not been completely established, most likely it will be different from the fungal strategy described here.

Phylogenetic Analysis of M β L-TEs

During phylogenetic analysis of PT domains, the Group V NRPKSs were branched into two major subgroups. Interestingly, their corresponding M β L-TEs can also be classified into two subgroups (Figure S17) with similar phylogenetic distribution to their paired NRPKSs, indicating that the NRPKS/M β L-TE may be coevolved. The first subgroup (Group V1) contains ACTE and AN0149, while the second subgroup (Group V2) contains AdaB, AptB and VrtG. AN0149 has been linked to the biosynthesis of the anthraquinone emodin, which is likely synthesized via an anthracenone intermediate as in the biosynthesis of **5** involving ACTE.¹⁶ Hence, the classification of M β L-TE is largely based on the sizes of the polyketide substrates. The fourth ring cyclization mechanism of Group V2 M β L-TEs should be common to the biosynthesis of fungal tetracyclic naphthacenedione compounds, such as **4**, hypomycesin and anthrotainin. To investigate if the Claisen cyclization activity is universal among both groups of M β L-TEs, we also purified and assayed ACTE with AdaA and AdaC. However, formation of **7** was not accelerated, which indicates that the Claisen cyclization activity is unique among the Group V2 M β L-TEs.

The presence of several uncharacterized PKS clusters with the Group V2 NRPKS/M β L-TE/FMO three-gene ensemble in other sequenced fungal genomes (Figure S6) indicates that there are potentially other undiscovered fungal tetracyclic compounds. Based on the presence of FMOs in all analyzed Group V2 PKS clusters and their absence in all Group V1 PKS clusters, it is likely that the Group V2 M β L-TEs were diverged from the hydrolytic Group V1 M β L-TEs and acquired the Claisen cyclase function during the recruitment of FMO into these pathways. The fact that AptB exhibits higher efficiency in catalyzing the Claisen cyclization than hydrolysis, and the presence of the unique FMO in the *apt* cluster suggest that AptB may have been part of a tetracyclic pathway, but reverted to its hydrolytic function more recently. This is perhaps due to the loss of ability to produce the longer

decaetide chain by the NRPKS AptA. Most identified enzymes in the M β L superfamily are known to catalyze hydrolytic reaction, such as the well-known β -lactamase that confers penicillin resistance. Therefore, the hydrolytic activity is likely the ancestral role of these newly identified M β L-TEs with fourth ring cyclization activity.

In conclusion, this study sheds light into the mechanism of biosynthesis of anthracenone and naphthacenedione in fungi. Compared to the biosynthesis of naphthacenedione in bacteria, such as premithramycinone and anhydrotetracycline that require over ten enzymes, the fungal strategy is concise and only involves three enzymes. The linear polyhydroxylated tetracyclic compounds have been a promising source of bioactive molecules, which include clinically important antibiotics and antitumor drugs. The establishment of the minimal three-gene ensemble (Group V NRPKS/ M β L-TE/FMO) required for the biosynthesis of a fungal naphthacenedione provides a useful signature feature for discovery of other related compounds from fungi using genome mining. Furthermore, the demonstrated ability to produce tetracyclic naphthacenedione in yeast has opened up new possibilities to engineer these pharmacologically attractive tetracyclic scaffolds.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Prof. Neil Garg, Dr. Stuart Haynes, Dr. Hui Zhou, Dr. Jaclyn Winter, Xue Gao and Wei Xu for their helpful discussions. This work was supported by NIH Grants 1R01GM085128 and 1R01GM092217 to YT.

ABBREVIATIONS

ATP	adenosine-5'-triphosphate
CoA	coenzyme A
FAD	flavin adenine dinucleotide
NADPH	reduced nicotinamide adenine dinucleotide phosphate
MβL-TEs	metallo- β -lactamase-type thioesterase
NRPKS	non-reducing polyketide synthase
SAT	starter unit:ACP transacylase
KS	ketosynthase
MAT	malonyl-CoA:ACP transacylase
PT	product template
ACP	acyl carrier protein
ICP-AES	inductively coupled plasma atomic emission spectrometer
SOE	splice by overlapping extension
RT-PCR	reverse transcription polymerase chain reaction
EtOAc	ethyl acetate
MeOH	methanol
CHCl₃	chloroform

CH₃CN	acetonitrile
AcOH	acetic acid
LC-MS	liquid chromatography mass spectrometry
IPTG	isopropylthio- β -D-galactoside

REFERENCES

- O'Hagan, D. *The Polyketide Metabolites*. Chichester, UK: Ellis Howard; 1991.
- (a) Das A, Khosla C. *Acc. Chem. Res.* 2009; 42:631. [PubMed: 19292437] (b) Hertweck C, Luzhetskyy A, Rebets Y, Bechthold A. *Nat. Prod. Rep.* 2007; 24:162. [PubMed: 17268612] (c) Zhou H, Li Y, Tang Y. *Nat. Prod. Rep.* 2010; 27:839. [PubMed: 20358042]
- Cox RJ. *Org. Biomol. Chem.* 2007; 5:2010. [PubMed: 17581644]
- Crawford JM, Thomas PM, Scheerer JR, Vagstad AL, Kelleher NL, Townsend CA. *Science*. 2008; 320:243. [PubMed: 18403714]
- Harris CM, Roberson JS, Harris TM. *J. Am. Chem. Soc.* 1976; 98:5380. [PubMed: 956563]
- Watanabe A, Fujii I, Sankawa U, Mayorga ME, Timberlake WE, Ebizuka Y. *Tetrahedron Lett.* 1999; 40:91.
- (a) Hutchison RD, Steyn PS, Van Rensburg SJ. *Toxicol. Appl. Pharmacol.* 1973; 24:507. [PubMed: 4122267] (b) Jesus A, Hull W, Steyn P, Heerden F, Vleggaar R. *J. Chem. Soc., Chem. Commun.* 1982; 1982:902. (c) Zheng C-J, Yu H-E, Kim E-H, Kim W-G. *J. Antibiot.* 2008; 61:633. [PubMed: 19168978]
- Crawford JM, Townsend CA. *Nat. Rev. Microbiol.* 2010; 8:879. [PubMed: 21079635]
- Crawford JM, Dancy BCR, Hill EA, Udway DW, Townsend CA. *Proc. Natl. Acad. Sci. U. S. A.* 2006; 103:16728. [PubMed: 17071746]
- (a) Zhang WJ, Li YR, Tang Y. *Proc. Natl. Acad. Sci. U. S. A.* 2008; 105:20683. [PubMed: 19075227] (b) Crawford JM, Korman TP, Labonte JW, Vagstad AL, Hill EA, Kamari-Bidkorpheh O, Tsai S-C, Townsend CA. *Nature*. 2009; 461:1139. [PubMed: 19847268]
- Wattana-amorn P, Williams C, Ploskon E, Cox RJ, Simpson TJ, Crosby J, Crump MP. *Biochemistry*. 2010; 49:2186. [PubMed: 20136099]
- Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon BG. *Proc. Natl. Acad. Sci. U. S. A.* 2003; 100:15670. [PubMed: 14676319]
- (a) Korman TP, Crawford JM, Labonte JW, Newman AG, Wong J, Townsend CA, Tsai SC. *Proc. Natl. Acad. Sci. U. S. A.* 2010; 107:6246. [PubMed: 20332208] (b) Fujii I, Watanabe A, Sankawa U, Ebizuka Y. *Chem. Biol.* 2001; 8:189. [PubMed: 11251292]
- Bailey AM, Cox RJ, Harley K, Lazarus CM, Simpson TJ, Skellam E. *Chem. Comm.* 2007:4053. [PubMed: 17912413]
- Li Y, Xu W, Tang Y. *J. Biol. Chem.* 2010; 285:22764. [PubMed: 20479000]
- Awakawa T, Yokota K, Funa N, Doi F, Mori N, Watanabe H, Horinouchi S. *Chem. Biol.* 2009; 16:613. [PubMed: 19549600]
- Szewczyk E, Chiang YM, Oakley CE, Davidson AD, Wang CC, Oakley BR. *Appl. Environ. Microbiol.* 2008; 74:7607. [PubMed: 18978088]
- Chooi Y-H, Cacho R, Tang Y. *Chem. Biol.* 2010; 17:483. [PubMed: 20534346]
- Ishimaru, T.; Tsuboya, S.; Saijo, T. *Japanese Patent*. JP0556861A. 1994.
- (a) Kodukula K, Arcuri M, Cutrone JQ, Hugill RM, Lowe SE, Pirnik DM, Shu YZ, Fernandes PB, Seethala R. *J. Antibiot.* 1995; 48:1055. [PubMed: 7490206] (b) Shu YZ, Cutrone JFQ, Klohr SE, Huang S. *J. Antibiot.* 1995; 48:1060. [PubMed: 7490207]
- Pfeifer BA, Admiraal SJ, Gramajo H, Cane DE, Khosla C. *Science*. 2001; 291:1790. [PubMed: 11230695]
- (a) Jones EW. *Methods Enzymol.* 1991; 194:428. [PubMed: 2005802] (b) Lee KK, Da Silva NA, Kealey JT. *Anal. Biochem.* 2009; 394:75. [PubMed: 19595983]

23. Pall ML, Brunelli JP. *Fungal. Genet. Newsl.* 1993; 40:59.
24. Andrianopoulos A, Hynes MJ. *Mol. Cell. Biol.* 1988; 8:3532. [PubMed: 3062382]
25. Chooi Y-H, Stalker DM, Davis MA, Fujii I, Elix JA, Louwhoff SHJJ, Lawrie AC. *Mycol. Res.* 2008; 112:147. [PubMed: 18280724]
26. Mutka SC, Bondi SM, Carney JR, Da Silva NA, Kealey JT. *FEMS Yeast Res.* 2006; 6:40. [PubMed: 16423069]
27. Bradford MM. *Anal. Biochem.* 1976; 72:248. [PubMed: 942051]
28. An JH, Kim YS. *Eur. J. Biochem.* 1998; 257:395. [PubMed: 9826185]
29. Pel HJ, et al. *Nat. Biotechnol.* 2007; 25:221. [PubMed: 17259976]
30. (a) Chiang Y-M, Szewczyk E, Davidson AD, Keller N, Oakley BR, Wang CCC. *J. Am. Chem. Soc.* 2009; 131:2965. [PubMed: 19199437] (b) Bergmann S, Schumann J, Scherlach K, Lange C, Brakhage AA, Hertweck C. *Nat. Chem. Biol.* 2007; 3:213. [PubMed: 17369821]
31. Thompson CJ, Movva NR, Tizard R, Cramer R, Davies JE, Lauwereys M, Botterman J. *EMBO J.* 1987; 6:2519. [PubMed: 16453790]
32. Chen Y, Fan K, He Y, Xu X, Peng Y, Yu T, Jia C, Yang K. *ChemBioChem.* 2010; 11:1055. [PubMed: 20422670]
33. (a) Garau G, Garcia-Saez I, Bebrone C, Anne C, Mercuri P, Galleni M, Frere JM, Dideberg O. *Antimicrob. Agents Chemother.* 2004; 48:2347. [PubMed: 15215079] (b) Bush K, Jacoby GA. *Antimicrob. Agents Chemother.* 2010; 54:969. [PubMed: 19995920]
34. (a) Roy A, Kucukural A, Zhang Y. *Nat. Protoc.* 2010; 5:725. [PubMed: 20360767] (b) Zhang Y. *BMC Bioinf.* 2008; 9:40.
35. Shimada A, Ishikawa H, Nakagawa N, Kuramitsu S, Masui R. *Proteins.* 2010; 78:2399. [PubMed: 20544975]
36. Stamp AL, Owen P, El Omari K, Nichols CE, Lockyer M, Lamb HK, Charles IG, Hawkins AR, Stammers DK. *Protein Sci.* 2010; 19:1897. [PubMed: 20669241]
37. (a) Wang Z, Fast W, Valentine AM, Benkovic SJ. *Curr. Opin. Chem. Biol.* 1999; 3:614. [PubMed: 10508665] (b) Xu D, Xie D, Guo H. *J. Biol. Chem.* 2006; 281:8740. [PubMed: 16423823]
38. (a) López V, Alarcón R, Orellana MS, Enríquez P, Uribe E, Martínez J, Carvajal N. *FEBS J.* 2005; 272:4540. [PubMed: 16128822] (b) Orellana MS, López V, Uribe E, Fuentes M, Salas M, Carvajal N. *Arch. Biochem. Biophys.* 2002; 403:155. [PubMed: 12139964] (c) Scolnick LR, Kanyo ZF, Cavalli RC, Ash DE, Christianson DW. *Biochemistry.* 1997; 36:10558. [PubMed: 9265637]
39. Daiyasu H, Osaka K, Ishino Y, Toh H. *FEBS Lett.* 2001; 503:1. [PubMed: 11513844]
40. Bebrone C. *Biochem. Pharmacol.* 2007; 74:1686. [PubMed: 17597585]
41. (a) Christianson DW. *Prog. Biophys. Mol. Biol.* 1997; 67:217. [PubMed: 9446936] (b) Dismukes GC. *Chem. Rev.* 1996; 96:2909. [PubMed: 11848845] (c) Christianson DW, Cox JD. *Annu. Rev. Biochem.* 1999; 68:33. [PubMed: 10872443]
42. Kanyo ZF, Scolnick LR, Ash DE, Christianson DW. *Nature.* 1996; 383:554. [PubMed: 8849731]
43. Heikinheimo P, Lehtonen J, Baykov A, Lahti R, Cooperman BS, Goldman A. *Structure.* 1996; 4:1491. [PubMed: 8994974]
44. Wilce MC, Bond CS, Dixon NE, Freeman HC, Guss JM, Lilley PE, Wilce JA. *Proc. Natl. Acad. Sci. U. S. A.* 1998; 95:3472. [PubMed: 9520390]
45. Ma SM, Li JW-H, Choi JW, Zhou H, Lee KKM, Moorthis VA, Xie X, Kealey JT, Da Silva NA, Vederas JC, Tang Y. *Science.* 2009; 326:589. [PubMed: 19900898]
46. Sharma KK, Boddy CN. *Bioorg. Med. Chem. Lett.* 2007; 17:3034. [PubMed: 17428661]
47. Khangulov SV, Sossong TM Jr, Ash DE, Dismukes GC. *Biochemistry.* 1998; 37:8539. [PubMed: 9622506]
48. Kabuto C, Silverton JV, Akiyama T, Sankawa U, Hutchison RD, Steyn PS, Vleggaar R. *J. Chem. Soc., Chem. Commun.* 1976:728.
49. (a) Hall DR, Leonard GA, Reed CD, Watt CI, Berry A, Hunter WN. *J. Mol. Biol.* 1999; 287:383. [PubMed: 10080900] (b) Galkin A, Li Z, Li L, Kulakova L, Pal LR, Dunaway-Mariano D, Herzberg O. *Biochemistry.* 2009; 48:3186. [PubMed: 19236002]
50. Lipscomb WN, Strater N. *Chem. Rev.* 1996; 96:2375. [PubMed: 11848831]

51. Ma SM, Zhan J, Xie X, Watanabe K, Tang Y, Zhang W. *J. Am. Chem. Soc.* 2008; 130:38. [PubMed: 18072779]
52. Pickens LB, Tang Y. *J. Biol. Chem.* 2010; 285:27509. [PubMed: 20522541]
53. Lombo F, Menendez N, Salas JA, Mendez C. *Appl. Microbiol. Biotechnol.* 2006; 73:1. [PubMed: 17013601]

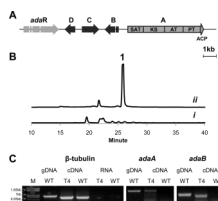


Figure 1. Induction of the *ada* gene cluster. (A) Organization of the *ada* gene cluster. (B) LC analysis (400 nm) of *A. niger* extracts: i. WT and ii. T4 transformant; (C) RT-PCR analysis of the *ada* genes with lane M as marker. The β -tubulin was used as a positive control and RNA only of both WT and T4 was used as negative controls. Regions of *adaA* and *adaB* were probed for induced transcription.

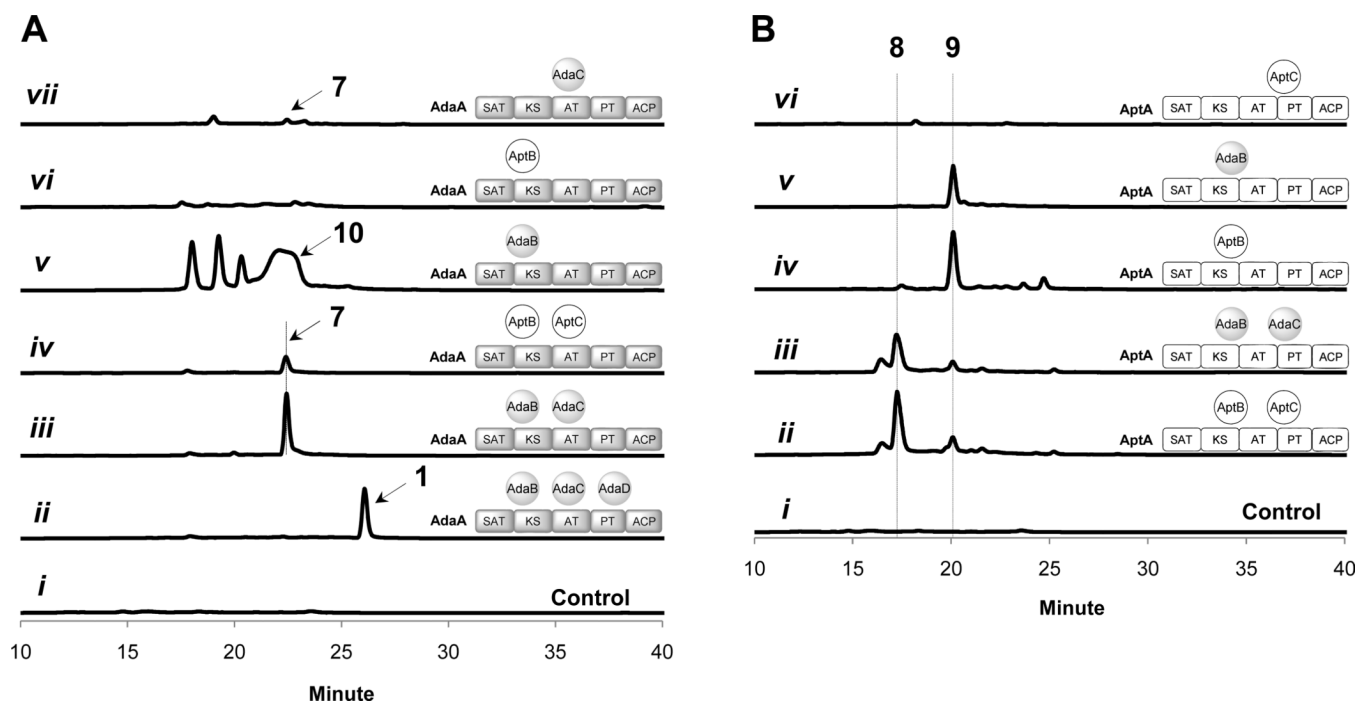


Figure 2. LC analysis (400 nm) of polyketides synthesized by coexpression of different combinations of *ada* and *apt* genes in *S. cerevisiae* BJ5464-NpgA. (A) AdaA coexpressed with different combinations of MβL-TEs and/or FMOs; (B) AptA coexpressed with different combinations of MβL-TEs and/or FMOs. Control trace is extracted from untransformed BJ5464-NpgA culture.

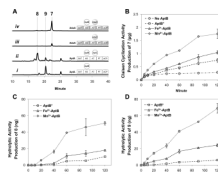


Figure 3.

In vitro reconstitution of AptB activities. **(A)** LC analysis (400 nm) of polyketide synthesized from different combinations of purified enzymes; **(B–D)**. Time-course analysis of production of **7**, **8**, or **9** over 120 minutes using AptB containing different levels of metal cofactors (AptB*, Fe²⁺-AptB and Mn²⁺-AptB) coincubated with AdaA and AdaC **(B)**, AptA and AptC **(C)**, and AptA only **(D)**. No hydrolysis can be observed in the absence of AptB.

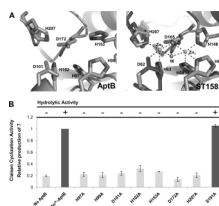
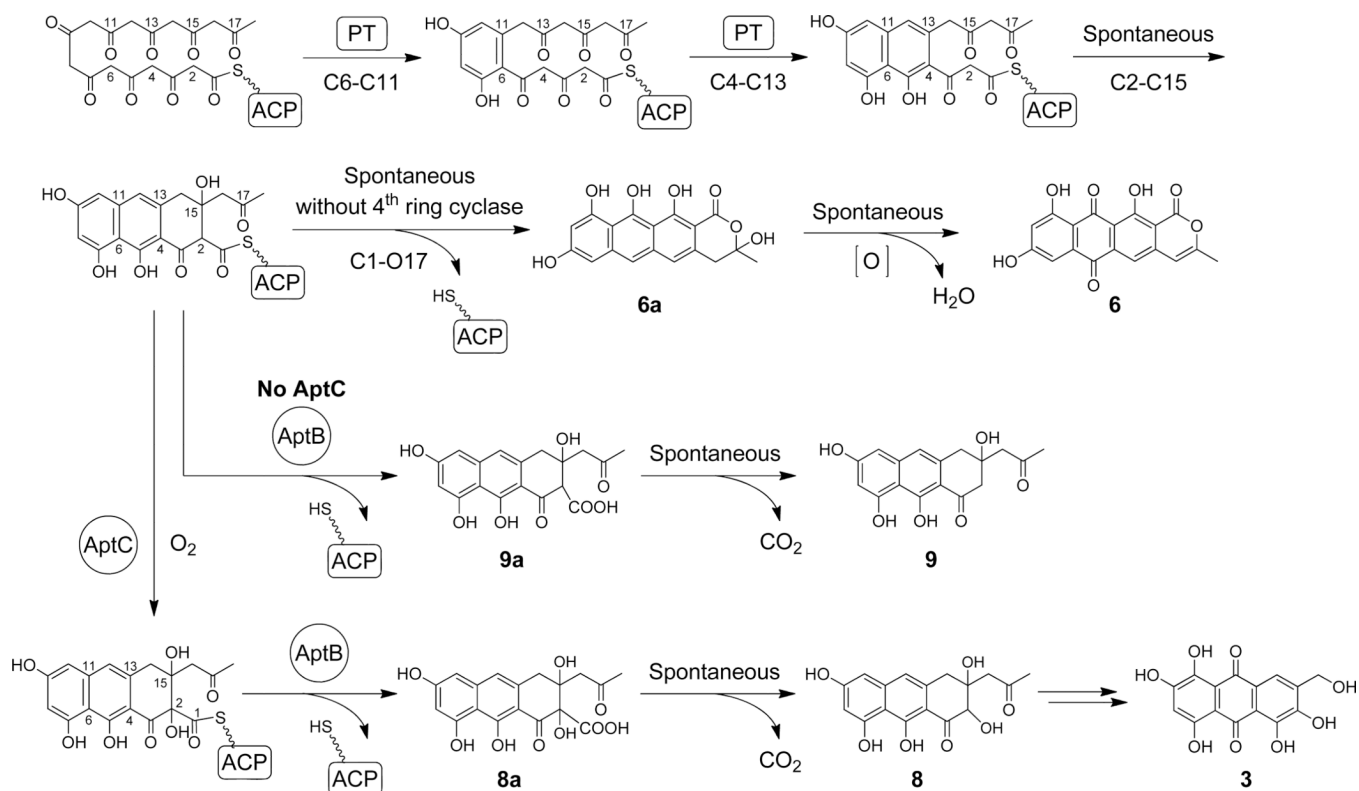


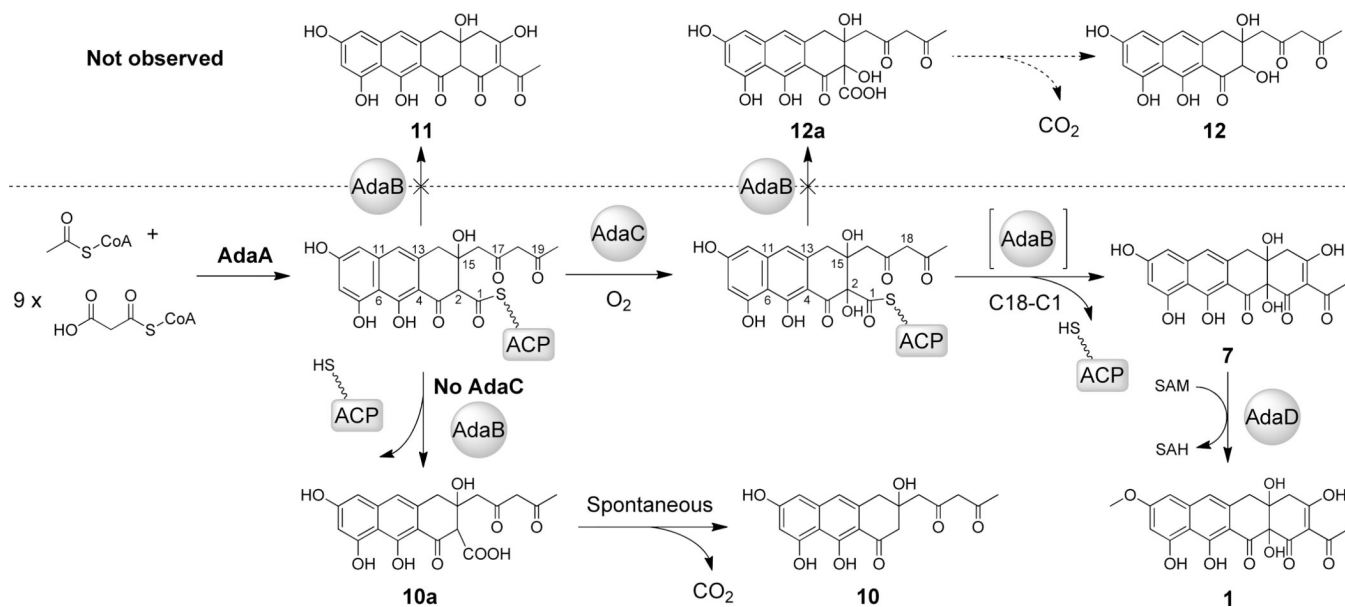
Figure 4. Structural modeling and mutational analysis of AptB. **(A)** Comparison between the modeled metal binding sites of AptB (left) predicted by the I-TASSER server and the zinc binding site of a MβL superfamily protein ST1585 (Chain B) (adapted from ref 35) (right). **(B)**. Site directed mutagenesis of AptB to probe the hydrolysis (formation of **9**) and Claisen condensation (formation of **7**) activities. The level of production of **7** in the wild-type Mn²⁺-AptB assay is defined as 1 for comparison of the catalytic activity of the Mn²⁺-AptB mutants.



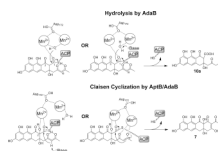
Scheme 1.
Anthracenone and Naphthacenedione natural products isolated from fungi.



Scheme 2.
Biosynthetic pathway of 3 and related shunt products.



Scheme 3.
Biosynthetic pathway of 1 and related shunt products.



Scheme 4.
Proposed enzymatic mechanism of MβL-TE.

Table 1

Genes Within the *ada* Cluster From *Aspergillus niger* ATCC 1015¹

Gene	Size (bp/aa)	<i>A. niger</i> CBS 513.88 (% identity)	BLASTP homolog	Identity/Similarity (%)	E-value	Conserved Domain
<i>adaA</i>	5382/1793	An11g07310 (99)	<i>aptA</i> <i>vrrA</i>	69/80	0.0	NRPKS, SAT-KS-MAT-PT-ACP
				59/74	0.0	
<i>adaB</i>	1005/317	An11g07320 (100)	<i>aptB</i> <i>vrrG</i>	69/79	1e-177	metallo- β -lactamase superfamily
				62/76	1e-107	
<i>adaC</i>	1245/414	An11g07320 (100)	<i>aptC</i> <i>vrrH</i>	63/79	8e-153	UbiH and related FAD-dependent oxidoreductases
				63/78	2e-153	
<i>adaD</i>	750/249	An11g07340 (99)	<i>vrrF</i>	60/71	2e-68	methyltransferase domain type12
<i>adaR</i>	2504/777	An11g07350 (100)	<i>vrrR2</i>	42/55	1e-128	GAL4-like Zn(II) ₂ Cys ₆ binuclear cluster DNA-binding domain

¹The sequence of the *ada* gene cluster in *A. niger* ATCC 1015 was deposited in GenBank with accession number JN257714.

Table 2Heterologous expression of different combinations of *apt* and *ada* genes and the resulting products¹.

NRPKS	MβL-TE	FMO	Product
<i>adaA</i>	<i>adaB</i>	<i>adaC</i>	7
<i>aptA</i>	<i>aptB</i>	<i>aptC</i>	8
<i>adaA</i>	<i>aptB</i>	<i>aptC</i>	7
<i>aptA</i>	<i>adaB</i>	<i>adaC</i>	8
<i>aptA</i>	<i>aptB</i>	-	9
<i>adaA</i>	<i>adaB</i>	-	10
<i>adaA</i>	<i>aptB</i>	-	<i>nd</i> ²
<i>aptA</i>	<i>adaB</i>	-	9
<i>adaA</i>	-	<i>adaC</i>	7
<i>aptA</i>	-	<i>aptC</i>	<i>nd</i>

¹ *S. cerevisiae* BJ5464-NpgA is the host;

² *nd* no polyketide metabolite was detected from this combination of genes.