ORGANIC LETTERS

2013
Vol. 15, No. 4
756–759

Engineering Fungal Nonreducing Polyketide Synthase by Heterologous Expression and Domain Swapping

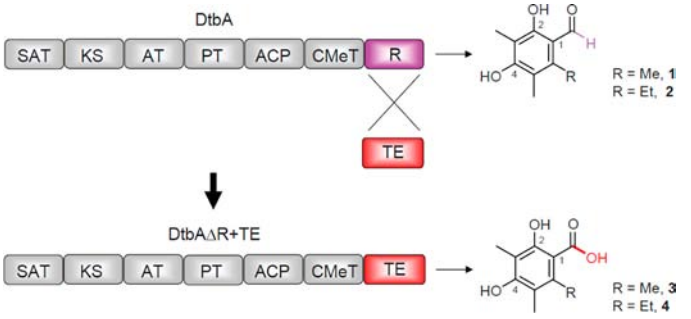
Hsu-Hua Yeh,[†,∥,▽] Shu-Lin Chang,[‡,∥,▽] Yi-Ming Chiang,[§,∥,▽] Kenneth S. Bruno,[⊥] Berl R. Oakley,[○] Tung-Kung Wu,[*,†] and Clay C. C. Wang[*,∥,#]

Department of Biological Science and Technology, National Chiao Tung University, Hsinchu City 30010, Taiwan, Department of Biotechnology, Chia Nan University of Pharmacy and Science, Tainan City 71710, Taiwan, Graduate Institute of Pharmaceutical Science, Chia Nan University of Pharmacy and Science, Tainan City 71710, Taiwan, Chemical and Biological Process Development Group, Energy and Environment Directorate, Pacific Northwest National Laboratory, Richland, Washington 99352, United States, Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66045, United States, Department of Pharmacology and Pharmaceutical Sciences, University of Southern California, School of Pharmacy, Los Angeles, California 90089, United States, and Department of Chemistry, University of Southern California, Dornsife College of Letters, Arts, and Sciences, Los Angeles, California 90089, United States

tkwmll@mail.nctu.edu.tw; clayw@usc.edu

Received December 5, 2012

ABSTRACT



We reannotated the *A. niger* NR-PKS gene, e_gw1_19.204, and its downstream R domain gene, est_GWPlus_C_190476, as a single gene which we named *dtbA*. Heterologous expression of *dtbA* in *A. nidulans* demonstrated that DtbA protein produces two polyketides, 2,4-dihydroxy-3,5,6-trimethylbenzaldehyde (1) and 6-ethyl-2,4-dihydroxy-3,5-dimethylbenzaldehyde (2). Generation of DtbAΔR+TE chimeric PKSs by swapping the DtbA R domain with the AusA (austinol biosynthesis) or ANID_06448 TE domain enabled the production of two metabolites with carboxylic acids replacing the corresponding aldehydes.

Filamentous fungi produce numerous polyketides that exhibit a wide range of bioactivities. Fungal polyketides are synthesized by multidomain polyketide synthases (PKSs) which can be classified as nonreducing (NR), partially reducing (PR), and highly reducing (HR) PKSs.[1−5] Canonical NR-PKSs catalyze the production of aromatic

[†] National Chiao Tung University.
[‡] Department of Biotechnology, Chia Nan University of Pharmacy and Science.
[§] Graduate Institute of Pharmaceutical Science, Chia Nan University of Pharmacy and Science.
[∥] Department of Pharmacology and Pharmaceutical Sciences, University of Southern California.
[⊥] Pacific Northwest National Laboratory.
[○] University of Kansas.
[#] Department of Chemistry, University of Southern California.
[▽] These authors contributed equally.

(1) Nicholson, T. P.; Rudd, B. A.; Dawson, M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. *Chem. Biol.* **2001**, *8*, 157–178.
(2) Kroken, S.; Glass, N. L.; Taylor, J. W.; Yoder, O. C.; Turgeon, B. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15670–15675.
(3) Cox, R. J. *Org. Biomol. Chem.* **2007**, *5*, 2010–2026.
(4) Chiang, Y. M.; Oakley, B. R.; Keller, N. P.; Wang, C. C. *Appl. Microbiol. Biotechnol.* **2010**, *86*, 1719–1736.
(5) Chooi, Y. H.; Tang, Y. *J. Org. Chem.* **2012**, *77*, 9933–9953.

10.1021/ol303328t
Published on Web 01/31/2013

polyketides and contain starter unit ACP transacylase (SAT), β-ketoacyl synthase (KS), acyl transferase (AT), product template (PT), and acyl carrier protein (ACP) domains at a minimum, and sometimes contain a methyltransferase (CMeT) domain as well.[6−8] Release of the polyketide chain from an NR-PKS can be mediated by several different types of domains found in NR-PKSs including thioesterase/Claisen-cyclase (TE/CLC), thioesterase (TE), and reductase (R) domains.[9−11] An additional discrete β-lactamase can also release the polyketide from NR-PKSs lacking a TE/CLC domain.[12] NR-PKS genes are common and widespread among *Aspergillus* species. Using a variety of genome mining methods, we and others have successfully identified the products of all 14 NR-PKSs of *A. nidulans*. As an extension to our study in *A. nidulans* we have examined the eight species in the Broad Institute *Aspergillus* Comparative Database (http://www.broadinstitute.org/) and identified at least 71 NR-PKSs that can be classified into seven major groups.[13] We are interested in deciphering the chemical products of these NR-PKSs, especially ones with non-canonical domain architectures. Ultimately, our goal is to understand the relationship between NR-PKS DNA sequences and chemical structures in order to develop bioinformatic tools capable of predicting the chemical products of NR-PKSs from DNA sequences generated by fungal genome projects.

Through bioinformatic analyses of *A. niger* strain ATCC 1015, we identified a putative TE-less NR-PKS encoded by e_gw1_19.204, belonging to Group VI NR-PKSs. A PKS in this group typically has an SAT-KS-AT-PT-ACP-CMeT-TE domain architecture.[13] Sequence analysis indicated that the 3′ coding region of e_gwl_19.204 is incomplete. Interestingly, the gene next to e_gw1_19.204, est_GWPlus_C_190476, encodes a protein with a putative R domain, but both 5′ and 3′ coding regions are missing. These data combined suggested that e_gw1_19.204 and est_GWPlus_C_190476 are both misannotated, and the two fragments are in fact part of a single gene. If this is true, then the fused NR-PKS should have the SAT-KS-AT-PT-ACP-CMeT-R domain architecture, which is unusual in group VI. To verify this hypothesis, we used an *in vivo* heterologous expression approach to determine the PKS product of this fused gene.

First, a PCR fragment ranging from the start site of e_gwl_19.204 to a predicted stop site of est_GWPlus_C_190476 was generated and fused to a selectable marker and an inducible *alcA* promoter cassette (Figure S1).[14] This DNA construct was heterologously expressed in *A. nidulans*. The parent strain used for transformation featured an *nkuA*Δ, *stcJ*Δ genetic background which not only facilitates the transformation success rate but also eliminates the major secondary metabolite sterigmatocystin.[15,16] After induction, total RNA was extracted, and the cDNA was synthesized using the reverse primer specific to the stop site of est_GWPlus_C_190476. The cDNA was then PCR amplified and underwent sequencing analysis. The cDNA sequence differed from the predicted coding sequence (Figure S2). NCBI protein blast analysis (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) of the deduced amino acids revealed that a coding region was incorrectly annotated, thereby splitting the actual R domain from the C terminus of the NR-PKS.

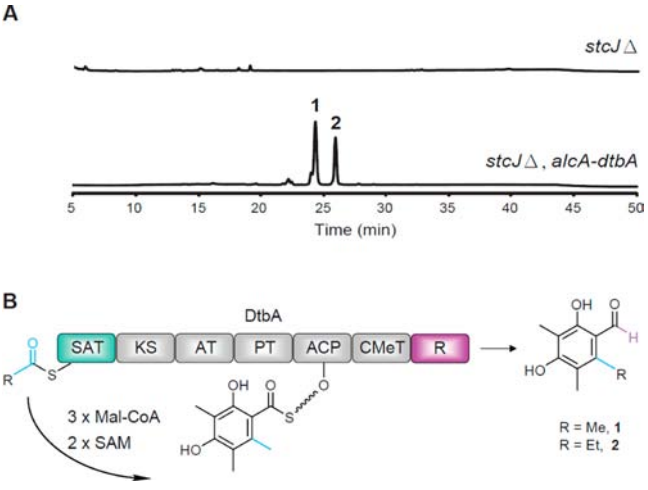


Figure 1. (A) HPLC-DAD-MS analysis of metabolite profiles. (B) Domain architecture, putative starter units, unreleased intermediates, and released polyketides for DtbA.

Next, we focused on determining the product of this reannotated NR-PKS. Upon induction, the heterologous expression strain produced two UV−vis detectable aromatic compounds **1** and **2** (Figure 1). To determine the structure of **1** and **2**, both compounds were isolated from a large-scale culture and structurally elucidated by spectroscopic methods. Compound **2** has been structurally elucidated as 6-ethyl-2,4-dihydroxy-3,5-dimethylbenzaldehyde by our group previously.[13] Compound **1** has a similar

(6) Crawford, J. M.; Dancy, B. C.; Hill, E. A.; Udwary, D. W.; Townsend, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16728–16733.
(7) 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*, 1139–1143.
(8) Fisch, K. M.; Skellam, E.; Ivison, D.; Cox, R. J.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J. *Chem. Commun. (Camb)* **2010**, *46*, 5331–5333.
(9) Fujii, I.; Watanabe, A.; Sankawa, U.; Ebizuka, Y. *Chem. Biol.* **2001**, *8*, 189–197.
(10) Schroeckh, V.; Scherlach, K.; Nutzmann, H. W.; Shelest, E.; Schmidt-Heck, W.; Schuemann, J.; Martin, K.; Hertweck, C.; Brakhage, A. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 14558–14563.
(11) Bailey, A. M.; Cox, R. J.; Harley, K.; Lazarus, C. M.; Simpson, T. J.; Skellam, E. *Chem. Commun. (Camb)* **2007**, 4053–4055.
(12) Awakawa, T.; Yokota, K.; Funa, N.; Doi, F.; Mori, N.; Watanabe, H.; Horinouchi, S. *Chem. Biol.* **2009**, *16*, 613–623.
(13) Ahuja, M.; Chiang, Y. M.; Chang, S. L.; Praseuth, M. B.; Entwistle, R.; Sanchez, J. F.; Lo, H. C.; Yeh, H. H.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2012**, *134*, 8212–8221.
(14) Waring, R. B.; May, G. S.; Morris, N. R. *Gene* **1989**, *79*, 119–130.
(15) Nayak, T.; Szewczyk, E.; Oakley, C. E.; Osmani, A.; Ukil, L.; Murray, S. L.; Hynes, M. J.; Osmani, S. A.; Oakley, B. R. *Genetics* **2006**, *172*, 1557–1566.
(16) Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Keller, N.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2009**, *131*, 2965–2970.

UV−vis spectrum to **2** (Figure S5). The ^{1}H and ^{13}C NMR spectra of **1** exhibited three methyl groups attached to the benzene ring (Figures S6, S7). This together with the fact that **1** is 14 amu less than **2** suggested that **1** is 2,4-dihydroxy-3,5,6-trimethylbenzaldehyde. We designated this intact NR-PKS gene as *dtbA*, from 2,4-dihydroxy-3,5,6-trimethylbenzaldehyde. The result implies that the SAT domain of DtbA is capable of accepting both acetyl-CoA and propionyl-CoA as starter units, which is then extended with three malonyl-CoAs and modified with two S-adenosyl methionines (SAMs) followed by the reductive release.

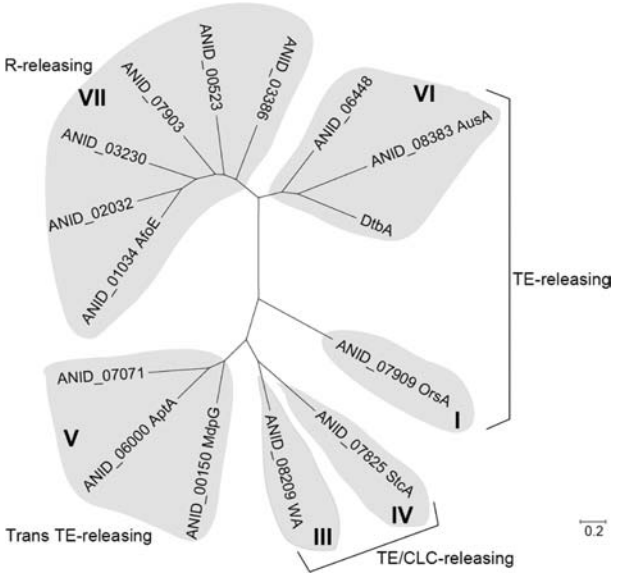


Figure 2. Phylogenetic analysis of DtbA and 14 NR-PKSs in *A. nidulans*.

Phylogenetic analysis revealed this reannotated NR-PKS belongs in group VI NR-PKS, all members of which possess a TE domain but not an R domain for product release (Figure S3). To date, it is still not clear whether a programmed releasing mechanism can be modulated by modifying the native building blocks. We reasoned that if we could engineer this NR-PKS with a TE release domain instead of the R domain, we would be able to further expand the structural diversity of products produced by fungal NR-PKSs. To test this hypothesis, by taking advantage of the fact that all chemical products of NR-PKSs from *A. nidulans* have been solved recently,[13] we generated chimeric modules in which the SAT-KS-AT-PT-ACP-CMeT fragment from DtbA (DtbAΔR) was fused to a TE or TE/CLC domain from other NR-PKSs of *A. nidulans* (Figure S4).

We analyzed the phylogenetic relationship of the recipient NR-PKS DtbA with the donor NR-PKSs from *A. nidulans* (Figure 2). Phylogenetic analysis revealed that ANID_08383 (AusA) and ANID_06448 from *A. nidulans* are positioned at the shortest phylogenetic distance from DtbA and both of them contain a TE domain (Figure 2). Therefore, first, we replaced the R domain in DtbA with the TE from AusA or ANID_06448 of *A. nidulans*, respectively.

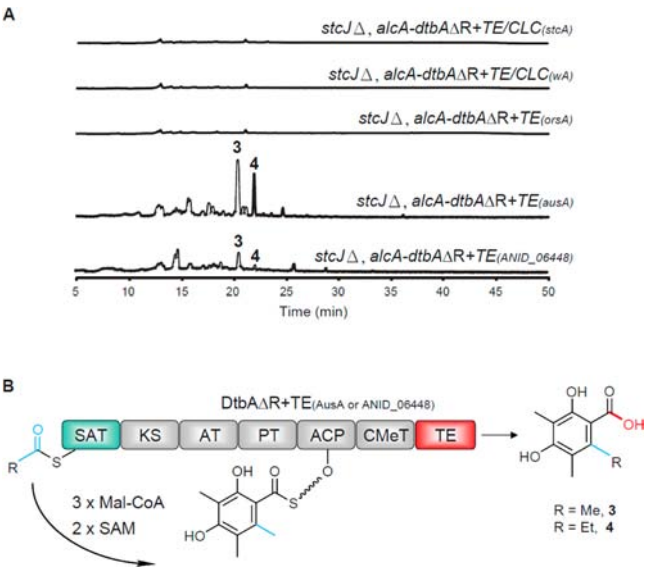


Figure 3. (A) HPLC-DAD-MS analysis of metabolite profiles. (B) Domain architecture, putative starter units, unreleased intermediates, and released polyketides for hybrid DtbAΔR+TE(AusA or ANID_06448).

Heterologous expression of the chimeric PKS, DtbAΔR+TE(AusA) and DtbAΔR+TE(ANID_06448), led to the production of **3** and **4** after 48 h of induction (Figure 3). Compounds **3** and **4** were purified from a large-scale culture for structural elucidation. Spectroscopic data confirmed that **3** is 2,4-dihydroxy-3,5,6-trimethylbenzoic acid, an aromatic polyketide produced by AusA.[13] Compound **4** has a similar UV−vis spectrum to **3** (Figure S5). The ^{1}H and ^{13}C NMR spectra of **4** exhibited two methyl groups and one ethyl group attached to the benzene ring (Figures S8 and S9). This together with the fact that **4** is 14 amu more than **3** suggested that **4** is 6-ethyl-2,4-dihydroxy-3,5-dimethylbenzoic acid. Therefore, by swapping the R domain with the TE domain, we were able to engineer the chimeric PKS to produce the predicted carboxylic derivatives of **1** and **2**.

Next we tested whether TE or TE/CLC from NR-PKSs further away phylogenetically would also allow polyketide production. We created three chimeras by fusing DtbAΔR with the TE domain from OrsA, or the TE/CLC domain from StcA or WA. However, no metabolites were detected by LC/MS analysis from any strains expressing the chimeric proteins (Figure 3). These results suggested that phylogenetic analysis could serve as a guide for the success of the domain swapping experiment. It seems that the close phylogenetic relationship between domain swapped synthases might provide better domain−domain interaction of the engineered synthases which leads to the successful product release. It is not surprising that TE/CLC did not release the polyketides since TE/CLC catalyzes Claisen cyclization and exhibits low sequence similarity with TE.[5] In the cases of NR-PKSs in group VI and VII, the location of the CMeT domain after the ACP domain might imply that the two domains in the CMeT-TE or CMeT-R

didomains have evolved together. This might explain why TE from OrsA failed to release the product since OrsA lacks the CMeT domain;[10] moreover the TE from OrsA is phylogenetically far away from that of AusA and ANID_06448 (Figure S10). We also performed phylogenetic analysis on the DtbA R domain, five TE domains for which we conducted domain swapping, and six R domains excised from the NR-PKSs of *A. nidulans*. This analysis indicated that the DtbA R domain is phylogenetically more closely related to the R domains than other TE domains (Figure S11). Lastly we determined the production yield of the products for both WT DtbA and the chimeric DtbAΔR+$TE_{(AusA)}$. We showed that, under induction conditions, we were able to isolate 11.0 mg/L of **1** and 8.4 mg/L of **2** from WT DtbA and 3.1 mg/L of **3** and 2.3 mg/L of **4** from the chimeric DtbAΔR+$TE_{(AusA)}$. Therefore, we observed a 3-fold decrease in production of the chimeric NR-PKS compared with that of native DtbA (Figure S12).

The TE from ANID_06448, $TE_{(ANID\_06448)}$, is responsible for the release of 2,4-dihydroxy-3,6-dimethylbenzoic acid, whereas the TE from AusA, $TE_{(AusA)}$, is responsible for the release of 2,4-dihydroxy-3,5,6-trimethylbenzoic acid (**3**) which harbors an additional methyl group at the C-5 position in contrast to 2,4-dihydroxy-3,6-dimethylbenzoic acid. This implies that $TE_{(AusA)}$ may provide a larger and more flexible pocket for substrate binding, which might explain the observation that the production yields of **3** and **4** are more abundant in DtbAΔR+$TE_{(AusA)}$ than in DtbAΔR+$TE_{(ANID\_06448)}$ (Figure 3).

TEs have been classified into 23 families based on enzyme function and substrate identity.[17] We performed phylogenetic analysis of TE domains that we applied for domain swapping and 47 representative TE family members that belong to 23 known TE families, TE1 to TE23. Interestingly, phylogenetic results showed that $TE_{(AusA)}$ and $TE_{(ANID\_06448)}$ are phylogenetically close to TE21, whereas $TE_{(OrsA)}$, TE/$CLC_{(StcA)}$, and TE/$CLC_{(WA)}$ are phylogenetically close to TE16 (Figure S10). Members of TE21 possess the canonical TE catalytic triad consisting of a conserved serine-histidine-aspartate active site in which serine served as the reactive residue which relied on the histidine and aspartate residues to stabilize the transition state.[17−20] To better define the role of $TE_{(AusA)}$ in product release, we applied a RaptorX protein structure prediction server (http://raptorx.uchicago.edu)[21] to predict the tertiary structure of $TE_{(AusA)}$. Together with the fact that nucleophilic S123 is located within the characteristic active site motif G-X-S-X-G, the putative tertiary structure suggested that S123-H317-D285 in $TE_{(AusA)}$ could be the catalytic triad that is responsible for $TE_{(AusA)}$-mediated product release (Figure S13). To address this hypothesis, we generated active site mutants followed by heterologous expression and metabolite profile analysis. The proposed TE catalytic triad residues were individually mutated (S123A, H317F, and D285N) to assess their involvement in TE-catalyzed product release. These mutants showed no detectable product (Figure 4). As a control, two serines which are not involved in the triad were individually mutated (S152A and S162A) and gave only a slightly reduced level of product generation (Figure S14). Our results confirmed that the proposed catalytic triad (S123-H317-D285) is essential for product release.

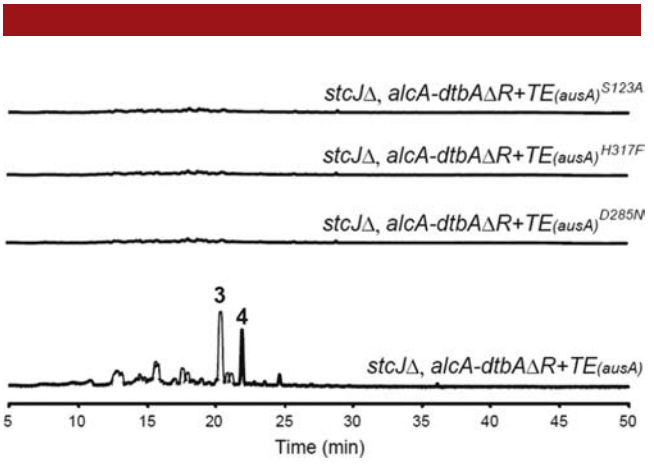


Figure 4. HPLC-DAD-MS analysis of TE triad mutant strains.

In summary, using a heterologous expression approach, we first identified the products of DtbA, a reannotated NR-PKS in *A. niger*. We further created chimeric PKSs by replacing the R domain with the TE domain which led to the production of new polyketides with functional groups that were generated by the releasing domains. Our experiments provided evidence that the replacement of the R domain with a phylogenetically close TE domain can facilitate alternative NR-PKS product release, creating a carboxylic moiety instead of an aldehyde functionality. These results indicated that rational domain swaps may provide a route for engineering the functionalization of valuable chemicals.

Acknowledgment. The project described was supported by Grant No. PO1-GM084077 from the National Institute of General Medical Science to B.R.O. and C.C.C.W. Research conducted at the Pacific Northwest National Laboratory was supported by the Department of Energy, Office of the Biomass Program.

Supporting Information Available. The primers and *Aspergillus nidulans* strains used in this study, nucleotide alignment analysis, experimental details, and characterization of **1** and **4** (^{1}H, ^{13}C NMR and HRMS). This material is available free of charge via the Internet at http://pubs.acs.org.

The authors declare no competing financial interest.

(17) Cantu, D. C.; Chen, Y.; Reilly, P. J. *Protein Sci.* **2010**, *19*, 1281–1295.
(18) Carter, P.; Wells, J. A. *Nature* **1988**, *332*, 564–568.
(19) Kim, K. K.; Song, H. K.; Shin, D. H.; Hwang, K. Y.; Choe, S.; Yoo, O. J.; Suh, S. W. *Structure* **1997**, *5*, 1571–1584.
(20) 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*, 6246–6251.
(21) Kallberg, M.; Wang, H.; Wang, S.; Peng, J.; Wang, Z.; Lu, H.; Xu, J. *Nat. Protoc.* **2012**, *7*, 1511–1522.