VIP Combinatorial Biosynthesis

DOI: 10.1002/anie.201208550

Combinatorial Domain Swaps Provide Insights into the Rules of Fungal Polyketide Synthase Programming and the Rational Synthesis of Non-Native Aromatic Products**

*Anna L. Vagstad, Adam G. Newman, Philip A. Storm, Katherine Belecki, Jason M. Crawford, and Craig A. Townsend**

Polyketide natural products encompass a structurally and functionally diverse family of molecules ranging from those of pharmaceutical importance to environmental toxins adversely affecting plant and animal health. Apart from the giant, modular polyketide synthases (PKSs), best exemplified by the trio of DEBS proteins that synthesize in an "assembly line" manner the macrocyclic core of the antibiotic erythromycin,[1,2] all other classes of PKSs function iteratively, in which a basis set of catalytic domains is re-used a fixed, or "programmed", number of times. Fatty acid synthase (FAS) and PKS enzymes similarly use acyl-CoA precursors, tether intermediates as thioesters, and catalyze two-carbon chain extensions by decarboxylative (thio)Claisen reactions with malonyl units, but differ by the extent of β-carbon processing. At one extreme are the FASs where a starter unit, typically acetyl, is extended and stepwise reductions and dehydration occur at every β-carbonyl to generate saturated fatty acid products.[3] At the other extreme are the fungal non-reducing polyketide synthases (NR-PKSs) in which the same chain-extension machinery acts, but no modifications occur, to afford a poly-β-ketone intermediate that by intramolecular aldol reaction(s) cyclizes to aromatic products (Figure 1).[4,5]

Dissection of an NR-PKS into single and multiple domain segments and their systematic reassembly ("deconstruction") enabled the synthetic contribution of each domain to the overall product formation to be deduced.[6–9] Above a specific concentration threshold, domains from a deconstructed synthase can efficiently interact to catalyze wild-type chemistry, mimicking the *in trans* associations observed in discrete type II PKSs.[7] Herein, we extend this experimental approach to combinatorial swaps of functionally analogous domains from different NR-PKSs to evaluate their compatibility for synthesis and to gain insight into the rules governing programming. We demonstrate that polycyclic aromatic products can be predictably synthesized by the rational heterocombination of NR-PKS catalytic domains.

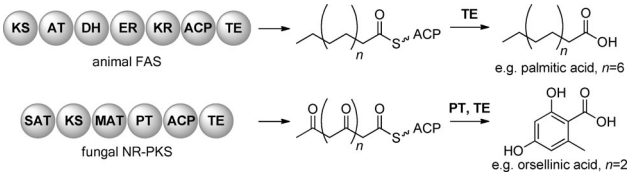

Figure 1. Comparison of animal FAS to a representative NR-PKS.

In NR-PKSs, the three N-terminal domains: starter-unit acyl transferase (SAT), ketosynthase (KS), and malonyl acyl transferase (MAT) are responsible for precursor selection and chain elongation. Of the C-terminal domains, the product template (PT) and thioesterase (TE) control backbone cyclization(s) and product release, respectively. The acyl carrier protein (ACP) domain tethers the substrates and intermediates as terminal thioesters of a conserved 4′-phosphopantetheinylated serine residue. SAT and PT are unique domains to the fungal NR-PKS family of enzymes. We hypothesized that swapping N- and C-terminal halves among NR-PKSs could generate insights into their substrate specificities and domain–domain interactions and potentially facilitate new, non-native polyketide scaffolds. The linear poly-β-ketone intermediates are intrinsically unstable and prone to spontaneous cyclizations; hence, extended ACP-tethered substrates must be generated in situ from a "minimal PKS" consisting of the substrate-loading, condensation, and carrier domains. Enzyme activity is then inferred from the backbone connectivity of the isolated products.

NR-PKSs with defined in vitro product profiles that span a variety of chain lengths (C_{12} to C_{18}), PT cyclization modes (C2-C7, C4-C9, and C6-C11), release mechanisms (pyrone formation, hydrolysis, and Claisen cyclization), and domain organizations (tandem or solo ACPs and discrete or fused TEs) were selected for analysis, five in total. The products of *Cercospora nicotianae* CTB1, *Aspergillus nidulans* wA, *Giberella fujikuroi* Pks4, *Colletotrichum lagenarium* Pks1, and *A. terreus* ACAS have been characterized and produce nor-toralactone (**1**),[10] YWA1 (**2**),[11] pre-bikaverin/SMA76a (**3**),[12] 1,3,6,8-tetrahydroxynaphthalene (T4HN, **4**),[13] and atrochrysone carboxylic acid (**5**),[14] respectively (Figure 2). Percent similarity ranges from 43% to 61% for pairwise global protein sequence alignments among these NR-PKSs

[*] Dr. A. L. Vagstad, A. G. Newman, P. A. Storm, K. Belecki, Prof. C. A. Townsend
Department of Chemistry, The Johns Hopkins University
3400 N. Charles St, Baltimore, MD 21218 (USA)
E-mail: ctownsend@jhu.edu
Prof. J. M. Crawford
Department of Chemistry, Chemical Biology Institute, Yale University
PO Box 27392, West Haven, CT 06516 (USA)

[**] An authentic sample of pannorin was generously provided by Prof. Keiji Hasumi (Tokyo University of Agriculture and Technology, Tokyo (Japan)). This work was supported by NIH Grant ES001670.

Supporting information for this article (experimental details) is available on the WWW under http://dx.doi.org/10.1002/anie.201208550.

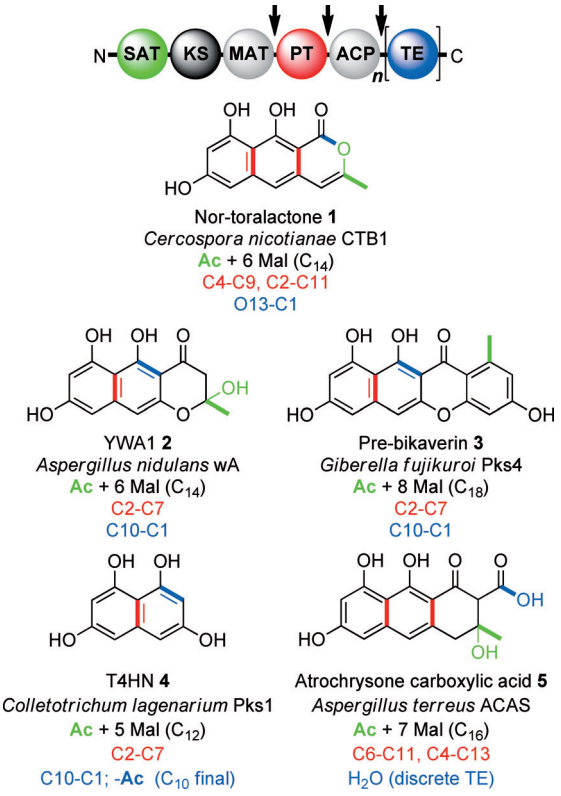

Figure 2. Native products of NR-PKSs used for domain swaps. Starter units (green), bonds made by PT-catalyzed cyclizations (red), and bonds made by TE-catalyzed release (blue) are shown. Arrows indicate sites of protein "deconstruction". Pks1, wA, and CTB1 have tandem ACPs ($n=2$); Pks4 and ACAS have a single ACP ($n=1$).

(Supporting Information, Table S2). Interdomain dissection sites were modeled on successful deconstruction of the norsolorinic acid anthrone synthase, PksA, from *A. parasiticus*[7] to generate a library of SAT-KS-MAT, PT, ACP, and TE proteins (Table S3 and Figure S1). Tandem ACPs from CTB1, wA, and Pks1 were kept intact as their didomains. CTB1 SAT-KS-MAT was selected as the parent test case because it produces a medium chain-length, C_{14} intermediate that could potentially be cyclized by all PT-types. In vitro reactions with selected sets of individually expressed and purified PT, ACP, and TE heterodomains were screened. Organic products were extracted and analyzed by high-pressure liquid chromatography (HPLC) and electrospray ionization mass spectrometry (ESI-MS).

The native chemistry of the cercosporin pathway NR-PKS, CTB1, has been established and produces heptaketides from condensation of an acetyl starter unit with six malonyl extension units.[10] Reactions lacking the TE domain demonstrated low metabolic conversion (approximately 2% of the wild-type level) indicating the critical role of the CTB1 TE in regulating metabolic flux and catalyzing pyrone-cyclization to release nor-toralactone (**1**). More commonly, fungal TEs catalyze hydrolysis or C–C bond formation by Claisen/Dieckmann cyclization (TE/CLC) to release carboxylic acid or acyl side-chain substituted aromatic products, respectively.[4] CTB1 reactions lacking a PT failed to catalyze C4-C9/C2-C11 cyclizations and instead produced low levels of spontaneously cyclized full-length shunt metabolites.[10]

ACP domains proved to be completely interchangeable in all CTB1 SAT-KS-MAT + ACP combinations, producing virtually identical product profiles with yields comparable to the matched CTB1 ACP reaction (Figure 3a, top; Figure S2). This observation is of particular note where ACP monodomains dissected from Pks4 and ACAS efficiently complemented heptaketide production with the CTB1 SAT-KS-MAT fragment, a synthase that has tandem ACPs in its native form. Fujii et al. have previously demonstrated that each of the paired ACPs is functional in the wA synthase.[11]

Next, the influence of PT domains on polyketide backbone cyclizations was investigated (Figure 3b). The ACAS PT predictably directed the linear C_{14} intermediate to form C6-C11 and C4-C13 bonds. The resulting bicyclic ACP intermediate was spontaneously released by attack of the phenolic oxygen at C5 to produce the α-pyrone pannorin (**6**). Pannorin is a native fungal metabolite of *Chrysosporium pannorum* and a 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitor, although not as potent as the medically useful statins.[15] The ACAS PT, which natively accepts a C_{16} substrate, was able to catalyze the regiospecific cyclization of the shorter C_{14} intermediate. All other synthases showed lower conversion for this domain set, which we attribute to the lack of viable release modes. This result is consistent with previous reports for Pks1 and CTB1 cognate systems that, in the absence of a functional TE, formed only low levels of spontaneously released pyrone metabolites.[10,13]

Addition of TE domains showed a sharp improvement in yields for the wA and Pks1 combinations (Figure 3c). Not surprisingly, wA—also a heptaketide synthase like CTB1—was the most efficient combination and made its wild-type product YWA1 (**2**) from conidial pigment biosynthesis. The melanin pathway Pks1 natively makes T4HN (**4**) and has an interesting TE domain with both Claisen cyclase and deacetylase activities resulting in an apparent pentaketide product from a hexaketide precursor.[13] Pks1 PT and TE domains could accommodate the longer C_{14} substrate to produce **2**; however, the TE domain lacked deacylase activity towards the intermediate bearing an acetoacetyl side chain and **4** was not detected. The final C2-C7 cyclase examined was Pks4 from the nonaketide bikaverin pathway. Production of **2** was only slightly enhanced and derailment products slightly suppressed for the Pks4 combination, and it is unclear if the PT and/or TE were limited in their activity toward the short intermediate. The discrete β-lactamase type TE, ACTE which natively releases **5**, does not appear to accept the C6-C11/C4-C13 ACP-bound intermediate as a substrate for hydrolysis. No additional carboxylic acid products were observed, and **6** was the only PT-programmed metabolite detected.

From these combinatorial domain-swapping experiments, particularly in light of isolated results from other studies, we are able to formulate some rules of thumb for NR-PKS programming, although there are certain to be exceptions. The MAT and ACP domains have consistent activities across all synthases. MAT domains demonstrate robust substrate specificity for transferring malonyl[8] and, in contrast to all other catalytic domains, are highly, if not fully, populated to

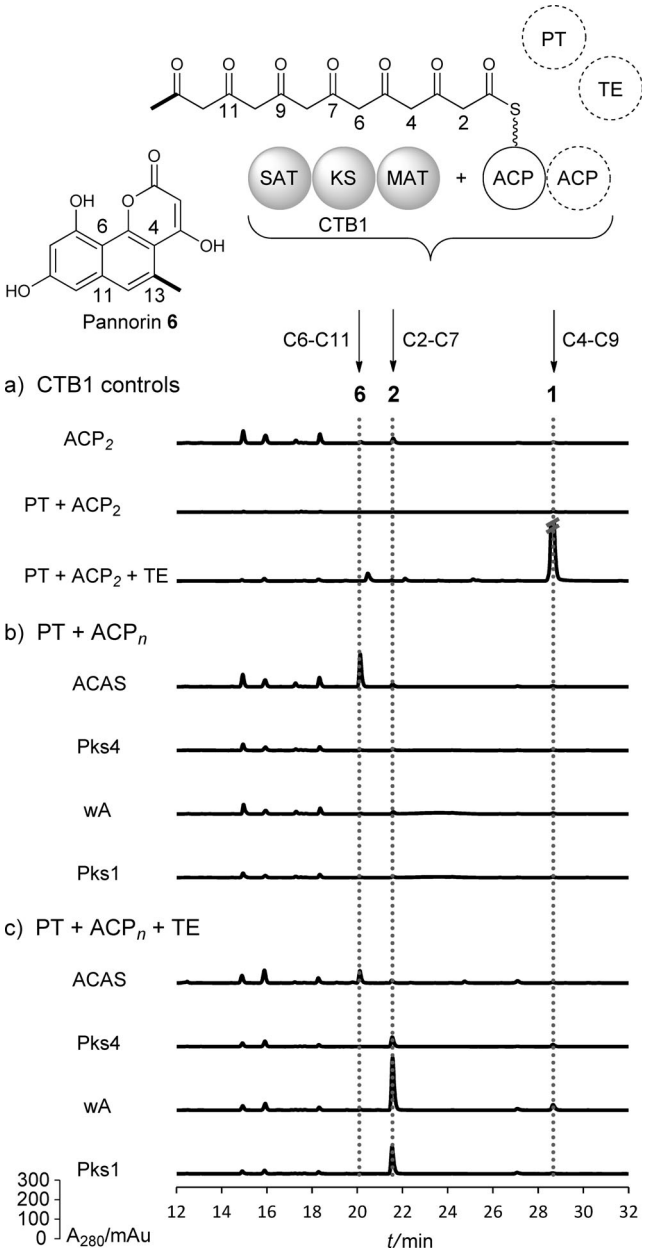

Figure 3. Product analysis of in vitro combinatorial domain-swap experiments. All reactions contain the CTB1 SAT-KS-MAT tridomain fragment. a) CTB1 control reactions. b) Reactions additionally contain PT and ACP heterodomains. c) Reactions additionally contain PT, ACP, and TE heterodomains. Minor derailment products at retention times <19 min correspond to those produced by the "minimal" set of catalytic domains (top control chromatogram and Figure S2).

support the high processivity of NR-PKSs.[9] The cross-compatibility of ACPs indicates that recognition motifs for its client domains are well conserved among NR-PKSs and generation of an elongated polyketide backbone is not the limiting factor in heterocombination experiments.

Product structural diversity among NR-PKSs arises from programmed selection of the starter unit, chain length, backbone cyclization(s), and mode of product release. The KS domain is the major determinant of chain length. In the case of CTB1 discussed herein, the KS demonstrated stringent chain-length control to faithfully produce heptaketides as the principal species, regardless of which C-terminal heterodomains were partnered. The chain length of products from PksA,[7,9] Pks4,[12] Pks1,[16] and AspE[17] (asperfuranone pathway) NR-PKS experiments further support this chain-length control function for KS domains.

The PT "register" for cyclization is measured from the phosphopantetheine thioester. The C2-C7, C4-C9, and C6-C11 type PTs were all capable of fully controlling cyclization of the linear C_{14} intermediate. ACAS, wA, and Pks1 PTs, which natively accept C_{16}, C_{14}, and C_{12} substrates, respectively, worked well, whereas the Pks4 PT—a C_{18} synthase—may be less active, indicating that similarity of chain length to the natural substrate favors efficient catalysis. Tang and co-workers observed that AptA and VrtA PTs from C6-C11 closure type asperthecin and viridicatumtoxin pathways, respectively, predictably directed cyclizations of a C_{18} polyketide intermediate.[18] AptA natively accepts a matched C_{18} chain-length substrate and VrtA uses an effective C_{20} substrate (malonamoyl starter unit and eight malonyl extensions, $C_{19}N$). The data suggest that PTs can easily accept substrates within one acetate unit of their native chain length and may additionally be compatible with shorter intermediates.

PTs are dimeric double hot-dog fold enzymes related to dehydratases. The X-ray crystal structure of PksA PT shows a large active-site cavity consisting of a starter-unit binding region (hexanoyl in this case), a cyclization chamber, and a phosphopantetheine arm tunnel.[19] The linear polyketide is thought to insert in an extended conformation, as inferred from the co-crystal structure with palmitate bound and in silico substrate-docking studies; however it is not clear how the acetyl starter-unit substrates bind into the other PT-domains to accommodate both shorter and longer polyketide intermediates. Further structural studies with meaningful intermediates bound are needed to answer these questions.

TEs are important for catalytic turnover and efficient accumulation of products. Because they act last in the pathway (apart from their editing function[9]), their activity is masked unless a compatible substrate is handed down from their upstream domains. The TE/CLC domains—Pks1, wA, and Pks4—could all act on the C_{14} C2-C7 intermediate to varying extents. The identity of the acyl side chain may not be crucial for substrate orientation to facilitate C–C bond formation in these enzymes. Detached ACTE acts by a distinct mechanism to the α/β-hydrolase fold TE enzymes that make up the C-terminus of NR-PKSs.[14] ACTE natively accepts a tricyclic substrate and no evidence for hydrolysis of a bicyclic intermediate was detected. The C6-C11/C4-C13 intermediate was released, rather, by spontaneous pyrone formation to yield **6**, much like the reaction lacking a TE.

It can be appreciated from these examples that rational choice of the starter-unit transacylase (SAT), KS to determine length of the poly-β-ketone intermediate, and a PT/TE combination can be expected to faithfully carryout its "programmed" cyclization chemistry to afford a predicted product. The synthetic efficiency of this process can be further anticipated to improve as these heterodomains are increasingly linked.[7,9]

Received: October 24, 2012
Published online: January 2, 2013

Keywords: domain swapping · natural products · pathway engineering · polyketides

[1] R. S. Gokhale, D. Hunziker, D. E. Cane, C. Khosla, *Chem. Biol.* **1999**, *6*, 117.
[2] J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* **2001**, *18*, 380.
[3] T. Maier, M. Leibundgut, N. Ban, *Science* **2008**, *321*, 1315.
[4] J. M. Crawford, C. A. Townsend, *Nat. Rev. Microbiol.* **2010**, *8*, 879.
[5] R. J. Cox, T. J. Simpson, *Methods Enzymol.* **2009**, *459*, 49.
[6] J. M. Crawford, B. C. R. Dancy, E. A. Hill, D. W. Udwary, C. A. Townsend, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16728.
[7] J. M. Crawford, P. M. Thomas, J. R. Scheerer, A. L. Vagstad, N. L. Kelleher, C. A. Townsend, *Science* **2008**, *320*, 243.
[8] J. M. Crawford, A. L. Vagstad, K. P. Whitworth, K. C. Ehrlich, C. A. Townsend, *ChemBioChem* **2008**, *9*, 1019.
[9] A. L. Vagstad, S. B. Bumpus, K. Belecki, N. L. Kelleher, C. A. Townsend, *J. Am. Chem. Soc.* **2012**, *134*, 6865.
[10] A. G. Newman, A. L. Vagstad, K. Belecki, C. A. Townsend, *Chem. Commun.* **2012**, *48*, 11772.
[11] I. Fujii, A. Watanabe, U. Sankawa, Y. Ebizuka, *Chem. Biol.* **2001**, *8*, 189.
[12] S. M. Ma, J. Zhan, K. Watanabe, X. Xie, W. Zhang, C. C. Wang, Y. Tang, *J. Am. Chem. Soc.* **2007**, *129*, 10642.
[13] A. L. Vagstad, E. A. Hill, J. W. Labonte, C. A. Townsend, *Chem. Biol.* **2012**, *19*, 1525.
[14] T. Awakawa, K. Yokota, N. Funa, F. Doi, N. Mori, H. Watanabe, S. Horinouchi, *Chem. Biol.* **2009**, *16*, 613.
[15] K. H. Haruo Ogawa, K. Sakai, S. Murakawa, A. Endo, *J. Antibiot.* **1991**, *44*, 762.
[16] A. Watanabe, Y. Ebizuka, *Tetrahedron Lett.* **2002**, *43*, 843.
[17] T. Liu, Y.-M. Chiang, A. D. Somoza, B. R. Oakley, C. C. C. Wang, *J. Am. Chem. Soc.* **2011**, *133*, 13314.
[18] Y. Li, W. Xu, Y. Tang, *J. Biol. Chem.* **2010**, *285*, 22764.
[19] J. M. Crawford, T. P. Korman, J. W. Labonte, A. L. Vagstad, E. A. Hill, O. Kamari-Bidkorpeh, S.-C. Tsai, C. A. Townsend, *Nature* **2009**, *461*, 1139.

