



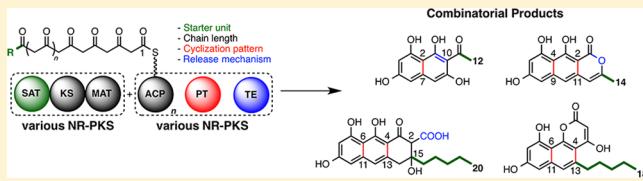
# Systematic Domain Swaps of Iterative, Nonreducing Polyketide Synthases Provide a Mechanistic Understanding and Rationale For Catalytic Reprogramming

Adam G. Newman, Anna L. Vagstad,<sup>†</sup> Philip A. Storm, and Craig A. Townsend\*

Department of Chemistry, The Johns Hopkins University, 3400 N. Charles Street, Baltimore, Maryland 21218, United States

## Supporting Information

**ABSTRACT:** Iterative, nonreducing polyketide synthases (NR-PKSs) are multidomain enzymes responsible for the construction of the core architecture of aromatic polyketide natural products in fungi. Engineering these enzymes for the production of non-native metabolites has been a long-standing goal. We conducted a systematic survey of *in vitro* “domain swapped” NR-PKSs using an enzyme deconstruction approach. The NR-PKSs were dissected into mono- to multidomain fragments and recombined as noncognate pairs *in vitro*, reconstituting enzymatic activity. The enzymes used in this study produce aromatic polyketides that are representative of the four main chemical features set by the individual NR-PKS: starter unit selection, chain-length control, cyclization register control, and product release mechanism. We found that boundary conditions limit successful chemistry, which are dependent on a set of underlying enzymatic mechanisms. Crucial for successful redirection of catalysis, the rate of productive chemistry must outpace the rate of spontaneous derailment and thioesterase-mediated editing. Additionally, all of the domains in a noncognate system must interact efficiently if chemical redirection is to proceed. These observations refine and further substantiate current understanding of the mechanisms governing NR-PKS catalysis.



## INTRODUCTION

Polyketides are a structurally and functionally diverse family of natural products containing environmental toxins and pigments as well as pharmaceutical agents. Several drugs on the market are of polyketide origin including the antibiotics tetracycline and erythromycin, immunosuppressant rapamycin, anticholesterol drug lovastatin, and anticancer drug epothilone B.<sup>1</sup> The direct engineering of polyketide biosynthetic pathways for the production of non-native metabolites has been an attractive goal for those wishing to expand the potential polyketide drug pool. However, such approaches require a greater and nuanced understanding of the underlying mechanisms at play in polyketide assembly. As a consequence, the methodology of rational redirection of polyketide pathways is still in its infancy.<sup>2,3</sup>

Many of the known polyketides are assembled in a linear fashion by large, multidomain proteins dubbed type I polyketide synthases (PKS).<sup>4</sup> While the modular type I PKSs have been extensively studied for 20 years, the biochemistry of fungal, iterative, nonreducing PKSs (NR-PKS) is only now beginning to be elucidated.<sup>5,6</sup> The NR-PKSs are responsible for the biosynthesis of a variety of aromatic polyketide products and share a common domain architecture that is intrinsically linked to their function (Figure 1). The mode of biosynthesis is analogous to that of fatty acids by animal fatty acid synthases (FASs), but simplified.<sup>7</sup> The three amino-terminal domains, the starter unit:acyl-carrier protein transacylase (SAT), ketosynthase (KS), and malonyl acyl transferase (MAT) domains, are responsible for the initiation and polyketide elongation phases.

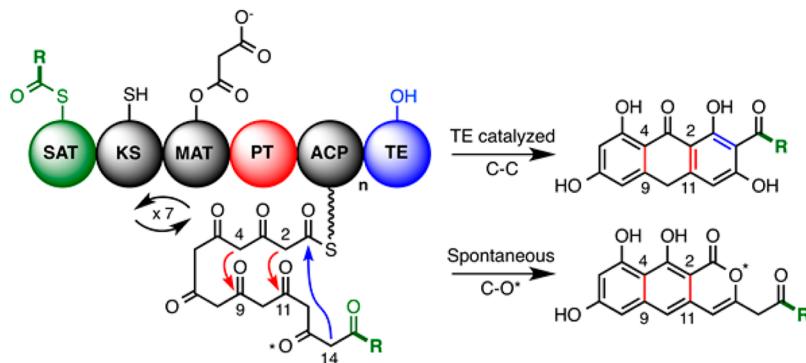
The SAT domain selects a precursor or starter unit substrate as an acyl thioester, while the MAT domain introduces ketide extender units from malonyl-CoA. The KS works in collaboration with the acyl-carrier protein (ACP) to catalyze the decarboxylative Claisen condensation of these substrates generating a linear, ACP-bound poly-β-ketone intermediate. The carboxy-terminal domains of NR-PKSs control the final stage of biosynthesis, which includes regiospecific aldol cyclizations/aromatizations by the product template (PT) domain and release through either hydrolysis,<sup>8</sup> Claisen (Dieckmann) cyclization,<sup>9,10</sup> or pyrone formation<sup>11</sup> by the thioesterase (TE). In this way, the four factors governing chemical diversity in aromatic polyketides are entirely controlled by the enzyme, with the amino-terminal half determining starter unit selection and chain length, and the carboxy-terminal half controlling regiospecific cyclization/aromatization and mechanism of product release.

Protein deconstruction, in which the NR-PKS is dissected into smaller mono- to multidomain fragments, has been a crucial tool for mechanistic understanding of these enzymes.<sup>9,12</sup> The method enables the rapid and selective *in vitro* recombination of NR-PKS activity, allowing for individual domain functions to be deduced and assayed. In these reactions, the enzymes work in *trans*, mimicking the discrete type II PKSs. The approach also allows for the recombination of an unnatural NR-PKS through domain swapping from

Received: January 22, 2014

Published: April 25, 2014





**Figure 1.** Core domain architecture of NR-PKS with highlighted enzyme bound intermediates and products of TE-directed Claisen cyclization or spontaneous O–C bond closure.

**Table 1.** Native Polyketide Products of the NR-PKSs Used in This Study

Protein	SAT Starter Unit Selection	Chain Length	PT Cyclization Register	TE Release Mechanism	Product Formed
PksA	Hexanoyl	C <sub>20</sub>	C4-C9 C2-C11	C14-C1	noranthrone, <b>1</b>
Pks4	Acetyl	C <sub>18</sub>	C2-C7	C10-C1	pre-bikaverin, <b>2</b>
ACAS	Acetyl	C <sub>16</sub>	C6-C11 C4-C13	Hydrolysis <sup>a</sup>	atrochrysone carboxylic acid, <b>3</b>
CTB1	Acetyl	C <sub>14</sub>	C4-C9 C2-C11	O13-C1	nor-torlactone, <b>4</b>
wA	Acetyl	C <sub>14</sub>	C2-C7	C10-C1	YWA1, <b>5</b>
Pks1	Acetyl	C <sub>12</sub>	C2-C7	C10-C1 <sup>b</sup>	1,3,6,8-tetrahydroxynaphthalene, <b>6</b>

<sup>a</sup>TE of the ACAS system exists as a discrete  $\beta$ -lactamase type TE, ACTE. <sup>b</sup>Pks1 TE has an additional deactylase activity.

functionally analogous domains of different NR-PKSs. Genetic domain swapping and the creation of chimeric modular, type I PKSs has been extensively examined *in vivo*, primarily with the DEBS PKSs of erythromycin biosynthesis.<sup>13,14</sup> These systems have been uniquely attractive because the domains of these canonical modular enzymes are only used once in an assembly line fashion. Additionally, the KS domains often demonstrate sufficient substrate promiscuity to allow for a degree of synthetic flexibility. Similarly, combinatorial studies of discrete type II PKSs have led to the generation of aromatic metabolite libraries.<sup>15</sup> Such approaches have focused on site-specific reductions and cyclizations catalyzed by ketoreductase (KR) and aromatase/cyclase (ARO/CYC) proteins accepting alternative chain-length polyketide intermediates.<sup>16,17</sup>

Recently, our laboratory has explored *in vitro* domain swapping as a method of developing noncognate NR-PKSs.<sup>18</sup> We postulated that swapping N- and C-terminal halves of NR-PKSs could lead to the generation of novel products. In this way starter unit selection and chain-length control would be

decoupled from the tailoring steps of late stage polyketide biosynthesis. Using the N-terminal SAT-KS-MAT tridomain fragment from the cercosporin biosynthetic NR-PKS, CTB1,<sup>11,19</sup> we were able to efficiently complement enzymatic activity with the C-terminal half of four noncognate NR-PKSs, representing a variety of cyclization and release pathways (Figure S7). These limited data codified several proposed “rules” for efficient redirection of biosynthesis, which are further tested and validated in experiments described herein. While the MAT and ACP domains behave consistently regardless of the parent synthase, the KS domain exerts stringent chain-length control, as has been demonstrated in multiple systems. The PT domain catalyzes cyclization using the bound phosphopantetheine thioester as the benchmark from which to establish regiochemistry. In this way, aromatization always occurs with the proper “register” (i.e., C2–C7, C4–C9, or C6–C11) for the given PT, even if it accepts a linear polyketide intermediate of non-native length. This observation is in keeping with the complementation of

cyclization register observed by Tang and co-workers with respect to the AptA and VrtA PTs.<sup>20</sup> Finally, the TE domain is crucial for catalytic turnover and will only exert its effect if a compatible intermediate is generated.

With the success of *in vitro* C-terminal domain swapping with the CTB1 SAT-KS-MAT, we wanted to systematically investigate the potential of this approach. We elected to study six enzymes that represent the wide span of known chemical diversity controlled by NR-PKSs: alkyl starter unit selection (acetyl or hexanoyl), chain-length control ( $C_{12}$  to  $C_{20}$ ), PT cyclization mode ( $C_2-C_7$ ,  $C_4-C_9$ , or  $C_6-C_{11}$ ), and TE release mechanism (hydrolysis, pyrone formation, or Claisen cyclization). The native activity of each enzyme investigated has been determined: *Aspergillus parasiticus* PksA, *Giberella fujikuroi* Pks4, *Aspergillus terreus* ACAS, *Cercospora nicotianae* CTB1, *Aspergillus nidulans* wA, and *Colletotrichum lagenarium* Pks1 each produce norsolorinic acid anthrone (1, noranthrone),<sup>9</sup> pre-bikaverin (2),<sup>21</sup> atrochrysone carboxylic acid (3),<sup>22</sup> nor-torvalactone (4),<sup>11</sup> YWA1 (5),<sup>23</sup> and 1,3,6,8-tetrahydroxynaphthalene (THN, 6),<sup>24</sup> respectively (Table 1).

## ■ EXPERIMENTAL SECTION

**Cloning and Heterologous Expression.** DNA manipulations were carried out in accordance with established procedures.<sup>25</sup> Details of the expression plasmids for deconstructed NR-PKSs used in this study are summarized in Table S2. Exons were artificially spliced together using the overlap extension polymerase chain reaction (PCR). Primers used for cloning new constructs used in this study are presented in Table S3. The cloning strategy for preparation of the full length CTB1 construct is presented in the Supporting Information. A chimeric full-length NR-PKS (CTB1 SAT-KS-MAT + Pks1 PT-ACP-TE) was prepared through Gibson assembly cloning. Details of its construction are presented in the Supporting Information. Cut sites for protein deconstruction were guided by a variety of bioinformatics analyses including multiple sequence alignment, secondary structure prediction, and the UMA algorithm for predicting interdomain regions. All expression plasmids were maintained in *E. coli* BL21(DE3) cells stored in 20% glycerol at  $-80^{\circ}\text{C}$ . All SAT-KS-MAT, PT, and full-length NR-PKS proteins had a C-terminal 6×-His-tag. All ACP and TE proteins had an N-terminal 6×-His-tag.

Proteins were heterologously expressed from *E. coli* BL21(DE3) harboring pET-24a(+) or pET-28a(+) (Novagen, Madison, WI) based expression vectors. Cells were grown in Luria Broth (LB) medium at  $37^{\circ}\text{C}$  in baffled flasks with constant shaking (250–300 rpm) until  $\text{OD}_{600\text{ nm}}$  0.7–0.8 at which time protein expression was induced with the addition of 1 mM isopropyl- $\beta$ -D-thiogalactoside (GoldBio, St. Louis, MO). Protein expression was conducted overnight at  $18^{\circ}\text{C}$ . Following expression, cells were collected by centrifugation at  $4100 \times g$ . The cell pellets were frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until purification.

Cell pellets were suspended in lysis buffer (50 mM potassium phosphate, pH 8.0, 300 mM NaCl, 10 mM imidazole, 10% glycerol) and lysed by sonication. Cleared lysates were collected by centrifugation at  $28000 \times g$ . Proteins were purified by  $\text{Ni}^{2+}$ -affinity chromatography according to standard practices using high-density nickel agarose resin (GoldBio, St. Louis, MO). Purified proteins were dialyzed overnight at  $4^{\circ}\text{C}$  into 100 mM potassium phosphate pH 7.0 and 10% glycerol. Proteins were concentrated using Amicon Ultra centrifugation devices (Millipore, Billerica, MA), and final protein concentrations were determined by Bradford assay using protein dye reagent (Bio-Rad, Hercules, CA) and bovine serum albumin (New England Biolabs, Ipswich, MA) as a standard. ACP-containing proteins were activated to the *holo* form through the action of the promiscuous phosphopantetheinyl transferase Svp, as previously described.<sup>9</sup>

**In Vitro Reactions.** *In vitro* reactions were conducted using S-acetyl-N-acetylcysteamine (acyl-SNAC) substrates in place of acyl-CoAs, as previously described.<sup>11,18</sup> Acetyl- and hexanoyl-SNACs were synthe-

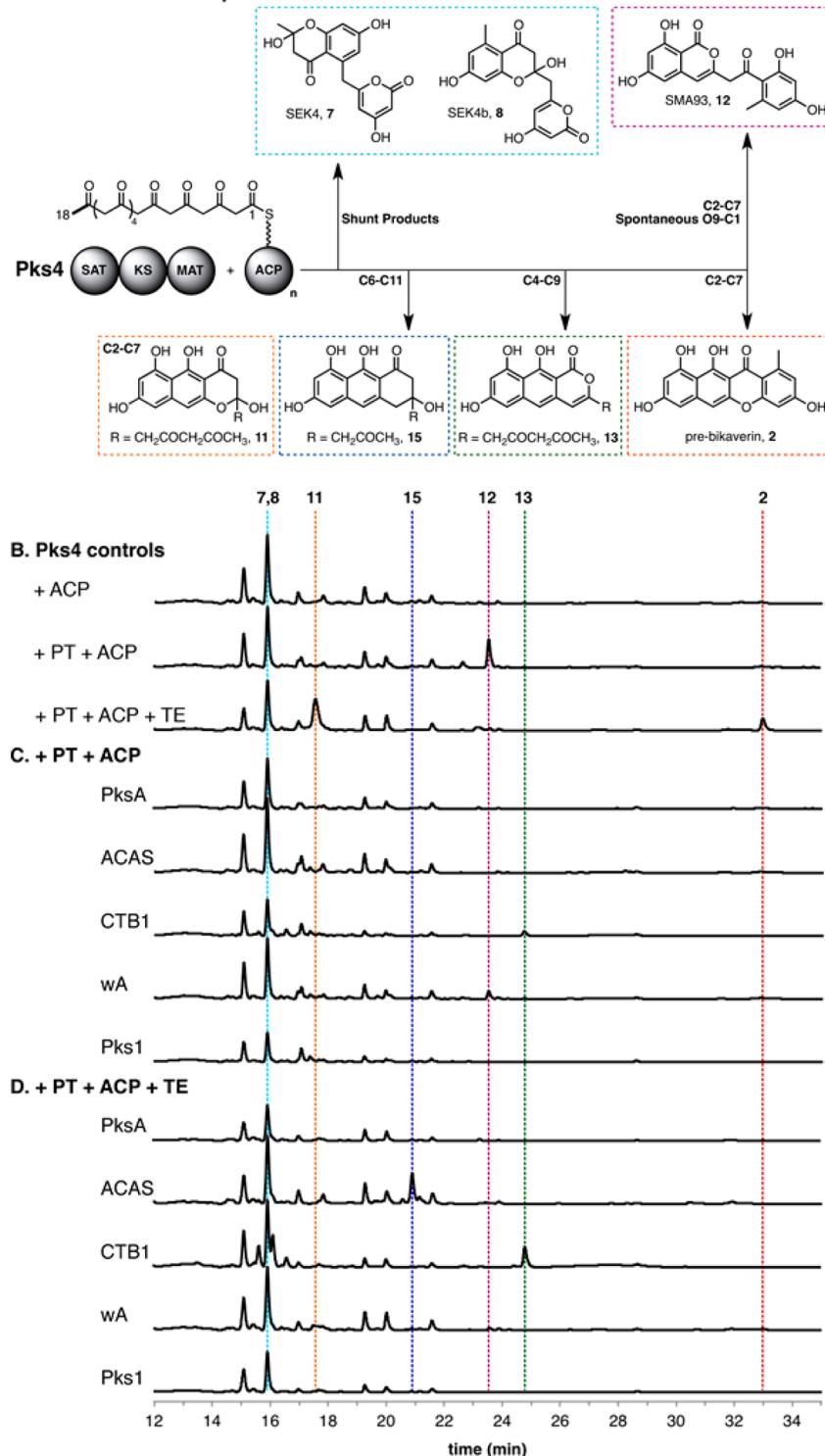
sized from the corresponding acid chlorides. The malonyl-CoA synthetase from *Rhizobium leguminosarum*, MatB, was used to produce HPLC purified malonyl-SNAC, as previously described.<sup>26,27</sup> Purified enzyme fragments were selectively recombined *in vitro* at 10  $\mu\text{M}$  final concentration for each protein in 100 mM potassium phosphate pH 7.0, 1 mM tris(2-carboxyethyl)phosphine (TCEP), and 10% glycerol. Reactions were initiated with the simultaneous addition of either 0.5 mM hexanoyl-SNAC (for reactions containing PksA SAT-KS-MAT) or 0.5 mM acetyl-SNAC (all other reactions) and 2 mM malonyl-SNAC (all reactions) and were conducted for 4 h at room temperature. Reactions were quenched by acidification with HCl, and products were extracted into ethyl acetate. Organic fractions were combined, dried *in vacuo*, and dissolved in 80% water, 20% acetonitrile at a final volume equivalent to the initial reaction volume.

**Product Analysis.** Reactions were analyzed by reverse phase HPLC on an Agilent 1200 instrument (Agilent Technologies, Santa Clara, CA). Solvent A was water + 0.1% formic acid. Solvent B was acetonitrile + 0.1% formic acid. Purified reaction extracts were injected onto a linear gradient of 5% to 85% solvent B over 30 min at 1 mL/min on a Prodigy 5u ODS3 column (4.6 × 250 mm, 5  $\mu$ ; Phenomenex, Torrance, CA). Chromatograms were recorded at 280 nm, 4 nm bandwidth with a background reference of 650 nm, 100 nm bandwidth. Mass data, unless otherwise stated, were collected on a Shimadzu LC-IT-TOF (Shimadzu Corporation, Kyoto, Japan) in positive ion mode fitted with a Luna C18(2) column (2.0 × 150 mm, 3  $\mu$ ; Phenomenex, Torrance, CA) using a linear gradient of 5% to 85% solvent B over 30 min at 0.2 mL/min. Due to low titers, product identification by NMR spectroscopy was not feasible. Notwithstanding, as NR-PKSs act through defined chemistry, distinct UV-vis absorption profiles and high resolution MS data are sufficient for metabolite characterization. Unequivocal structural assignment was available for the majority of products through comparison to synthetic standards and literature values of fully characterized materials. Provisional assignments based upon available data are made for selected species for which unambiguous identification was not possible. A complete collection of characterization data for each product is presented in the Supporting Information.

## ■ RESULTS

**Deconstruction of NR-PKSs.** The dissection of NR-PKSs into component mono- to multidomain fragments was patterned on the previously reported deconstruction of PksA.<sup>9</sup> The percent similarity between the NR-PKSs used in this study ranged from 43% to 61% as calculated from global pairwise protein sequence alignments (Table S1). Equivalent cut sites for the other NR-PKSs used in this study were selected based on primary sequence alignments. Alternative cut sites were selected to optimize yields for proteins that suffered from low expression, insolubility or instability. SAT-KS-MAT tridomain, PT, ACP<sub>n</sub>, and TE monodomain fragments were cloned as hexahistidine fusions for heterologous expression in *E. coli* followed by nickel affinity purification (Figure S1). The tandem ACPs from CTB1, wA, and Pks1 were prepared as intact didomain fragments, here referred to as ACP<sub>2</sub>. The discrete  $\beta$ -lactamase type TE of atrochrysone biosynthesis, ACTE,<sup>22</sup> was expressed separately. SAT-KS-MAT tridomain fragments for ACAS and wA could not be obtained despite repeated attempts to optimize cut-site selection, fusion tag identity and location, and culture conditions. Details of individual protein constructs used in this study are presented in Table S2. All constructs were sequence verified.

**Noncognate Minimal NR-PKS Compatibility.** Given the successful redirection of biosynthesis in combinatorial reactions of CTB1,<sup>18</sup> we sought to systematically evaluate if noncognate NR-PKS domains, in general, could predictably alter biochemistry. Combinatorial reactions were modeled on the *in*

**A. Pks4 combinatorial products**

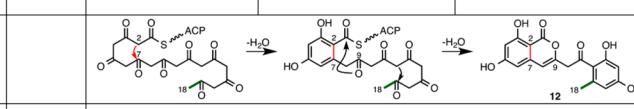
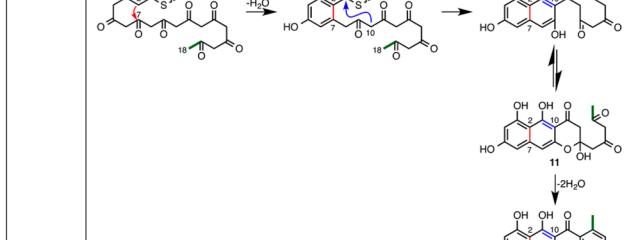
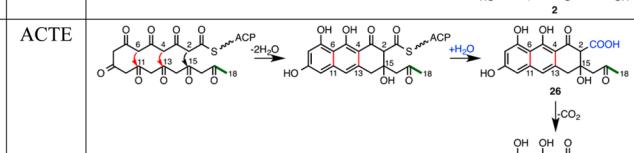
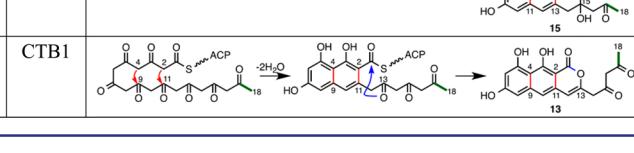
**Figure 2.** Product analysis of combinatorial reactions with Pks4 SAT-KS-MAT. A) Proposed structures for products of chemical redirection. B) Pks4 control reactions. C) Combinatorial reactions containing PT and ACP<sub>n</sub> for the given parent PKS. D) Combinatorial reactions containing PT, ACP<sub>2</sub>, and TE for the given parent PKS.

*in vitro* reconstitution of PksA, using 10  $\mu$ M protein concentrations and acyl-N-acetylcysteamine (SNAC) thioester substrates.<sup>9</sup> Extracted products from these reactions were analyzed by reverse phase HPLC and LC-ESI-MS. Product identification was accomplished by comparing UV-vis spectra and masses to those of authentic standards or literature values. As in the previous study,<sup>18</sup> selected sets of PT, ACP<sub>n</sub>, and TE

monodomains from a single parent synthase were recombined with a noncognate SAT-KS-MAT tridomain. By systematically varying the SAT-KS-MAT tridomain, we were able to achieve 72 individual combinations, representing the complete set of 2-part noncognate configurations.

We first established that all minimal NR-PKSs were active, regardless of the identity of the ACP<sub>n</sub> domain. We define a

Table 2. Intermediates and products produced by combinatorial reactions with Pks4 SAT-KS-MAT

SAT-KS-MAT	PT + ACP	TE	Linear Poly- $\beta$ -ketone	PT-Cyclized Intermediate	Product Released
Pks4	Pks4				12
Pks4	Pks4	Pks4			11 2
Pks4	ACAS	ACTE			26 15
Pks4	CTB1	CTB1			13

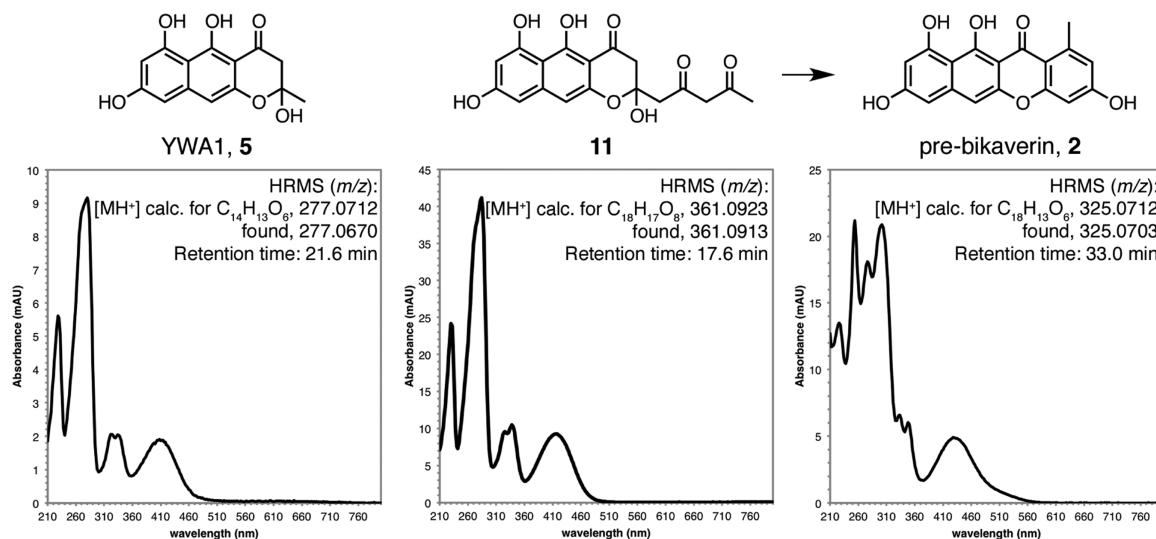
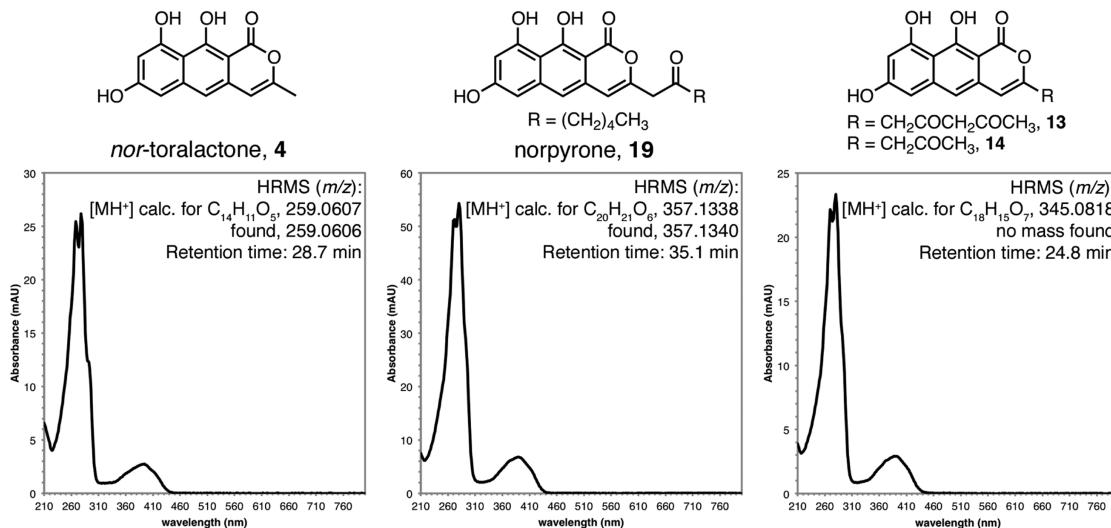


Figure 3. UV-vis spectra and high-resolution mass spectra (HRMS) for YWA1 (5) core containing molecules. Product 11 converts to pre-bikaverin (2).

minimal NR-PKS as the requisite set of domains for proficient ketide homologation accomplished in the initial stages of polyketide biosynthesis. In practice, the minimal NR-PKS consists of the SAT-KS-MAT and ACP<sub>n</sub> domains. The combination of SAT-KS-MAT with noncognate ACP<sub>n</sub> gave nearly identical product profiles in similar yields to that of the cognate reaction in all cases (Figures S2–S5). It is noteworthy that the native ACP monodomains (PksA, Pks4 and ACAS) were able to efficiently complement activity with SAT-KS-MAT tridomains from proteins with native tandem ACPs (CTB1 and Pks1). It has been established that in wild-type wA both ACPs are independently functional.<sup>23</sup> These data further strengthen the view that independent functionality is generally true of

tandem ACPs. The role of multiple ACPs may then be to improve biosynthetic efficiency, as has been shown in polyunsaturated fatty acid synthases, which contain up to six tandem ACPs.<sup>28</sup> The complete compatibility of ACPs indicates that the recognition motifs for this domain are highly conserved among the NR-PKSs used in the current study. It is expected that ACP compatibility is a universal feature of NR-PKSs and an essential precondition of the planned domain swapping experiment.

Overall, the products of the minimal NR-PKSs were consistent, by mass, with shunt products of full-length poly- $\beta$ -ketones produced by the respective SAT-KS-MAT protein, with one exception. The major products of the minimal Pks4



**Figure 4.** UV-vis spectra and HRMS for naphthopyrone containing products *nor-toralactone* (**4**) and *norpyprome* (**19**) are presented with those for likely naphthopyrone **13** or naphthopyrone **14**.

are a pair of octaketide ( $C_{16}$ ) truncations coincident with the actinorhodin type II PKS products, SEK4 (**7**) and SEK4b (**8**, Figure 2B).<sup>29,30</sup> Pks4 is expected to yield  $C_{18}$ -chain-length, nonaketide products (**9**, **10**, Figure S6), as has been observed in the reactions of 50  $\mu$ M Pks4 KS-MAT and 10  $\mu$ M Pks4 ACP.<sup>31</sup> It should be noted that in the current study, the concentration of Pks4 SAT-KS-MAT is 10  $\mu$ M, indicating that enzyme concentration might be a factor influencing chain length for deconstructed synthases. It is likely that Pks4 exhibits reduced processivity as the growing linear intermediate approaches the full chain length. The reduction in processivity is in agreement with a similar effect in PksA, where truncated heptaketide shunt products were observed *in vitro*.<sup>27</sup> This effect might be an artifact of the enzyme deconstruction approach. If the Pks4 KS domain has a decreased affinity for the  $C_{16}$ -linear poly- $\beta$ -ketone intermediate, the KS and ACP domains could freely dissociate in the deconstructed system allowing for derailing cyclizations and off-loading. For the intact system and perhaps for deconstructed enzymes at higher concentrations, dissociation is less likely.

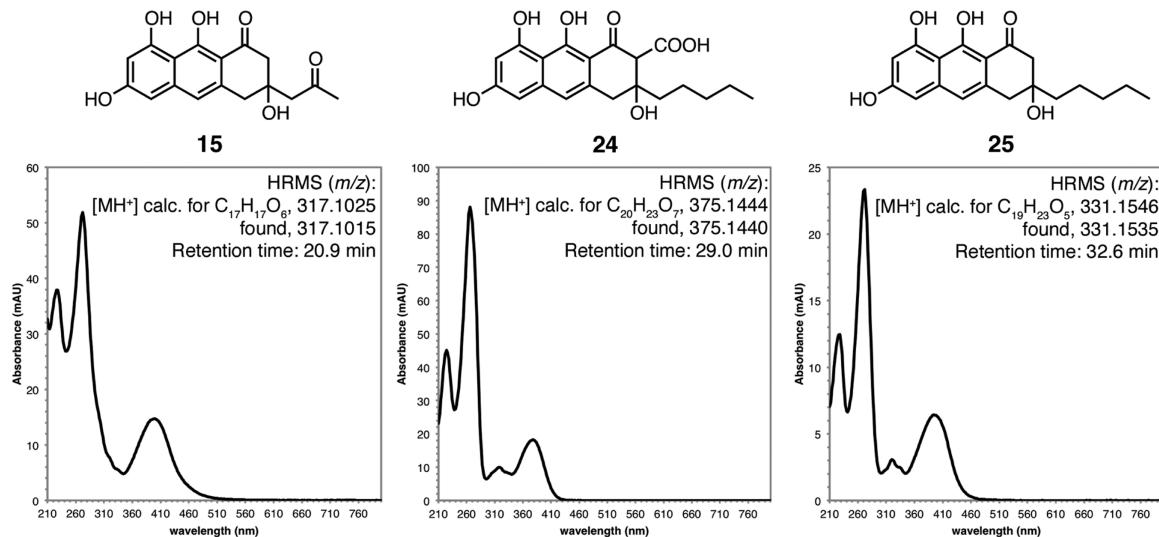
**Combinatorial Reactions with Pks4 SAT-KS-MAT.** HPLC product profiles for combinatorial reactions utilizing Pks4 SAT-KS-MAT are presented in Figure 2. Reactions for successful combinations are delineated in Table 2. As has been reported, the Pks4 PT catalyzes C2–C7 aldol cyclization/aromatization while the Pks4 TE catalyzes C10–C1 Claisen/Dieckmann cyclization on a nonaketide ( $C_{18}$ ) linear intermediate to generate pre-bikaverin (**2**, Table 2).<sup>21</sup> The third ring (O9–C13) and fourth ring (C12–C17) are postulated to form spontaneously, followed by two dehydrations. Indeed, pre-bikaverin is observed upon complete reconstitution of a cognate Pks4 (Figure 2B). Interestingly, an additional product **11** of this reaction was observed to have a virtually identical UV-vis spectrum to that of YWA1 (**5**) suggesting that they share the same core architecture (Figures 2B and 3).<sup>23</sup> The mass of this product was found to correspond to a formula of  $C_{18}H_{16}O_8$  and there is an apparent complete conversion of it to pre-bikaverin with time. Given these data, we postulate that this product is the  $C_{18}$ –YWA1 analogue **11**, in which the methyl substituent bears an acetoacetyl side chain extension (Table 2). This hypothesis would be consistent with proper PT and TE

cyclization followed by hemiketal formation prior to cyclization/aromatization of the fourth ring.

Neither pre-bikaverin nor the  $C_{18}$ –YWA1 analogue **11** is detected in reactions of Pks4 lacking the TE. Instead, SMA93 (**12**)<sup>32</sup> is observed (Figure 2B). SMA93 arises from proper C2–C7 PT-catalyzed cyclization of the  $C_{18}$ -linear intermediate followed by spontaneous O9–C1 closure, a common release mechanism (Table 2). In reactions of Pks4 SAT-KS-MAT with wA PT and ACP<sub>2</sub>, small quantities of SMA93 were also identified, indicating that wA PT directed C2–C7 cyclization (Figure 2C). However, wA TE could not complement C10–C1 cyclization, and only vanishingly small amounts of pre-bikaverin were formed (Figure 2D).

On the other hand, combinations with CTB1 ( $C_{14}$ ) PT, ACP<sub>2</sub> and TE redirected biosynthesis toward a product of similar retention time and spectrum to that of *nor-toralactone* (**4**), the native CTB1 product (Figure 2D).<sup>11</sup> Smaller quantities of this product are observed in reactions lacking the CTB1 TE domain (Figure 2C). The product has the characteristic spectrum of a naphthopyrone and can be compared to both *nor-toralactone* and *norpyprome* (**19**, Figure 4).<sup>9,11</sup> In keeping with these data, we propose a  $C_{18}$ –*nor-toralactone* analogue **13** arising from CTB1 PT-catalyzed C4–C9 and C2–C11 cyclization, followed by CTB1 TE-assisted O13–C1 pyrone formation (Table 2). Because a mass for this product could not be detected, the  $C_{16}$ –*nor-toralactone* analogue **14** (in which the methyl group has been extended with a single acetyl substituent) cannot be ruled out. Two additional unique products were observed in these reactions, eluting at 15.6 and 16.1 min (Figure 2D). We did not conduct structural characterization of these products but speculate that they arise from CTB1 PT- and TE-influenced spontaneous cyclization and release.

ACAS ( $C_{16}$ ) was also able to complement Pks4 SAT-KS-MAT, producing a product with a retention time of 20.9 min and a molecular formula of  $C_{17}H_{16}O_6$  as determined by LC-ESI-MS (Figure 2D). The UV-vis spectrum for this product is in agreement with that for atrochrysone, the decarboxylated native product of ACAS, atrochrysone carboxylic acid (Table 1).<sup>22,33</sup> If the ACAS PT domain catalyzed its native C6–C11 and C4–C13 cyclizations followed by ACTE-catalyzed hydro-



**Figure 5.** UV-vis spectra and HRMS for atrochrysone core containing molecules are presented for comparison. Spectral data are consistent for literature values for atrochrysone.<sup>33</sup>

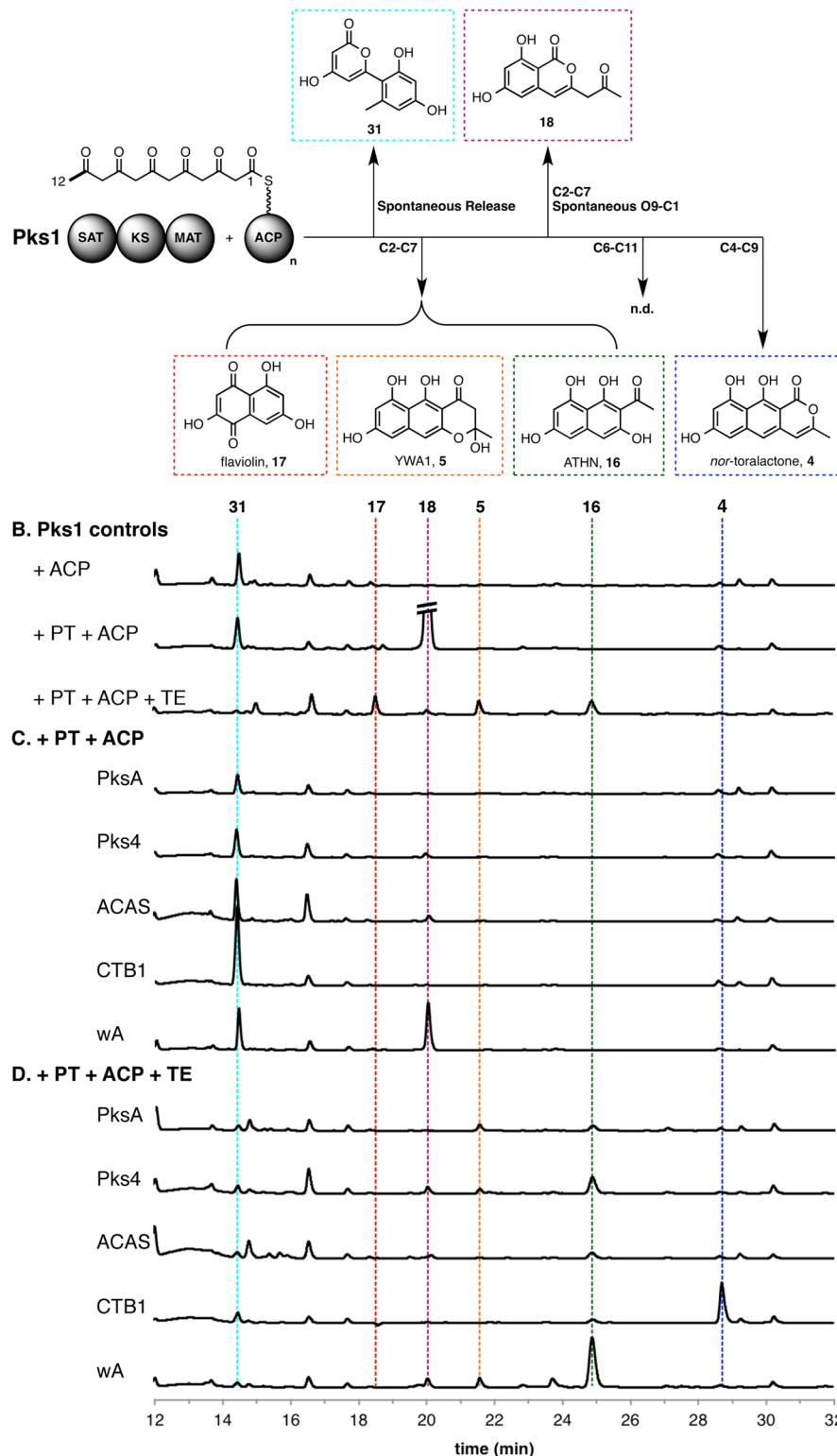
lytic release on the  $C_{18}$  linear poly- $\beta$ -ketone intermediate, the resulting product, **26**, would contain the atrochrysone core architecture (Table 2). We propose the structure **15**, which would arise from decarboxylation of acid **26** and would satisfy both the mass and spectral data (Figure 5). A similar decarboxylation is observed in reactions of native ACAS, with atrochrysone carboxylic acid decomposing spontaneously to atrochrysone.<sup>22</sup>

Notably, reactions of Pks4 SAT-KS-MAT with PksA ( $C_{20}$ ) PT, ACP and TE domains showed no evidence for redirection of chemistry (Figure 2D). This observation was unexpected because PksA normally processes a  $C_{20}$ -length intermediate and could, therefore, easily accommodate the expected  $C_{18}$ -length nonaketide intermediate of Pks4. Instead, reactions with all combinations of PksA domains showed activity similar to that of the minimal Pks4 alone. The inability of the PksA domains to complement activity with the Pks4 intermediate likely arises from the unique hexanoyl starter unit of PksA biosynthesis.<sup>34</sup> The crystal structure of PksA PT shows a structurally distinct, hydrophobic binding pocket deep in its active site that presumably binds the hexyl moiety of the PksA linear poly- $\beta$ -ketone.<sup>35</sup> It is probable that this binding pocket is incompatible with the poly- $\beta$ -ketone of acetyl-initiated polyketides, nullifying binding and cyclization by the PksA PT.

**Combinatorial Reactions of Pks1 SAT-KS-MAT.** HPLC product profiles for combinatorial reactions utilizing Pks1 SAT-KS-MAT are presented in Figure 6. Reactions for successful combinations are delineated in Table 3. The native product of Pks1 is 1,3,6,8-tetrahydroxynaphthalene (THN, **6**), an apparent pentaketide ( $C_{10}$ ) product.<sup>24</sup> However, as has been previously demonstrated, Pks1 in fact catalyzes the formation of a hexaketide ( $C_{12}$ ) intermediate from the condensation of an acetyl starter unit with five malonyl equivalents. The Pks1 PT domain catalyzes C2–C7 cyclization with the TE domain catalyzing the formation of the second ring through C10–C1 Claisen condensation cyclization, generating 2-acetyl-1,3,6,8-tetrahydroxynaphthalene (ATHN, **16**). The Pks1 TE domain has an additional, unique activity and will catalyze concomitant deacetylation to form THN directly. THN is further auto-oxidized to the naphthoquinone flaviolin (**17**), which is the observed product of *in vitro* reactions (Table 3).<sup>24</sup> These

activities are clearly demonstrated in the present Pks1 control reactions (Figure 6B). In reactions lacking the TE domain, the major product is an isocoumarin hexaketide product **18** that arises from PT-catalyzed C2–C7 cyclization followed by spontaneous O9–C1 release (Table 3).<sup>24</sup> The fully reconstituted system, although inefficient, produces ATHN and flaviolin as well as the heptaketide product YWA1 (**5**). The deconstructed Pks1 shows a marked reduction in efficiency versus the intact or minimally deconstructed (SAT-KS-MAT-PT-ACP<sub>2</sub> + TE) proteins, which produce flaviolin nearly exclusively and in high yields.<sup>24</sup> The appearance of heptaketide products in reactions with the TE is concurrent with a reduction of overall biosynthetic capacity (relative to reactions lacking the TE). This observation indicates that chain-length control, while highly regulated, is likely tied to overall turnover.

Indeed, reactions of Pks1 SAT-KS-MAT with CTB1 ( $C_{14}$ ) PT, ACP<sub>2</sub>, and TE produce *nor*-toralactone (**4**) as the clear major product (Figure 6D). *Nor*-toralactone is created by CTB1 PT-mediated C4–C9 and C2–C11 aldol cyclizations followed by CTB1 TE-catalyzed O13–C1 bond closure of a heptaketide intermediate (Table 3) and is the native product of CTB1.<sup>11</sup> The generation of *nor*-toralactone is attended by a simultaneous reduction in shunt product formation indicating that both the PT and TE domains of CTB1 have a marked influence over chain-length control in this particular non-cognate PKS. It is likely that the CTB1 PT domain is capturing both hexaketide (native,  $C_{12}$ ) and heptaketide (native +  $C_2$ ,  $C_{14}$ ) linear intermediates but only efficiently catalyzing cyclization of the  $C_{14}$  intermediate. In native CTB1, the TE domain affects a dramatic enhancement of overall turnover.<sup>11</sup> A similar kinetic role for the CTB1 TE is probably at play in the present study. In reactions of Pks1 SAT-KS-MAT with CTB1 PT and ACP<sub>2</sub> only, very small quantities of *nor*-toralactone are formed, while the  $C_{12}H_{10}O_5$  hexaketide shunt product **31** is greatly enhanced (Figure 6C). This observation follows native CTB1 reactions lacking the TE domain where overall biosynthetic efficiency is greatly reduced.<sup>11</sup> We propose that after CTB1 captures and cyclizes the heptaketide intermediates, the resulting enzyme-bound bicyclic acyl intermediate is slowly released through spontaneous pyrone formation. Therefore, the only products that result in enzymatic turnover are the

**A. Pks1 combinatorial products**

**Figure 6.** Product analysis of combinatorial reactions with Pks1 SAT-KS-MAT. (A) Proposed structures for products of chemical redirection. (B) Pks1 control reaction. Combinatorial reactions containing (C) PT and ACP<sub>n</sub> and (D) PT, ACP<sub>2</sub>, and TE for the given parent PKS.

hexaketide shunt products that have a higher rate of spontaneous release. This effect is completely masked by the highly efficient CTB1 TE domain, which quickly releases the bicyclic intermediates through pyrone formation, effectively enriching *nor*-toralactone in the product pool.

Reactions of Pks1 SAT-KS-MAT with wA (C<sub>14</sub>) PT, ACP<sub>2</sub>, and TE domains stand in contrast to the results of the Pks1/CTB1 noncognate system. While the native wA produces and functions on a heptaketide intermediate (as in CTB1),<sup>23</sup> the Pks1/wA noncognate system does not alter the native Pks1 hexaketide chain-length control. Reactions of Pks1 SAT-KS-

Table 3. Intermediates and Products Produced by Combinatorial Reactions with Pks1 SAT-KS-MAT

SAT-KS-MAT	PT + ACP	TE	Linear Poly- $\beta$ -ketone	PT-Cyclized Intermediate	Product Released
Pks1	Pks1 wA				
Pks1	Pks1	Pks1			
Pks1	Pks4 wA	Pks4 wA			
Pks1	CTB1	CTB1			

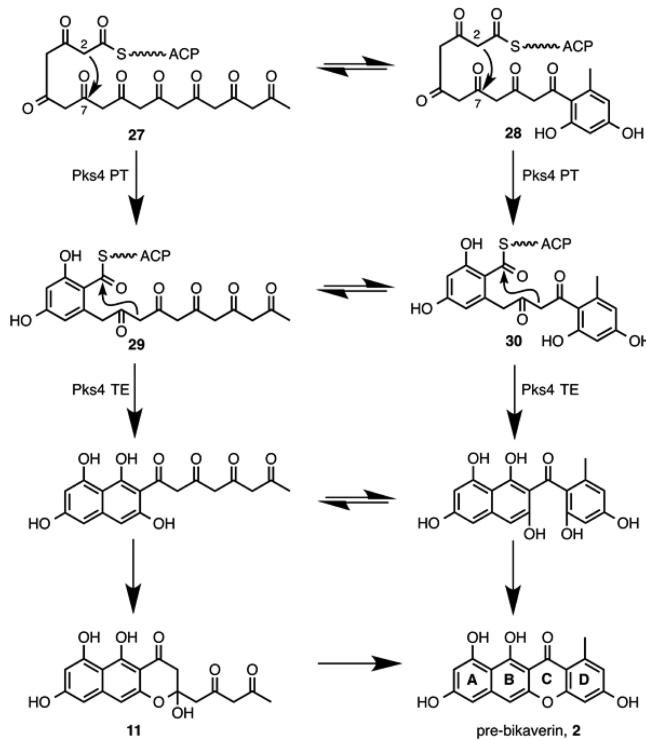
MAT with wA PT and ACP<sub>2</sub> greatly enhance the production of the isocoumarin **18** indicating that the wA PT domain is catalyzing its expected C2–C7 cyclization of the non-native hexaketide linear poly- $\beta$ -ketone (Figure 6C). In reactions including the wA TE, the chemistry is redirected toward ATHN formation, consistent with catalyzed release through C10–C1 Claisen cyclization (Figure 6D). As with the Pks1/CTB1 noncognate system, inclusion of the wA TE led to a parallel reduction in shunt product formation. This effect was not observed in Pks1/wA reactions lacking the TE domain. As with CTB1, this finding implicates the wA TE domain in the many-fold enhancement of overall efficiency through kinetic competition.

It is evident that reactions of Pks1 SAT-KS-MAT with Pks4 (C<sub>18</sub>) PT, ACP, and TE redirected chemistry toward ATHN, albeit without the efficiency of the Pks1/wA noncognate pair (Figure 6D). Reactions lacking the Pks4 TE domain produced the hexaketide isocoumarin **18** (Figure 6C). This outcome represents the expected Pks4 PT- and TE-catalyzed cyclizations but on a much smaller intermediate (Table 3). In fact, the hexaketide intermediate of Pks1 is a third smaller than the nonaketide intermediate of Pks4. This behavior is a testament to the considerable versatility of the Pks4 PT and TE domains, even if the redirection is inefficient. It also stands in stark contrast to the activity of the Pks1/CTB1 noncognate pairs, where there is apparent chain-length discrimination. There is no evidence for the Pks4 domains exerting any influence over Pks1 KS chain-length control.

This result is even more remarkable when it is considered that neither PksA nor ACAS was able to complement Pks1 SAT-KS-MAT, with these reactions producing only shunt products (Figure 6). As with Pks4, both PksA and ACAS accept larger intermediates, C<sub>20</sub> and C<sub>16</sub>, respectively.<sup>9,22</sup> The fact that both of these enzymes demonstrate chain-length discrimination while Pks4 PT and TE do not further confirms the flexibility of Pks4. Similarly, it has been previously shown that the Pks4 PT and TE domains also process the heptaketide intermediate produced by CTB1 SAT-KS-MAT (Figure S7). The Pks4 PT and TE may be inherently promiscuous with respect to substrate selectivity. While it is known that both domains process C<sub>18</sub> intermediates, the exact chemical natures of these intermediates are unclear. The native metabolite of Pks4 is pre-bikaverin. It is certain that the PT and TE domains set the A

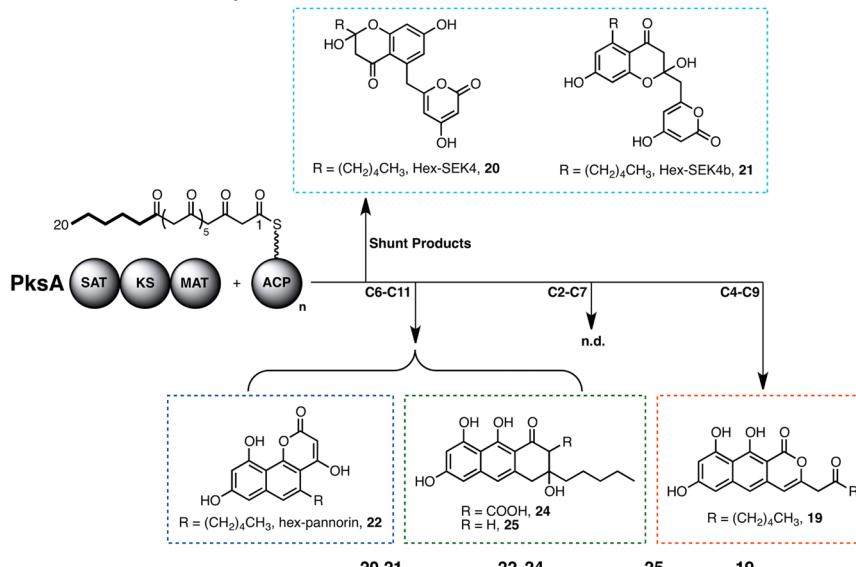
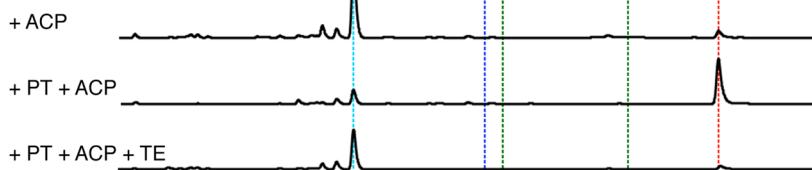
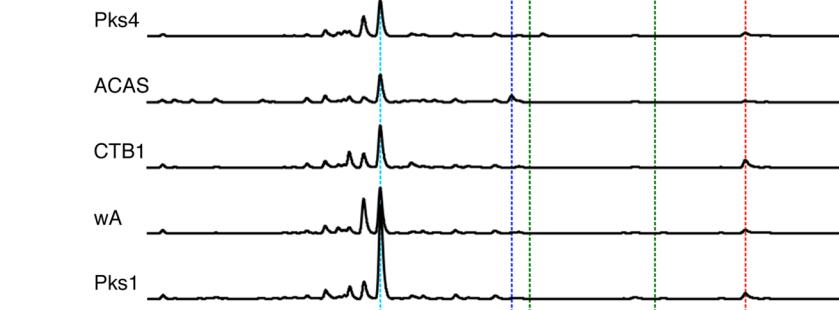
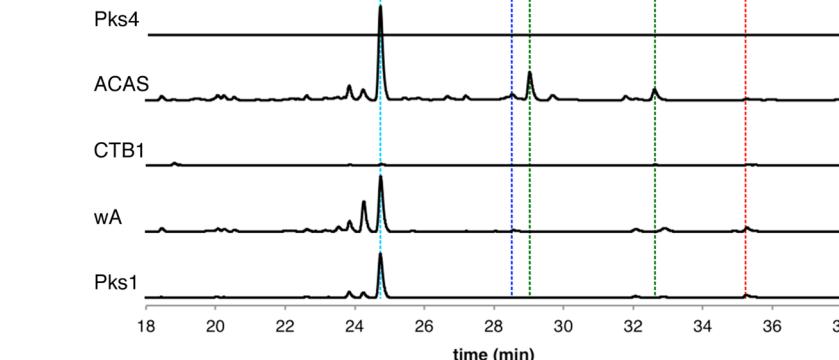
and B rings of pre-bikaverin, however the exact timing of the C and D ring formation is unclear (Scheme 1). It is possible that

### Scheme 1. Possible Intermediates of the Pks4 PT and TE Domains



the Pks4 PT could accept a linear C<sub>18</sub> poly- $\beta$ -ketone **27**, a monocyclic C<sub>18</sub> intermediate **28**, or both. Either one of these species could converge on pre-bikaverin. If it is the case that both of the species exist in the population of reactive intermediates and the PT domain accepts either as a substrate, it may explain the apparent lack of chain-length discrimination by Pks4 PT. Similarly, Pks4 TE could accept either a monocyclic intermediate **29**, a "pre-cyclized" intermediate **30**, or both, explaining its lack of specificity.

In all combinatorial reactions lacking a TE domain, an additional derailment product, 6-(2',4'-dihydroxy-6'-methyl-

**A. PksA combinatorial products****B. PksA controls****C. + PT + ACP****D. + PT + ACP + TE**

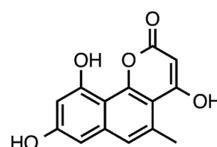
**Figure 7.** Product analysis of combinatorial reactions with PksA SAT-KS-MAT. (A) Proposed structures for products of chemical redirection. (B) PksA control reaction. (C) Combinatorial reactions containing PT and ACP<sub>n</sub> for the given parent PKS. (D) Combinatorial reactions containing PT, ACP<sub>n</sub>, and TE for the given parent PKS.

phenyl)-4-hydroxy-2-pyrone (**31**), is observed. The derailment arises from improper cyclization of the C<sub>12</sub> intermediate with C–O bond closure release yielding the pyrone. While this

product has been observed in type III PKSs, it occurs here from uncatalyzed cyclization.<sup>36,37</sup> Higher levels of production in reactions lacking a TE domain underscore the importance of

Table 4. Intermediates and Products Produced by Combinatorial Reactions with PksA SAT-KS-MAT

SAT-KS-MAT	PT + ACP	TE	Linear Poly- $\beta$ -ketone	PT-Cyclized Intermediate	Product Released
PksA	PksA				
PksA	ACAS				
PksA	ACAS	ACTE			



pannorin, 23

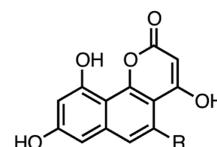
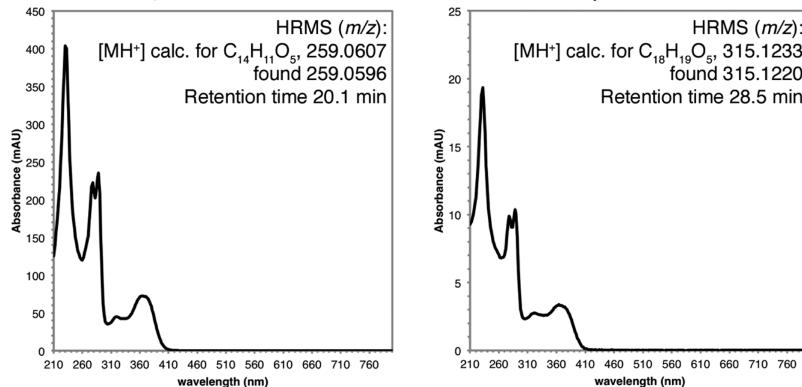
hex-pannorin, 22  
R = (CH<sub>2</sub>)<sub>4</sub>CH<sub>3</sub>

Figure 8. UV-vis spectra and HRMS are presented for pannorin (23) and hex-pannorin (22).

the TE domain for overall catalytic efficiency in PKSs. Without a TE domain forcing catalytic turnover, off-loading of derailments like product 31 can outpace even PT-catalyzed cyclizations. Further, the TE domain's editing role reduces the accumulation of this product by ensuring off-pathway routes are minimized.

**Combinatorial Reactions of PksA SAT-KS-MAT.** HPLC product profiles for combinatorial reactions utilizing PksA SAT-KS-MAT are presented in Figure 7. Reactions for successful combinations are delineated in Table 4. Of the NR-PKSs investigated in this study, PksA is unique in that it accepts an abbreviated fatty acid hexanoyl starter unit.<sup>34</sup> As such, hexanoyl-SNAC was used as the starter unit in noncognate combinations using the PksA SAT-KS-MAT tridomain. It is well established that PksA produces noranthrone (1), which is undetectable in the current assay due to low solubility, from the condensation and cyclization of hexanoyl with seven units of malonyl.<sup>9</sup> Thus, the linear octaketide poly- $\beta$ -ketone intermediate has a C<sub>20</sub> chain length. The PksA PT domain catalyzes C4–C9 and C2–C11 aldol cyclizations and the PksA TE domain catalyzes release through C14–C1 Claisen cyclization. Reactions lacking the

PksA TE domain will make the spontaneous release product norpyrone (19, Table 4).<sup>9</sup> Furthermore, the derailment products of the minimal PksA have been characterized as derivatives of SEK4 (7) and SEK4b (8) carrying a hexanoyl side chain derived from the starter unit (hex-SEK4 20, hex-SEK4b 21).<sup>27</sup> Control reactions consisting of a cognate PksA system form these expected products (Figure 7B).

Of all the noncognate pairs with PksA SAT-KS-MAT, only ACAS (C<sub>16</sub>) was able to complement catalysis. Reactions of PksA SAT-KS-MAT with ACAS PT and ACP led to the formation of a pannorin-like molecule that we dubbed hex-pannorin (22, Figure 7C). Hex-pannorin displays the same spectral features as pannorin (23), as is to be expected for products sharing the same core architecture (Figure 8).<sup>38</sup> Interestingly, hex-pannorin arises from a chain length one extension unit shorter (C<sub>18</sub>) than expected from the parent PksA SAT-KS-MAT, implicating ACAS PT in some level of chain-length control. This product would arise from proper C6–C11 and C4–C13 aldol cyclizations by the ACAS PT domain followed by spontaneous pyrone formation with release from the enzyme (Table 4). A similar reaction occurs with the

noncognate pair of CTB1 SAT-KS-MAT with ACAS PT and ACP to make pannorin itself (Figure S7).<sup>18</sup> Previous evidence suggests that the minimal PksA generates its linear C<sub>20</sub> intermediate in a highly processive fashion.<sup>27</sup> These data imply that the ACAS PT can sample maturing chain-length intermediates, particularly those approaching the full, native chain length, a result that is in keeping with other noncognate pairs described above. The low level of production could signify a slow off-loading rate, a feature that could also explain the decreased amounts of hex-SEK4 and hex-SEK4b derailment products. Intriguingly, the addition of ACTE to the system rescued the PksA KS chain-length control and led to the formation of a C<sub>20</sub> product (Figure 7D) displaying a spectrum consistent with the atrochrysone core architecture (Figure 5).<sup>33</sup> Assuming native processing of the C<sub>20</sub> linear intermediate by the ACAS PT and ACTE domains, the resulting product would be an atrochrysone carboxylic acid analogue bearing a hexanoyl side chain (24, Table 4). Additionally, we observed a species 25 analogous to atrochrysone, likely arising from decarboxylation of acid 24 and displaying a characteristic spectrum for a molecule bearing the atrochrysone core architecture (Figure 5 and Table 4).

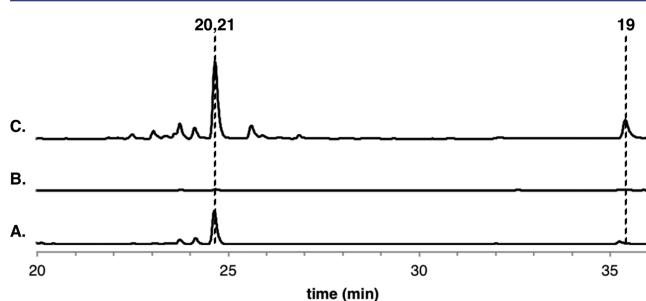
It is noteworthy that combinations of PksA SAT-KS-MAT containing the PT, ACP,<sub>n</sub>, and TE domains of either Pks4 or CTB1 led to an elimination of nearly all production, even shunt products (Figure 7D). This observation points toward the intrinsic editing role for all TE domains of these PKSs. As has been described for PksA, the TE domains of NR-PKSs not only catalyze the final release of product but also intervene when catalysis has been stalled.<sup>27</sup> Derailment products are expected to accumulate on the ACP domain of a given NR-PKS if they do not have an efficient, spontaneous off-loading mechanism. The TE domain can catalyze the hydrolysis of these products, freeing the ACP active site for another round of productive catalysis. It is likely that in the case of the PksA/Pks4 or PksA/CTB1 noncognate pairs, the hexanoyl-loaded ACPs appear as improperly loaded, derailment products to the noncognate TEs. Thus, they hydrolyze the starter unit more rapidly than extension can occur, thereby shutting down the catalytic cycle at the initiation stage.

These competing kinetic processes are complicated by reactions in which the ACP is held cognate with the SAT-KS-MAT tridomain. In reactions of PksA SAT-KS-MAT and ACP with CTB1 (C<sub>14</sub>) PT and TE, production is restored (Figure 9). Remarkably, this noncognate system is proficient in norpyrone (19) formation as well as enhanced in shunt production. Norpyrone would be the product of proper CTB1 PT and TE-mediated cyclization, but on the C<sub>20</sub>-PksA linear

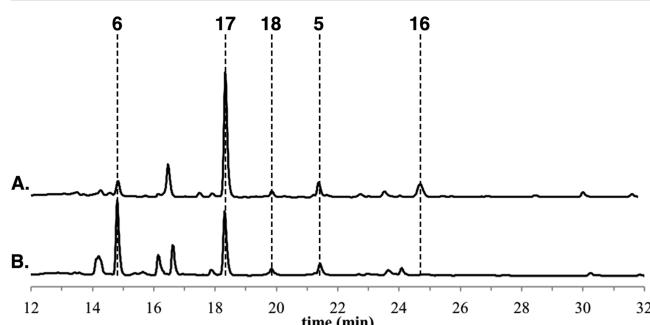
intermediate, a 43% longer chain than the native CTB1 intermediate. Although seemingly contradictory to the observed CTB1 TE editing of the previous reaction, this result is entirely consistent with the current understanding of TE editing. Editing must always be negotiated through the ACP domain. The noncognate PksA ACP and CTB1 TE pair likely interacts poorly relative to the cognate pair. Moreover, it is established that the hexanoyl-loaded PksA ACP initiates rapid and processive extension by the PksA SAT-KS-MAT, meaning the CTB1 TE cannot compete with the cognate minimal PksA. Overall, the competing interdomain kinetics would imply an enhanced efficiency toward extension with the CTB1 TE editing occurring later in the catalytic cycle after the C<sub>20</sub> linear intermediate has been constructed. Thus, the editing by the CTB1 TE domain is suppressed through effective native substrate channeling, allowing for later stage redirection of chemistry toward norpyrone.

**Reactions of Reassembled and Chimeric NR-PKSs.** In general, soluble full-length NR-PKSs are difficult to obtain by heterologous expression in *E. coli*, a key experimental advantage of more reliably expressed, smaller dissected fragments for the deconstruction approach. The rapid reconstitution of component parts exemplified in 72 native and non-native combinations described in this study is a second experimental advantage to both deduce the function of each domain and observe what limits are imposed by their heterocombination on product determination and overall flux. While not the focus of this paper, we prepared intact Pks1 and a CTB1–Pks1 chimera to assess the penalty deconstruction has on net synthetic efficiency.

Equivalent reactions containing either 10 μM full-length Pks1 or 10 μM each of Pks1 SAT-KS-MAT, PT, ACP<sub>2</sub>, and TE (4-part combination) were compared and found to each produce flaviolin (17) as a major product (Figure 10).



**Figure 9.** Combinatorial reactions containing PksA SAT-KS-MAT with the addition of (A) PksA PT, ACP, TE; (B) CTB1 PT, ACP<sub>2</sub>, TE; and (C) CTB1 PT, PksA ACP, CTB1 TE.

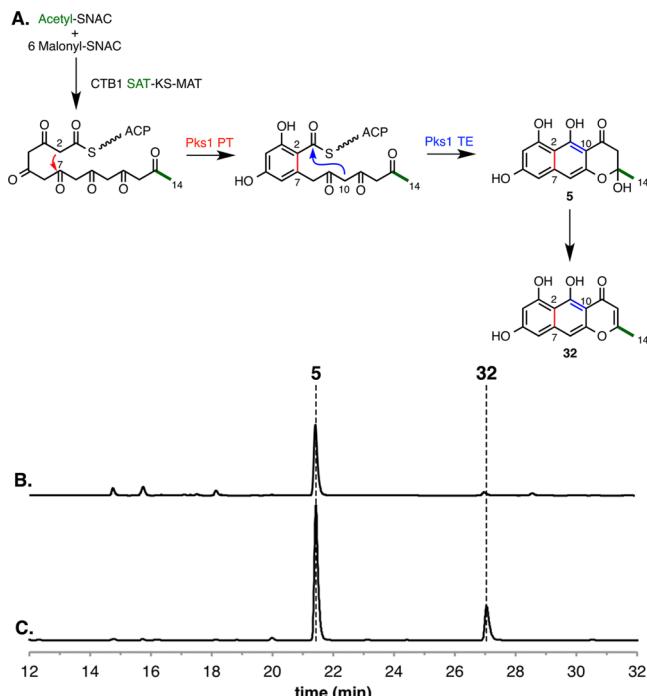


**Figure 10.** Reactions of (A) deconstructed Pks1 (SAT-KS-MAT + PT + ACP<sub>2</sub> + TE) and (B) intact Pks1.

Additionally, THN (6) remained the principal product for intact Pks1. Because THN oxidizes to flaviolin spontaneously, quantification of their relative productivity can be approximated. Interestingly, ATHN (16) is not observed in the intact system, indicating that it is a more efficient enzyme. Derailment products 18 and YWA1 (5) are observed in comparable yields for both systems. Additionally, the intact Pks1 was able to produce flaviolin in reactions containing 1 μM of protein. At these concentrations deconstructed systems do not produce any detectable products. These results are in keeping with previous experiments with deconstructed NR-PKSs where it was shown for PksA that the extent of deconstruction is inversely related to overall productivity.<sup>9</sup> Taken together, these

observations suggest that at the extreme of *holo*-Pks1 and its 4-part reconstitution at 10  $\mu$ M, a penalty of only 2- to 3-fold is exacted but will be less as 3- or 2-part combinations.<sup>9,24</sup> All remain faithful in product synthesis.

In a further test of the penalty for deconstruction, we prepared a chimeric, full-length NR-PKS harboring CTB1 SAT-KS-MAT and Pks1 PT-ACP<sub>2</sub>-TE. While this protein, dubbed M4P6, was soluble, we were hampered by low yields. Nevertheless, we were able to conduct *in vitro* reactions with 1  $\mu$ M M4P6. The chimeric protein produced YWA1 (**5**) in much higher yields than the deconstructed system, which contained 10 times the concentration of reconstituted protein domains (Figure 11). The dehydration product of YWA1, *nor*-



**Figure 11.** Formations of (A) YWA1 (**5**) and *nor*-rubrofusarin (**32**) by (B) 10  $\mu$ M each CTB1 SAT-KS-MAT and Pks1 PT, ACP<sub>2</sub>, and TE; and (C) 1  $\mu$ M M4P6.

rubrofusarin (**32**), was the only other product observed in the reaction of M4P6. Assuming product yield scales linearly with enzyme concentration under a given set of reaction conditions and excess substrates, this outcome represents a *ca.* 20-fold increase in YWA1 production by M4P6 over the 4-part deconstructed system as estimated by HPLC.

## DISCUSSION

The central facet of successful combinatorial biosynthesis in NR-PKSs revealed in this study is that production of a non-native species is contingent upon enzyme-directed cyclization and release outpacing spontaneous reactions and TE-mediated editing. Therefore, an intimate understanding of the underlying mechanisms and relative rates governing the NR-PKS catalytic cycle is crucial for designing an efficient combinatorial enzyme. Recently, it was shown in our laboratory that fidelity of biosynthesis in PksA has been attributed to three main characteristics: processivity of extension, coordinated domain interactions leading to balanced active site occupancy, and TE-mediated editing of spurious intermediates.<sup>27</sup> The results of the

current study suggest that these features may be generally true for the NR-PKSs, as successful complementation required the coordinated control of each of these mechanisms, while failure of just one led to the derailment of efficient redirected catalysis.

The complete interchangeability of ACPs in the minimal NR-PKS underscores the importance of the processivity of extension in these enzymes. Helix II of the ACP has been implicated in guiding client domain interactions.<sup>39</sup> While conservation of these recognition residues may explain some noncognate compatibility, the uniform behavior of the mixed minimal systems with respect to the identity of the SAT-KS-MAT tridomain demonstrates the importance of rapid and complete extension. In PksA, transfer of the starter unit from the SAT domain to the ACP is both slow and crucial for proper extension.<sup>27</sup> It is not until the hexanoyl starter unit is delivered to the KS active site via the ACP that extension commences despite the MAT domain maintaining a high steady-state level of malonyl occupancy. Indeed, when acetyl and malonyl are improperly loaded on the ACP, extension will not begin, a behavior rationalized as being mediated through negative cooperation between the SAT and KS domains. However, once the correct hexanoyl starter unit is loaded onto the PksA KS, extension to the full-length C<sub>20</sub> linear poly- $\beta$ -ketone occurs rapidly and without detectable accumulation of intermediate chain lengths.<sup>27</sup> The noncognate minimal NR-PKSs show similar activity. ACPs that normally accept acetyl starter units will accept hexanoyl starter units and *vice versa*. In either case, extension is complete, and truncation products of intermediate chain lengths are not observed. It is fortuitous that we carried out all our reactions with an intact SAT-KS-MAT tridomain fragment from a single parent NR-PKS, as the interaction of these domains govern correct extension.<sup>27</sup> We hypothesize that, at the very least, a matched SAT and KS pair is crucial in the development of a noncognate NR-PKS. All efforts in our laboratory to express and purify a deconstructed SAT-KS didomain or KS monodomain have failed to date, making this hypothesis unverifiable by *in vitro* recombination.

This is not to say the KS domain exclusively controls the chain length. Early studies of combinatorial synthesis in bacterial type II PKSs also yielded contradictory results.<sup>16</sup> While the majority of minimal type II PKSs demonstrated complete chain-length control,<sup>40</sup> the *Streptomyces coelicolor* spore pigment *whiE* minimal PKS produced a wide assortment of polyketide products ranging in length from 14 to 24 carbons.<sup>15</sup> This observation was in contrast to the prevailing hypothesis that chain length in type II PKSs was controlled by the minimal PKS, consisting of a KS, a chain-length control factor, and an ACP.<sup>40</sup> In the case of *whiE*, an aromatase/cyclase (ARO/CYC) was implicated in chain-length stabilization, extending the role for downstream, tailoring enzymes in the central process of polyketide extension.<sup>41</sup> In NR-PKSs, the PT domain may exert similar control of chain length by binding and sampling growing chains, especially as they near their mature “programmed” length, as has been suggested for the PksA PT.<sup>9,27</sup> Acyl-ACP species have been shown to negotiate client domain interactions in modular PKSs and FASs.<sup>42,43</sup> Indeed, in some of the combinatorial reactions containing the CTB1 and ACAS PT domains, the noncognate PT influenced the final chain length, directing formation of a longer or shorter intermediate, respectively.

Efficient redirection of chemistry depends on the PT domain not only capturing a linear intermediate but also catalyzing the correct aldol cyclization(s). The cyclization register in

successful combinations is always set by the PT domain and determined by the acyl-ACP thioester benchmark, as has been previously observed.<sup>18,44</sup> PT domains, with some exceptions (e.g., PksA PT), display a degree of substrate promiscuity and can proficiently accept chains with plus or minus one extension unit. This permissiveness is in keeping with previous investigations of NR-PKS PTs.<sup>18,20</sup> Some PTs however can also accept linear substrates with dramatically different chain lengths. Surprisingly, this level of promiscuity is usually in the direction of longer intermediates implying that some PT active sites are unexpectedly malleable. The PksA PT active site was completely incompatible with any noncognate intermediate, suggesting that the starter unit effect in this case is carried over to the cyclization stage of biosynthesis. The PksA PT structure contains a hexyl-binding region shown to be crucial for catalytic cyclization.<sup>35</sup> It is likely that acetyl-initiated linear poly- $\beta$ -ketones cannot bind this site, terminating any catalytic redirection by the PksA PT toward these intermediates.

The TE domain is the most crucial of all. As has been demonstrated in multiple NR-PKSs, the TE domain has roles both biosynthetic and editorial.<sup>10,11,24,27</sup> The transesterification of PT-cyclized intermediates onto the TE represents a likely slow step in the overall catalytic cycle. However, in most cases the TE-catalyzed reaction is rapid and drives the overall biosynthetic program. This factor is most evident in the reactions of CTB1, where the inclusion of the TE domain enhances *nor*-toralactone production about 50-fold relative to the reaction without the TE domain.<sup>11</sup> Without the TE domain, off-loading is inefficient, effectively stalling the catalytic cycle and allowing for the accumulation of derailment products. Furthermore, the TE domain will only process intermediates composed of their native substrate's core cyclization architecture. As has been recently suggested,<sup>45</sup> the TE domain serves as the catalytic lynchpin that determines success or failure in biosynthetic redirection.

The editing role of the TE domain of NR-PKSs has only recently been appreciated; however, it is essential for maintaining high catalytic turnover.<sup>27</sup> By intercepting spurious derailment products, the TE domain clears the ACP domain allowing for it to re-enter the kinetically more productive, programmed catalytic cycle. Nowhere is this effect more cleanly demonstrated than with the combinatorial reactions of PksA and CTB1 (Figure 9). The CTB1 TE domain recognizes the hexanoyl loaded CTB1 ACP as an aberrant species and effectively hydrolyses it, thereby shutting down any catalysis at an early stage. The fact that changing the identity of the ACP relative to the TE rescues extension highlights a fundamental factor that is likely at play in all reactions. TE-mediated editing can effectively shut down redirection of chemistry at any stage in the catalytic cycle. Furthermore, a desired intermediate may be deemed unsuitable by the TE domain and eliminated, compromising the success of the combinatorial approach to the synthesis of unnatural polyketides. Thus, the kinetics of redirected synthesis must outpace the rate of TE editing. These interactions are difficult to measure in these multistep processes but are the theoretical factors determining combinatorial success and failure.

The deconstruction approach brings practical advantages of generally reliable protein overproduction in *E. coli* and the rapid combinatorial assembly of functional units into reconstituted native and non-native catalytic systems. Although not assured at the outset, the reconstituted NR-PKSs have been found to recapitulate their wild-type synthetic capabilities. Similarly,

heterocombination of catalytic components reassemble to elicit chemistry in a largely predictable fashion as we demonstrate here. While examples are limited, as one might intuitively expect, *holo* native and chimeric NR-PKSs show improved synthetic efficiency over their reconstituted component parts and give a reduced incidence of shunt and truncated products. Although only two cases are compared, the data suggest that these improvements will be greater for reconstituted heterodomains over native systems. These observations auger well for their rational assembly in single multidomain designed proteins for non-natural product synthesis.

## CONCLUSIONS

In summary, we find that noncognate "domain swapped" NR-PKSs will redirect biosynthesis to non-native products if certain underlying constraints governing NR-PKS programming are met. While ketide extension appears to be complete in all noncognate reactions, the interplay of downstream domains (namely the PT and TE domains) determines catalytic turnover to non-native products. When these domains capture suitable intermediates and efficiently carry out their programmed catalysis, redirection of chemistry takes place. Effective redirection requires that these processes outpace the rate of spontaneous derailments and TE-mediated editing. For more effective engineered NR-PKSs, a greater understanding of the mechanisms and internal kinetics of the catalytic program is required.

## ASSOCIATED CONTENT

### S Supporting Information

Cloning details, product characterization, and tables and figures including protein details, noncognate minimal NR-PKS product profiles, noncognate CTB1 SAT-KS-MAT product profiles, and product spectra. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

ctownsend@jhu.edu

### Present Address

<sup>†</sup>ETH Zürich, Institute of Microbiology, HCI G 428, Vladimir-Prelog-Weg 1–5/10, 8093 Zürich, Switzerland

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors thank Dr. Katherine L. Fiedler and Prof. Dr. Robert J. Cotter of the Middle Atlantic Mass Spectrometry Laboratory at the Johns Hopkins School of Medicine and the Mid-Atlantic Regional Office of Shimadzu Scientific Instruments, Inc. (Columbia, MD) for use of their Shimadzu LC-IT-TOF. We thank Dr. Katherine Belecki (Johns Hopkins University, Baltimore, MD) for aid in collecting mass spectra. An authentic sample of pannorin was generously provided by Prof. Keiji Hasumi (Tokyo University of Agriculture and Technology, Tokyo, Japan). An authentic sample of 6-(2',4'-dihydroxy-6'-methylphenyl)-4-hydroxy-2-pyrone was generously provided by Prof. Ikuro Abe. This work was supported by NIH Grant ES001670.

## REFERENCES

- (1) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2012**, *75*, 311–335.

- (2) Floss, H. G. *J. Biotechnol.* **2006**, *124*, 242–257.
- (3) Wilkinson, B.; Micklefield, J. *Nat. Chem. Biol.* **2007**, *3*, 379–386.
- (4) Staunton, J.; Weissman, K. *J. Nat. Prod. Rep.* **2001**, *18*, 380–416.
- (5) Crawford, J. M.; Townsend, C. A. *Nat. Rev. Microbiol.* **2010**, *8*, 879–889.
- (6) Cox, R. J.; Simpson, T. J. In *Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates*; Hopwood, D. A., Ed.; Elsevier Academic Press Inc: San Diego, CA, 2009; Vol. 459, pp 49–78.
- (7) Maier, T.; Leibundgut, M.; Ban, N. *Science* **2008**, *321*, 1315–1322.
- (8) Sanchez, J. F.; Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Ahuja, M.; Oakley, C. E.; Bok, J. W.; Keller, N.; Oakley, B. R.; Wang, C. C. *Mol. BioSyst.* **2010**, *6*, 587–593.
- (9) Crawford, J. M.; Thomas, P. M.; Scheerer, J. R.; Vagstad, A. L.; Kelleher, N. L.; Townsend, C. A. *Science* **2008**, *320*, 243–246.
- (10) 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.
- (11) Newman, A. G.; Vagstad, A. L.; Belecki, K.; Scheerer, J. R.; Townsend, C. A. *Chem. Commun.* **2012**, *48*, 11772–11774.
- (12) Udwyar, D. W.; Merski, M.; Townsend, C. A. *J. Mol. Biol.* **2002**, *323*, 585–598.
- (13) Khosla, C.; Kapur, S.; Cane, D. E. *Curr. Opin. Chem. Biol.* **2009**, *13*, 135–143.
- (14) Pickens, L. B.; Tang, Y.; Chooi, Y. H. In *Annual Review of Chemical and Biomolecular Engineering*, Vol 2; Prausnitz, J. M., Ed.; Annual Reviews: Palo Alto, CA, 2011; Vol. 2, pp 211–236.
- (15) Shen, Y. M.; Yoon, P.; Yu, T. W.; Floss, H. G.; Hopwood, D.; Moore, B. S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3622–3627.
- (16) Moore, B. S.; Piel, J. *Antonie Van Leeuwenhoek* **2000**, *78*, 391–398.
- (17) Tang, Y.; Lee, T. S.; Lee, H. Y.; Khosla, C. *Tetrahedron* **2004**, *60*, 7659–7671.
- (18) Vagstad, A. L.; Newman, A. G.; Storm, P. A.; Belecki, K.; Crawford, J. M.; Townsend, C. A. *Angew. Chem., Int. Ed. Engl.* **2013**, *52*, 1718–1721.
- (19) Choquer, M.; Dekkers, K. L.; Chen, H. Q.; Cao, L. H.; Ueng, P. P.; Daub, M. E.; Chung, K. R. *Mol. Plant-Microbe Interact.* **2005**, *18*, 468–476.
- (20) Li, Y. R.; Xu, W.; Tang, Y. *J. Biol. Chem.* **2010**, *285*, 22762–22771.
- (21) Ma, S. M.; Zhan, J.; Watanabe, K.; Xie, X.; Zhang, W.; Wang, C. C.; Tang, Y. *J. Am. Chem. Soc.* **2007**, *129*, 10642–10643.
- (22) Awakawa, T.; Yokota, K.; Funa, N.; Doi, F.; Mori, N.; Watanabe, H.; Horinouchi, S. *Chem. Biol.* **2009**, *16*, 613–623.
- (23) Fujii, I.; Watanabe, A.; Sankawa, U.; Ebizuka, Y. *Chem. Biol.* **2001**, *8*, 189–197.
- (24) Vagstad, A. L.; Hill, E. A.; Labonte, J. W.; Townsend, C. A. *Chem. Biol.* **2012**, *19*, 1525–1534.
- (25) Sambrook, J.; Russell, D. W. *Molecular cloning: A laboratory manual*; Cold Spring Harbor Laboratory Press, Plainview, NY, 2001.
- (26) An, J. H.; Kim, Y. S. *Eur. J. Biochem.* **1998**, *257*, 395–402.
- (27) Vagstad, A. L.; Bumpus, S. B.; Belecki, K.; Kelleher, N. L.; Townsend, C. A. *J. Am. Chem. Soc.* **2012**, *134*, 6865–6877.
- (28) Jiang, H.; Zirkle, R.; Metz, J. G.; Braun, L.; Richter, L.; Van Lanen, S. G.; Shen, B. *J. Am. Chem. Soc.* **2008**, *130*, 6336–6337.
- (29) Fu, H.; Ebertkhosla, S.; Hopwood, D. A.; Khosla, C. *J. Am. Chem. Soc.* **1994**, *116*, 4166–4170.
- (30) Fu, H.; Hopwood, D. A.; Khosla, C. *Chem. Biol.* **1994**, *1*, 205–210.
- (31) Zhang, W. J.; Li, Y. R.; Tang, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 20683–20688.
- (32) Ma, S. M.; Zhan, J. X.; Xie, X. K.; Watanabe, K. J.; Tang, Y.; Zhang, W. J. *J. Am. Chem. Soc.* **2008**, *130*, 38–39.
- (33) Gill, M.; Morgan, P. M. *ARKIVOC* **2001**, *2001*, 145–156.
- (34) Crawford, J. M.; Dancy, B. C. R.; Hill, E. A.; Udwyar, D. W.; Townsend, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16728–16733.
- (35) 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–U1243.
- (36) Abe, I.; Oguro, S.; Utsumi, Y.; Sano, Y.; Noguchi, H. *J. Am. Chem. Soc.* **2005**, *127*, 12709–12716.
- (37) Springob, K.; Samappito, S.; Jindaprasert, A.; Schmidt, J.; Page, J. E.; De-Eknamkul, W.; Kutchan, T. M. *FEBS J.* **2007**, *274*, 406–417.
- (38) Ogawa, H.; Hasumi, K.; Sakai, K.; Murakawa, S.; Endo, A. J. *Antibiot.* **1991**, *44*, 762–767.
- (39) Wattana-Amorn, P.; Williams, C.; Ploskon, E.; Cox, R. J.; Simpson, T. J.; Crosby, J.; Crump, M. P. *Biochemistry* **2010**, *49*, 2186–2193.
- (40) McDaniel, R.; Ebertkhosla, S.; Fu, H.; Hopwood, D. A.; Khosla, C. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11542–11546.
- (41) Yu, T. W.; Shen, Y. M.; McDaniel, R.; Floss, H. G.; Khosla, C.; Hopwood, D. A.; Moore, B. S. *J. Am. Chem. Soc.* **1998**, *120*, 7749–7759.
- (42) Tran, L.; Tosin, M.; Spencer, J. B.; Leadlay, P. F.; Weissman, K. *J. ChemBioChem* **2008**, *9*, 905–915.
- (43) Ploskon, E.; Arthur, C. J.; Kanari, A. L. P.; Wattana-amorn, P.; Williams, C.; Crosby, J.; Simpson, T. J.; Willis, C. L.; Crump, M. P. *Chem. Biol.* **2010**, *17*, 776–785.
- (44) Xu, Y. Q.; Zhou, T.; Zhou, Z. F.; Su, S. Y.; Roberts, S. A.; Montfort, W. R.; Zeng, J.; Chen, M.; Zhang, W.; Lin, M.; Zhan, J. X.; Molnar, I. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 5398–5403.
- (45) Xu, Y.; Zhou, T.; Zhang, S.; Xuan, L.-J.; Zhan, J.; Molnar, I. *J. Am. Chem. Soc.* **2013**, *135*, 10783–10791.