JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Rational Domain Swaps Decipher Programming in Fungal Highly Reducing Polyketide Synthases and Resurrect an Extinct Metabolite

Katja M. Fisch,[†] Walid Bakeer,[†,‡] Ahmed A. Yakasai,[†] Zhongshu Song,[†] Jennifer Pedrick,[†] Zahida Wasil,[†] Andrew M. Bailey,[§] Colin M. Lazarus,[§] Thomas J. Simpson,[*,†] and Russell J. Cox[*,†]

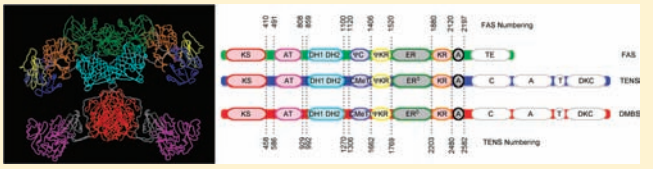
[†]School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, U.K.

[‡]Faculty of Pharmacy, Beni Suef University, Salah Salem Street, Beni Suef 32514, Egypt

[§]School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, U.K.

Supporting Information

ABSTRACT: The mechanism of programming of iterative highly reducing polyketide synthases remains one of the key unsolved problems of secondary metabolism. We conducted rational domain swaps between the polyketide synthases encoding the biosynthesis of the closely related compounds tenellin and desmethylbassianin. Expression of the hybrid synthetases in *Aspergillus oryzae* led to the production of reprogrammed compounds in which the changes to the methylation pattern and chain length could be mapped to the domain swaps. These experiments reveal for the first time the origin of programming in these systems. Domain swaps combined with coexpression of two cytochrome P450 encoding genes from the tenellin biosynthetic gene cluster led to the resurrection of the extinct metabolite bassianin.



INTRODUCTION

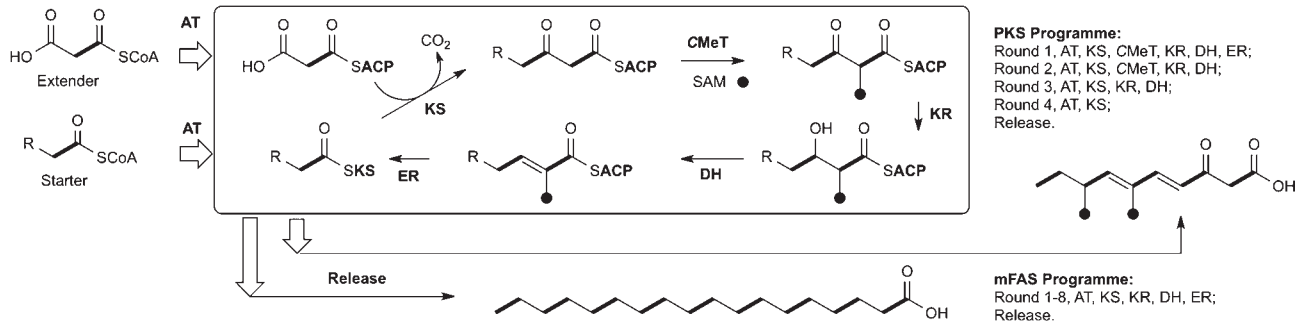
Fungal polyketides are astonishingly diverse natural products ranging from simple aromatic structures to extremely complex compounds.[1] They display many beneficial biological activities, e.g., antibiotic, antifungal, herbicidal, immunomodulation, and cholesterol biosynthesis inhibition, in addition to economically damaging toxicity (mycotoxins). Polyketide biosynthesis is closely related to fatty acid biosynthesis in which the fatty acid synthase (FAS) produces highly reduced carbon chains via acyl CoA starter and malonyl CoA extender units which are loaded onto the acyl carrier protein (ACP) domain of FAS by an acyl transferase (AT).[2] These are then condensed to give a β-ketothiolester in a decarboxylative Claisen reaction catalyzed by a ketosynthase enzyme (KS), followed by sequential β-keto reduction (KR), dehydration (DH), and enoyl reduction (ER) reactions using specialized domains resulting in deoxygenation and saturation of the acyl group (Scheme 1).[2] Polyketide synthases (PKSs) use the same protein machinery to assemble intermediates, but in contrast to FAS their products display very high chemical diversity. This is achieved by highly programmed control over the selection of starter and extender units, the number of condensation cycles, and the extent of the reduction and dehydration reactions following each extension cycle. There are two fundamental paradigms for this assembly. *Modular* PKSs consist of extremely large proteins in which individual modules are responsible for starter unit selection and each round of chain extension, with individual domains within each module controlling the extent of modification.[2] *Iterative* PKSs, on the other hand, consist of a single multidomain module which is responsible for every extension cycle, but programming may be different in each cycle.[1,2] The program of a modular PKS is usually obvious, as it is encoded by the number of modules and the presence or absence of modifying domains in each module. In contrast, the program of an iterative PKS is cryptic, and understanding how it is controlled is one of the remaining major challenges in secondary metabolism research.

Iterative PKSs are common in fungi and have been classified as *non-reducing* (nr-PKS), where there are no modifying reductive reactions after each chain extension, *partially reducing* (pr-PKS), where a single KR reaction occurs during biosynthesis, and *highly reducing* (hr-PKS), where a complex set of reductions and dehydrations are programmed.[1] Additional complexity can be introduced by chain methylation catalyzed by a C-methylation (CMeT) domain (Scheme 1). Substantial progress has been made in exploring the origin of programming in the nr-PKSs that make (poly)cyclic aromatic metabolites, notably through the efforts of Townsend and co-workers, who have mapped the role of domains unique to nr-PKSs in starter unit selection (SAT, starter-unit acyl transferase)[3] and control of chain length and cyclization (PT, product template)[4] in addition to the core KS, AT, and ACP domains.[5] However, no progress has been made in charting the origin of programming in hr-PKSs, where SAT and PT domains are absent and CMeT, KR, DH, and ER present additional elements for programmed control.

Our efforts at understanding programming by fungal hr-PKSs have been focused on the biosynthesis of tenellin[6] (**1**) and desmethylbassianin (DMB, **2**) in the insect pathogen *Beauveria*

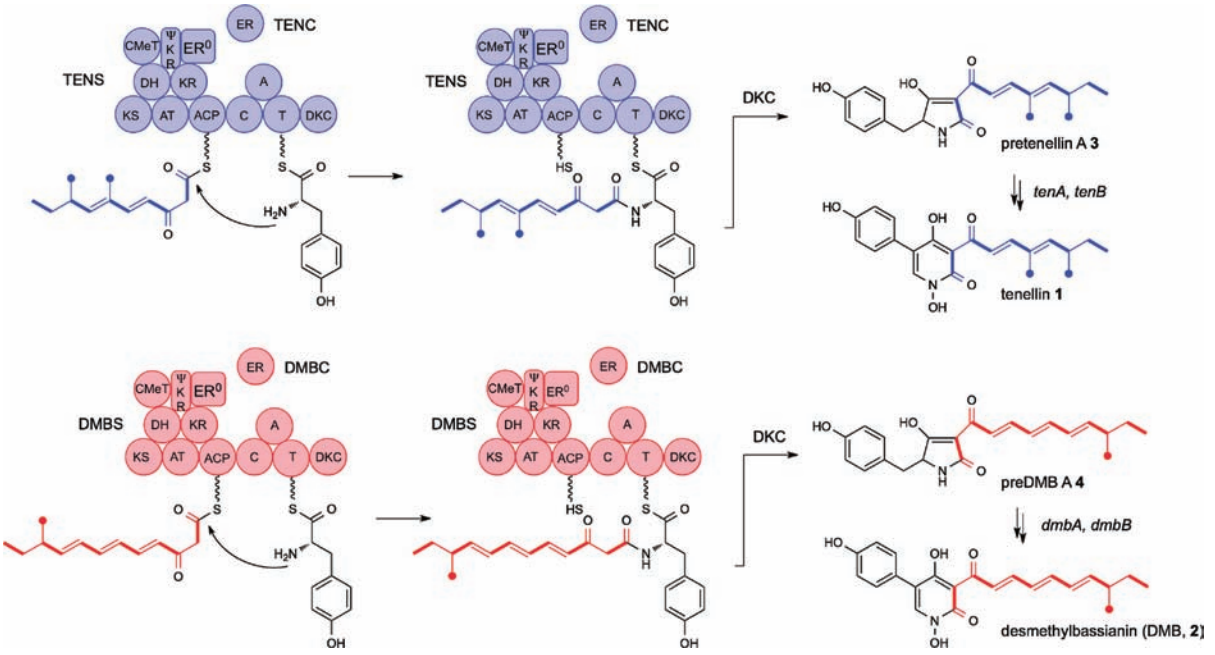
Received: July 24, 2011
Published: September 07, 2011

Scheme 1. Generic Reactions Catalysed by Iterative Fatty Acid Synthases (FASs) and Highly Reducing Polyketide Synthases (hr-PKSs)[a]



[a] Bold bonds indicate incorporation of intact acetate units. Filled circles indicate methyl derived from *S*-adenosylmethionine. Abbreviations are given in the text.

Scheme 2. Compounds Produced by the TENS and DMBS hr-PKS[a]



[a] Bold bonds indicate incorporation of intact acetate units. Filled circles indicate methyl derived from *S*-adenosylmethionine. Abbreviations are given in the text. Ψ-KR represents the KR structural domain as explained in the text.

bassiana.[7] These compounds belong to a wider class of 2-pyridones with interesting neuritogenic properties which are the focus of continuing synthetic efforts.[8] The biosynthesis of **1** and **2** is initiated by a typical fungal hr-PKS which is fused to a single module of a nonribosomal peptide synthetase (NRPS) consisting of condensation (C), adenylation (A), thiolation (T), and Dieckmann cyclase (DKC)[9−11] domains (Scheme 2). The NRPS acts as an efficient PKS release mechanism.[12] Previous work has shown that the NRPS does not influence the program of the PKS in the cases of **1** and **2** biosynthesis.[7] The fused tenellin PKS-NRPS is encoded by the *tenS* gene, and the protein is known as tenellin synthetase (TENS, 460.3 kDa), which directs the biosynthesis of pretenellin A **3**, a doubly methylated pentaketide fused to tyrosine.[13,14] Two oxidations convert the tetramic acid pretenellin A **3** to the *N*-hydroxypyridone tenellin **1**.[15] In predesmethylbassianin A (pre-DMB A, **4**), the programmed polyketide is a singly methylated hexaketide.[7] Thus, TENS and DMBS differ in programming with respect to both methylation pattern and chain length. The TENS and DMBS ER domains, like many fungal PKS ER domains, are inactive (ER^0), and the required reductive activities are supplied by the *tenC*- and *dmbC*-encoded *trans*-acting ERs of ca. 41.3 kDa. Previous work has shown that the *tenC*- and *dmbC*-encoded *trans*-acting ERs are interchangeable.[7]

Fungal hr-PKSs share a remarkable sequence, and presumably structural, similarity to vertebrate FASs (despite the separation of fungi and vertebrates during 500 million years of evolution) and are more closely related to vertebrate FASs than to fungal FASs. Ban and co-workers have provided a high-resolution structure of mammalian FAS (mFAS, 3.2 Å, Figure 1A)[16] which shows a

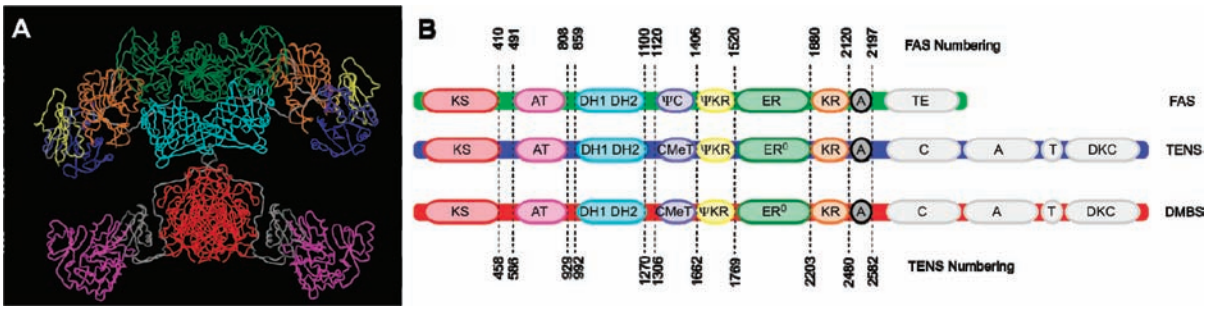


Figure 1. (A) Domain architecture of mammalian fatty acid synthase (mFAS): red, KS; magenta, AT; cyan, DH; blue, ΨCMeT; yellow, ΨKR; green, ER; orange, KR. Abbreviations are given in the text. Constructed using coordinates from Maier et al.[16] (B) Alignment between mFAS and TENS and DMBS (not to scale). Domain designations are as in text except A on dark gray background = ACP.

domain structure consisting of KS, AT, DH, ΨCMeT (inactive because of the lack of a SAM binding motif), ΨKR (a structural domain of the KR), ER, and KR. The ACP and TE (thiolesterase) domains do not appear in the crystal structure, and this was hypothesized to be due to the flexibility of these domains. Multiple sequence alignment between mFAS and hr-PKS, such as the PKS regions of TENS and DMBS (Figure 1B; see Supporting Information, section 3), shows that the location and order of the PKS catalytic domains closely mirrors that of mFAS, and that key structural and catalytic motifs are conserved. However, the hr-PKS catalytic domains are generally larger than their mFAS counterparts, and these differences are not yet understood. Sequence similarity for catalytically *inactive* domains such as ΨCMeT and ΨKR is relatively low, and in the experiments described below we treated the putative CMeT and ΨKR regions of TENS and DMBS as a single unit. Thus, the mFAS structure serves as a preliminary model for the structure of hr-PKS and guided our decisions about domain boundaries.

The key questions in programming of hr-PKSs are control of starter unit selection, normally acetate but other alkyl and aryl CoAs can be used; control of chain length, i.e., the number of assembly cycles (known to vary from one, e.g., lovastatin diketide synthase LDKS, to at least eight, e.g., lovastatin nonaketide synthase LNKS);[17] and the degree of *C*-methylation, keto reduction, dehydration, and enoyl reduction within each cycle. Ultimately control must reside in the structure of the protein and recognition of structurally ever-changing substrates (assembly intermediates) by some, or all, of the catalytic domains. Bioinformatic analyses of the overall PKS and individual domain sequences from hr-PKS have failed to provide any substantive information on where control might reside in the protein. We therefore chose to adopt the empirical but rational approach of swapping domains between closely related synthases that make similar but subtly different structures with the aim to begin to map the domains and subdomains that exercise demonstrable control. A rational domain swap approach, as has been successful in modular PKSs,[18] could provide answers to key questions about control of the modification pattern and chain length of the polyketides synthesized. We now report results of a systematic domain swapping approach applied to the closely related TENS and DMBS hr-PKS which have allowed methylation patterns and chain length to be altered predictably for the first time. The results of these experiments give the first insight into the mechanisms used by hr-PKS to control programming.

RESULTS

We used a procedure involving homologous recombination in yeast to construct chimaeric *tenS*/*dmbS* genes, which were then transferred to a fungal expression vector. In general this involved excising a fragment of *tenS* and replacing it with an equivalent fragment from *dmbS* (see Supporting Information for details). Thus, *tenS* should be regarded as the acceptor and *dmbS* as the donor. Reconstructed genes were shuttled from yeast to *Escherichia coli* and then transferred to the *amyB* expression cassette by Gateway LR recombination, and the final expression vectors were used to transform the fungal host *Aspergillus oryzae*. To provide the required ER functions, separate cotransformations were performed with *tenC* or *dmbC*. No differences in titers or product distributions were observed between cotransformants with *tenC* or *dmbC*, consistent with our previous observations that TENC and DMBC can be interchanged without affecting PKS-NRPS programming.[7] In an initial series of experiments, donor fragments from *dmbS* were swapped into *tenS* starting from the 5′ end corresponding to the KS and AT domains (Figure 2B). We then extended the swap up to the end of each following domain until the whole *tenS* PKS was replaced by the *dmbS* sequence (Figure 2C−G).

Donating KS-AT or KS-AT-DH sequences from *dmbS* to *tenS* had no effect on either the titer or the structure of the product produced by the hybrid synthetase in *A. oryzae*; wild type (Figure 2A) and both chimeras (Figure 2B,C) produced only pretenellin A **3** (Figure S1). Extending the sequence to include CMeT-ΨKR (Figure 2D), however, led to the production of a new compound with a mass of 342.1698 ([M]H$^+$), corresponding to a molecular formula of $C_{20}H_{23}NO_4$. This was purified and its structure determined by full NMR analysis to be desmethylpretenellin A **5**, which has the chain length of pretenellin A **3** but the single methylation frequency and position of pre-DMB A **4**. This chimera also produced pretenellin A **3**, but only as a minor product.

The further donation of ER^0 from *dmbS* into *tenS* also gave the monomethylated *pentaketide* desmethylpretenellin A **5** (Figure 2E), but extension to include the KR domain from *dmbS* into *tenS* led to a dramatic change, and this chimera produced the *hexaketide* preDMB A **4** as the exclusive product (Figure 2F). Finally, donation of the entire PKS from *dmbS* into *tenS* also produced only preDMB A **4** (Figure 2G). This is in agreement with our previous finding that the hybrid *tenS*PKS-*dmbS*NRPS chimera produced only pretenellin A **3** and confirms that the NRPS component acts only in the aminoacylation and off-loading of the polyketide and has no role in programming the PKS.[7]

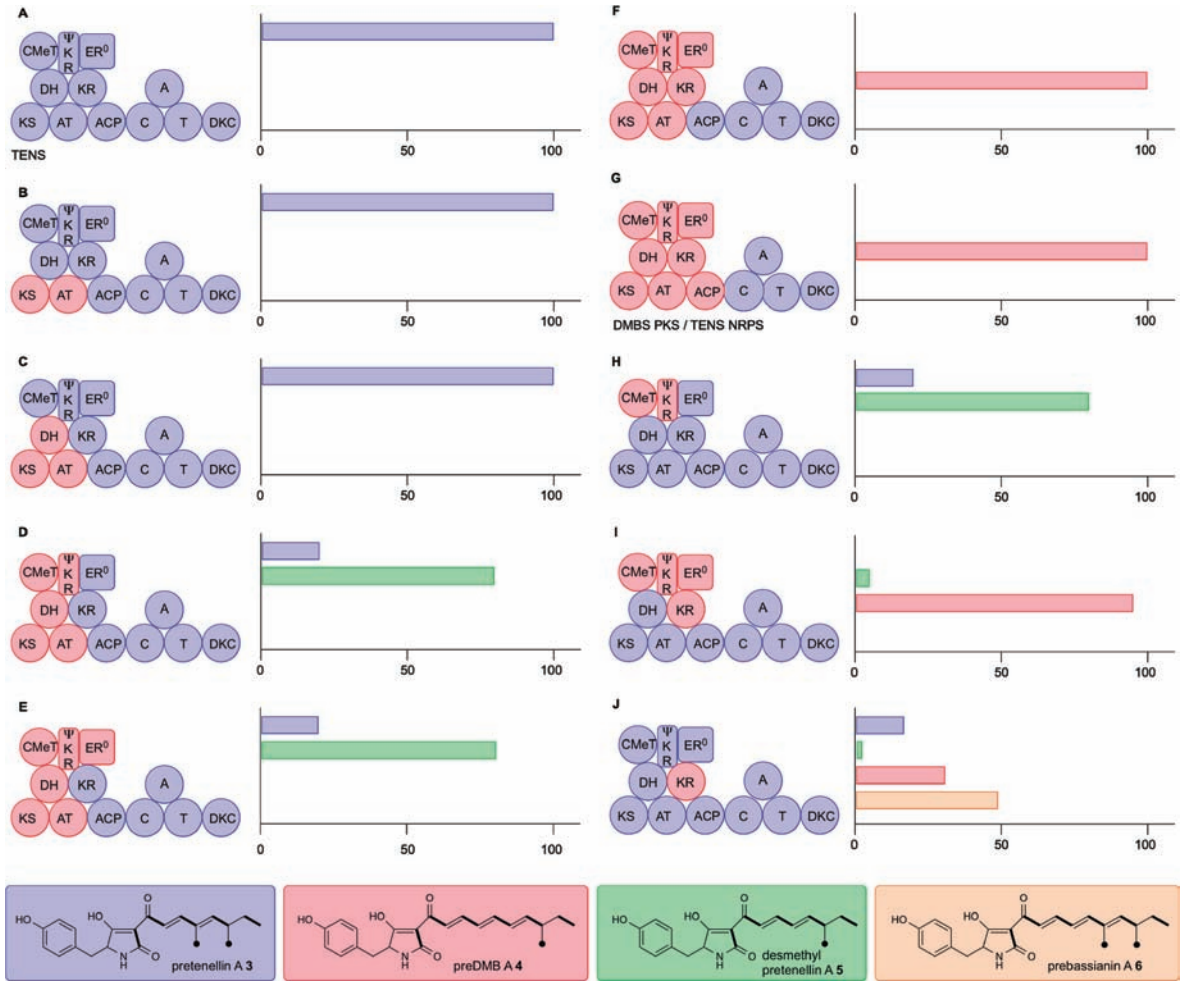


Figure 2. Domain architectures of chimeric PKS-NRPS constructed in this study and relative titers of small-molecule products. See text for abbreviations. No significant variations in titer were observed between WT and hybrid systems in any case.

We next performed more precise domain swaps to firmly establish the role of the CMeT-ΨKR and KR single domains on programming. Thus, the CMeT-ΨKR of *dmbS* was used to replace that in *tenS*, and as shown in Figure 2H this hybrid produced monomethylated pentaketide **5** as the main compound with minor amounts of dimethylated pentaketide **3** (cf. Figure 2D), suggesting that the CMeT-ΨKR didomain strongly controls the methylation pattern but lacks complete fidelity when placed in a different environment. The entire CMeT to KR fragment was then transferred from *dmbS* into *tenS*, and this gave the monomethylated hexaketide preDMB A **4** with a very minor amount (5%) of the corresponding pentaketide desmethylpretenellin A **5** (Figure 2I). Thus, we predicted that in a KR-only swap, chain length should be controlled by the *dmbS* KR, i.e., hexaketide, and methylation by the *tenS* CMeT, i.e., dimethylation. This experiment led to four products (Figures 2J and S1J) in which, gratifyingly, total hexaketides dominated (80%) over pentaketides (20%), and dimethylated products (66%) over monomethylated (34%). They consisted of all three compounds previously observed as well as a major new compound proven by NMR and HRMS ([M]H$^+$ 382.2017) to be prebassianin A **6**, i.e., the doubly methylated hexaketide.

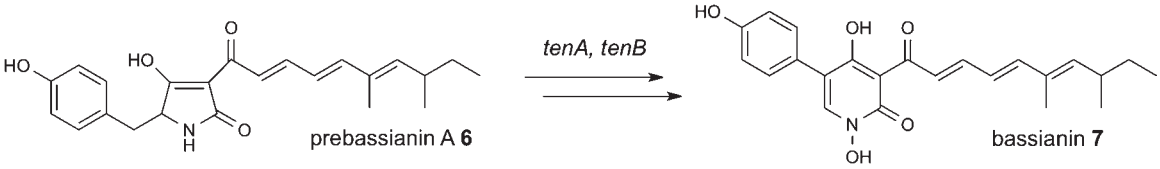
Previously we have reconstructed the entire tenellin biosynthetic gene cluster in the heterologous host *A. oryzae*.[19] This involved expressing *tenS* and *tenC* together with *tenA* and *tenB*, which encode cytochrome P450 monooxygenases (which achieve ring-expansion of pretenellin A and *N*-hydroxylation respectively), to form tenellin **1** (Scheme 2). In order to show that the chimaeric PKS-NRPS genes could also be used in pathway reconstruction, we transformed *A. oryzae* with the *tenS*(ΔKR:*dmbS*-KR) gene along with *tenA*, *tenB*, and *tenC*. LCMS analysis of transformants indicated the production of a new, albeit minor, compound as a mixture with prebassianin A **6** and other related compounds (Figure S1K). This new compound had a *m*/*z* 396 ([M]H$^+$, HRMS 396.1818), and ^{1}H, COSY, HSQC, and HMBC NMR confirmed it to be bassianin **7** (Scheme 3, Figures S2−S5).

In a final experiment to manipulate the domain structure of TENS, we deleted the NRPS module, but expression of the truncated gene encoding TENS PKS alone in *A. oryzae* did not lead to the observation of any new compounds.

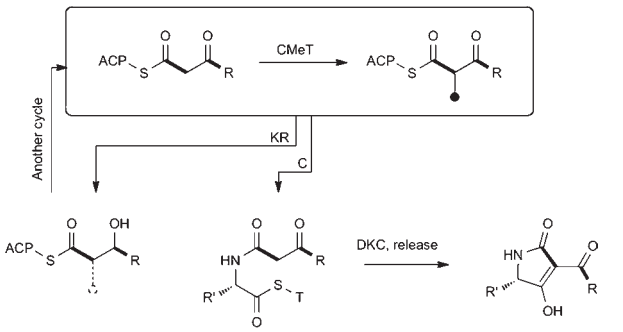
DISCUSSION

These experiments reveal for the first time how the catalytic domains of a pair of iterative fungal hr-PKS interact to control programming. One possible mechanism for the control of chain length would be for the KS or AT to recognize the length of the growing ACP-bound acyl group. In the case of iterative Type II[20] and Type III[21] PKSs, crystal structures of the KS components reveal the presence of binding tunnels within the proteins, which

Scheme 3. Coexpression of *tenS*(ΔKR:*dmbS*-KR) + *tenC* or *dmbC* with *tenA* and *tenB* To Produce Bassianin 7



Scheme 4. Potential Fates of β-Keto Acyl ACP Intermediates



have been postulated to control the length and cyclization of the polyketide intermediate. In these nonreduced systems, there is a prerequisite for stabilization of the highly reactive poly-keto thiolester intermediates in addition to control of cyclization. Neither of these factors applies to the highly reduced systems, so it would be expected that a different mechanism of chain length control would apply. One possibility for this would be for the AT to control programming by not recognizing and transferring a fully grown chain to the KS for further extension; likewise, the KS could either not accept or not extend a fully formed chain. Neither of these possibilities seems to be operative here because the DMBS KS and AT did not alter chain length when fused to the TENS modifying and ACP domains. In the only other report of a rational domain swap between hr-PKS, Du and co-workers showed that the KS domain of T-toxin PKS1 from *Cochliobolus heterostrophus* had no effect on the programming of fumonisin FUM1 when donated *in-cis* in the fungus *Fusarium verticillioides*.[22,23] Similarly, acyl ACP could be hypothesized to control chain length selection if the bound acyl group failed to act as a substrate for KS or AT when it reached the correct length. This mechanism is also not operative because the DMBS domains from KS to KR were sufficient to produce a hexaketide in the presence of the TENS ACP. These results suggest that the KS-AT together with the ACP extend acyl ACPs indiscriminately, leaving the programming functions to other domains of the PKS. It is thought that the KS-AT of mFAS acts similarly, so the chain length is governed by the substrate selectivity of the product-releasing TE.[24] Such a chain-length-dependent release mechanism would be conceivable for TENS and DMBS, but exchange of the NRPS release components did not affect the program of polyketide biosynthesis by TENS or DMBS (Figure 2G).

These results show that, at least in the case of TENS/DMBS, the regions of the hr-PKS responsible for controlling programming must be the modifying domains in the "top" half of the PKS, as depicted in Figure 1A, assuming these have an overall structure similar to that of mFAS. The DH domain might be expected to cause a change in programming because it consists of a double "hot-dog" fold,[16] very similar in architecture to the double hot-dog fold which forms the product template (PT) domain of nr-PKS.[4] In an elegant series of experiments, Townsend and co-workers have shown that the PT domain controls both chain length and cyclization pattern for the construction of nonreduced polyketides because it has an active-site cleft of defined geometry which can accommodate a poly β-keto intermediate.[4] Thus, the hr-PKS DH didomain could also harbor a cleft capable of accommodating the growing acyl group, but no change in product structure resulted from including the DH in swaps, ruling out a role in programming in the hr-PKS of TENS and DMBS.

The first observed change in programming arose from the inclusion of the CMeT-ΨKR region, which resulted in efficient production of desmethylpretenellin A **5**. In this monomethylated pentaketide, the methylation pattern matches that of pre-DMB A **4**, while the rest of the molecule is the same as pretenellin A **3**. Although a small amount of pretenellin A **3** is still produced (20%), this result suggests that the CMeT contributes significantly to its own programming. Thus, it appears that the DMBS CMeT will only recognize a diketide substrate, whereas its TENS counterpart will also recognize a triketide substrate. Other fungal PKS CMeT domains are known to be programmed by recognizing substrate structure. For example, we have shown that the CMeT domain of methylorcinaldehyde synthase (MOS), a nr-PKS from *Acremonium strictum*, must recognize and react with a triketide but not a diketide or tetraketide intermediate.[25,26] We have also made subtle mutations in the CMeT domain of squalestatin tetraketide synthase (SQTKS) which appear to abolish methylation, perhaps by interfering with substrate recognition.[27] Clearly, the selectivities of the TENS and DMBS CMeT domains need to be further probed by *in vitro* studies with expressed CMeT domain. It is perhaps not surprising that C-methylation is under specific CMeT domain control, but despite the high sequence similarity of the TENS and DMBS CMeT domains (87% amino acid identity), the factors controlling substrate recognition cannot be easily determined from simple sequence comparisons. That control of C-methylation resides mainly with the CMeT domain was confirmed by swapping the *dmbS* CMeT domain on its own into the otherwise *tenS* background (Figure 2H). The same result was obtained when the KS-AT-DH-CMeT-ΨKR-ER^0 region of *tenS* was replaced with that from *dmbS*, indicating that the ER^0 domain plays no role in programming (Figure 2E).

The next significant change occurred when the swap was extended to include the *dmbS* KR, which resulted in clean production of **4**, i.e., a change from pentaketide to hexaketide production (Figure 2F). In the case of TENS and DMBS, control of chain length by KR can be rationalized. The β-keto acyl ACP intermediate can be a substrate for three different catalysts (Scheme 4). Methylation of the β-keto thiolester by CMeT (which of course has its own selectivity) would lead to a

dx.doi.org/10.1021/ja206914q |*J. Am. Chem. Soc.* 2011, 133, 16635–16641

methylated β-keto thiolester, which could act as a substrate for ketoreduction to give a β-hydroxy species destined for dehydration and further chain extension. However, the action of the NRPS leads to release as the C-domain creates a substrate which can no longer be extended and which is set up for release by the DKC domain. For many iterative PKSs and FASs, pyrone formation appears to be a default release mechanism when programming has run awry; it has been observed in the cases of mFAS,[28] 6-methylsalicylic acid synthase (MSAS),[29] MOS,[25] LNKS,[17] and aspyridone synthetase (APDS)[30] but does not appear to be operative in the case of TENS, given that the NRPS-truncated PKS did not yield any discernible products. Thus, the only productive reaction when KR fails to act is chain release. We have previously shown that the TENS and DMBS C-domains are unselective, leaving the KR as the only control element for chain length determination in these systems.

While the TENS KR clearly recognizes and reduces β-keto di-, tri-, and tetraketides but not β-keto-pentaketides, the DMBS KR also reduces β-keto-pentaketides but presumably not β-keto-hexaketides. All β-hydroxy products formed in these two systems are dehydrated—and so DH does not require or apparently display selectivity in the case of TENS and DMBS. This was apparent in the *C*MeT-ΨKR-ER-KR swap (*vide infra*), in which both β-hydroxy-tetra- and β-hydroxy-pentaketides (in addition to β-hydroxy-di- and -triketides) were dehydrated by the TENS DH, which normally never encounters a β-hydroxy-pentaketide.

Further evidence for the predominant, but not exclusive, roles of the *C*MeT in control of extent of methylation and of the KR in chain length determination was obtained in the other single domain swap. Insertion of the *dmbS* KR into the *tenS* background led to the detection of four compounds, the major one being a dimethylated hexaketide-derived tetramic acid **6**. This was a particularly exciting result, as this compound corresponds to a precursor of bassianin **7**, which can be regarded as a hybrid of tenellin **1** and desmethylbassianin **2**. Isolation of bassianin **7** and biosynthetic studies were reported many years ago by Vining and Wright.[31] However, the original producing strain is no longer extant, and in the course of our recent work we could find no trace of bassianin production among some 30 *Beauveria* strains examined.[7] Thus, bassianin is effectively an extinct metabolite. However, coexpressing the *tenS*(ΔKR:*dmbS*-KR) with *tenA*, *tenB*, and *tenC* resulted in production of bassianin **7**, which, in keeping with its "raising from the dead", we class as a "Lazarus metabolite". This ability to not only produce new natural products but also recreate extinct compounds is a pertinent demonstration of the power of combinatorial biosynthesis as applied to fungal polyketides.

We have previously shown that fungal hr-PKSs, in which the ER component is inactive (ER^0) and where the ER function is provided by a *trans*-acting ER, appear to be fundamentally promiscuous in their programming. This is the case with TENS, DMBS, and the lovastatin nonaketide synthase (LNKS) investigated by Vederas, Tang, and co-workers.[32] However, in the presence of the cognate ER, provided by the *tenC*, *dmbC*, and *lovC* genes, respectively, the hr-polyketide synthases usually display high fidelity and produce single compounds. Thus, the issue of fidelity is closely allied with that of programming, as demonstrated in the experiments where we exchanged single catalytic domains. Two methylation patterns were observed in the case of *C*MeT-ΨKR exchange, while for the single KR domain swap all four observed compounds were produced. This contrasts with the experiment where the *C*MeT-ΨKR-ER^0-KR exchange cleanly produced **4** and no other compounds. This shows that while the *C*MeT-ΨKR and KR domains are the dominant determinants of the methylation and chain length programs, they are not the sole determinants. Presumably interactions between these domains perturb the structures of *C*MeT-ΨKR and KR such that they are incapable of fully exerting their usual selectivities in a heterologous setting.

Our results suggest a model in which the individual catalytic modifying domains themselves possess selectivity for specific substrates. In the cases of TENS and DMBS, the only differences in programming relate to chain length and methylation pattern, and these are controlled by the KR and *C*MeT-ΨKR domains, respectively. We have already shown that the DMBC and TENC ERs have selectivity for methylated diketide enoyl units, and this is conserved in the case of all hybrids examined here. For both TENS and DMBS, the DH appears to be unselective, whereas in other cases, such as that of LNKS/LOVC, the DH does show selectivity, acting early in the program but not later.[17,32] It seems reasonable to believe that the LNKS DH domain does possess selectivity, but this could not be explored in the TENS/DMBS system, as both show the same DH programming. LNKS also differs from TENS and DMBS in lacking an obvious release domain, so chain length may be controlled here by lactonization. Our model also suggests that the AT, KS, and ACP domains play no role in programming—being merely efficient engines for chain extension.

Perhaps significantly, programming differences between TENS and DMBS all map to the KR, ΨKR, and *C*MeT domains. Since the mFAS structure shows these to be physically in contact with one another (orange, yellow, and blue, respectively, in Figure 1) on the periphery of the synthase, and given the apparent conservation between them, the same may be true for TENS and DMBS. It is not yet known if or where the *trans*-ER docks to the structure, but it is likely that this may be in a similar region so that programming is controlled by a relatively small volume of the overall synthase. Docking by the *trans*-ER onto the hr-PKS would explain the observed increase in fidelity when *tenS* and *tenC* are coexpressed[7,14]— such an interaction could be envisaged to produce allosteric changes in the *C*MeT, KR, and ΨKR, which in turn could contribute to the fidelity of programming. Sequence comparison between the hr-PKS and mFAS shows that the PKSs are generally longer both within and between catalytic domains; these sequence differences may contribute to the significant differences in selectivity between the hr-PKS and mFAS, and such studies will form the focus of our future experiments.

The results of our rational domain swaps show for the first time how the TENS and DMBS fungal hr-PKSs are programmed and allow us to suggest a general mechanism by which the very large class of similar synthases are also programmed to produce the observed diversity of products. We propose a mechanism in which the KS, AT, and ACP domains perform much as the related Type II PKS (exemplified by the actinorhodin PKS)[33,34] and nr-PKS as efficient, but indiscriminate, chain initiation and extension catalysts.[5] This then leaves programmed roles to the modifying domains at the "top" of the synthase (as depicted in Figure 1A). The detailed chemical mechanisms and substrate—protein interactions of these programmed steps remain to be elucidated, and we are currently investigating this question with more focused domain swap experiments and structural work. However, we have already shown that these synthases can be

rationally manipulated to produce specific compounds, as demonstrated by the production of bassianin **7**.

■ ASSOCIATED CONTENT

Supporting Information. All chemical, genetic, and microbiological experimental details; full characterization of new compounds; and a multiple alignment of mFAS, TENS, and DMBS sequences. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
r.j.cox@bris.ac.uk; tom.simpson@bris.ac.uk

■ ACKNOWLEDGMENT

This work was supported by EPSRC grant to R.J.C. (EP/F066104/1, LCMS equipment, KMF, ZS). The authors are also grateful for funding from Kano State Government Nigeria, The MacArthur Foundation, Bayero University, and the Nigerian Petroleum Technology Fund (A.A.Y.); The Higher Education Commission of Pakistan (Z.W.); The University of Bristol (J.P.); and Beni Suef University and the Egyptian Ministry of Higher Education (W.B.).

■ REFERENCES

(1) Cox, R. J. *Org. Biomol. Chem.* **2007**, *5*, 2010–2026.
(2) Staunton, J.; Weissman, K. J. *Nat. Prod. Rep.* **2001**, *18*, 380–416.
(3) Crawford, J. M.; Dancy, B. C. R.; Hill, E. A.; Udwary, D. W.; Townsend, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16728–16733.
(4) 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.
(5) Crawford, J. M.; Thomas, P. M.; Scheerer, J. R.; Vagstad, A. L.; Kelleher, N. L.; Townsend, C. A. *Science* **2008**, *320*, 243–246.
(6) McInnes, A.; Smith, D.; Walter, J.; Vining, L.; Wright, J. *J. Chem. Soc., Chem. Commun.* **1974**, 282–284.
(7) Heneghan, M. N.; Yakasai, A. A.; Williams, K.; Kadir, K. A.; Wasil, Z.; Bakeer, W.; Fisch, K. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J.; Lazarus, C. M. *Chem. Sci.* **2011**, *2*, 972.
(8) Jessen, H. J.; Schumacher, A.; Shaw, T.; Pfaltz, A.; Gademann, K. *Angew. Chem., Int. Ed.* **2011**, *50*, 4222–4226.
(9) Eley, K. L.; Halo, L. M.; Song, Z.; Powles, H.; Cox, R. J.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J. *ChemBioChem* **2007**, *8*, 289–297.
(10) Sims, J. W.; Schmidt, E. W. *J. Am. Chem. Soc.* **2008**, *130*, 11149–11155.
(11) Liu, X.; Walsh, C. T. *Biochemistry* **2009**, *48*, 8746–8757.
(12) Du, L.; Lou, L. *Nat. Prod. Rep.* **2010**, *27*, 255.
(13) Moore, M.; Cox, R. J.; Duffin, G.; O'Hagan, D. *Tetrahedron* **1998**, *54*, 9195–9206.
(14) Halo, L. M.; Marshall, J. W.; Yakasai, A. A.; Song, Z.; Butts, C. P.; Crump, M. P.; Heneghan, M.; Bailey, A. M.; Simpson, T. J.; Lazarus, C. M.; Cox, R. J. *ChemBioChem* **2008**, *9*, 585–594.
(15) Halo, L. M.; Heneghan, M. N.; Yakasai, A. A.; Song, Z.; Williams, K.; Bailey, A. M.; Cox, R. J.; Lazarus, C. M.; Simpson, T. J. *J. Am. Chem. Soc.* **2008**, *130*, 17988–17996.
(16) Maier, T.; Leibundgut, M.; Ban, N. *Science* **2008**, *321*, 1315.
(17) Kennedy, J.; Auclair, K.; Kendrew, S.; Park, C.; Vederas, J.; Hutchinson, C. *Science* **1999**, *284*, 1368–1372.
(18) Oliynyk, M.; Brown, M. J.; Cortes, J.; Staunton, J.; Leadlay, P. F. *Chem. Biol.* **1996**, *3*, 833–839.
(19) Heneghan, M. N.; Yakasai, A. A.; Halo, L. M.; Song, Z.; Bailey, A. M.; Simpson, T. J.; Cox, R. J.; Lazarus, C. M. *ChemBioChem* **2010**, *11*, 1508–1512.
(20) Keatinge-Clay, A. T.; Maltby, D. A.; Medzihradszky, K. F.; Khosla, C.; Stroud, R. M. *Nat. Struct. Mol. Biol.* **2004**, *11*, 888–893.
(21) Morita, H.; Kondo, S.; Kato, R.; Wanibuchi, K.; Noguchi, H.; Sugio, S.; Abe, I.; Kohno, T. *Acta Crystallogr. F* **2007**, *63*, 947–949.
(22) Zhu, X.; Yu, F.; Bojja, R. S.; Zaleta-Rivera, K.; Du, L. *J. Ind. Microbiol. Biotechnol.* **2006**, *33*, 859–868.
(23) In a similar experiment, donation of the KS domain from LNKS into FUM1 appeared to result in the manufacture of distantly related dihydroisocoumarins, although these compounds could not be proven to be produced as a direct result of the domain swap in this case.Zhu, X.; Yu, F.; Li, X.-C.; Du, L. *J. Am. Chem. Soc.* **2007**, *129*, 36–37.
(24) Chakravarty, B.; Gu, Z.; Chirala, S. S.; Wakil, S. J.; Quiocho, F. A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15567–15572.
(25) Fisch, K. M.; Skellam, E.; Ivison, D.; Cox, R. J.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J. *Chem. Commun.* **2010**, *46*, 5331–5333.
(26) Bailey, A. M.; Cox, R. J.; Harley, K.; Lazarus, C. M.; Simpson, T. J.; Skellam, E. *Chem. Commun.* **2007**, 4053–4055.
(27) Skellam, E. J.; Hurley, D.; Davison, J.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. *Mol. BioSyst.* **2010**, *6*, 680–682.
(28) Zha, W.; Shao, Z.; Frost, J. W.; Zhao, H. *J. Am. Chem. Soc.* **2004**, *126*, 4534–4535.
(29) Spencer, J.; Jordan, P. *Biochem. J.* **1992**, *288*, 839–846.
(30) Xu, W.; Cai, X.; Jung, M. E.; Tang, Y. *J. Am. Chem. Soc.* **2010**, *132*, 13604–13607.
(31) Wat, C.-K.; McInnesS, A.; Smith, D.; Wright, J. L. C.; Vining, L. *Can. J. Chem.* **1977**, *55*, 4090–4098.
(32) Ma, S. M.; Li, J. W.-H.; Choi, J. W.; Zhou, H.; Lee, K. K. M.; Moorthie, V. A.; Xie, X.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C.; Tang, Y. *Science* **2009**, *326*, 589–592.
(33) Beltran-Alvarez, P.; Cox, R. J.; Crosby, J.; Simpson, T. J. *Biochemistry* **2007**, *46*, 14672–14681.
(34) Keatinge-Clay, A. T.; Maltby, D. A.; Medzihradszky, K. F.; Khosla, C.; Stroud, R. M. *Nat. Struct. Mol. Biol.* **2004**, *11*, 888–893.