

ARTICLE

<https://doi.org/10.1038/s41467-019-11896-1>

OPEN

Emulating evolutionary processes to morph aureothin-type modular polyketide synthases and associated oxygenases

Huiyun Peng¹, Keishi Ishida¹, Yuki Sugimoto¹, Holger Jenke-Kodama² & Christian Hertweck^{1,3} 

Polyketides produced by modular type I polyketide synthases (PKSs) play eminent roles in the development of medicines. Yet, the production of structural analogs by genetic engineering poses a major challenge. We report an evolution-guided morphing of modular PKSs inspired by recombination processes that lead to structural diversity in nature. By deletion and insertion of PKS modules we interconvert the assembly lines for related antibiotic and antifungal agents, aureothin (*aur*) and neo-aureothin (*nor*) (aka spectinabilin), in both directions. Mutational and functional analyses of the polyketide-tailoring cytochrome P450 monooxygenases, and PKS phylogenies give contradictory clues on potential evolutionary scenarios (generalist-to-specialist enzyme evolution vs. most parsimonious ancestor). The KS-AT linker proves to be well suited as fusion site for both excision and insertion of modules, which supports a model for alternative module boundaries in some PKS systems. This study teaches important lessons on the evolution of PKSs, which may guide future engineering approaches.

¹Department of Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology (HKI), Beutenbergstrasse 11a, 07745 Jena, Germany. ²Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan. ³Faculty of Biological Sciences, Friedrich Schiller University Jena, 07743 Jena, Germany. Correspondence and requests for materials should be addressed to C.H. (email: christian.hertweck@leibniz-hki.de)

The wealth of complex polyketides produced by bacteria is an essential source for antifungal, antitumor, antiparasitic, and immunosuppressive agents that are currently used in the clinics¹. Whereas their chemical synthesis is often challenging, bacteria produce these valuable compounds from simple carboxylic acid building blocks by means of modular (type I) polyketide synthases (PKSs)². Each module of these highly versatile assembly lines consists of a minimal set of domains for chain elongation: a ketosynthase (KS) catalyzing Claisen condensations of the activated acyl and malonyl building blocks, an acyltransferase (AT) for selecting and loading extender units, and an acyl carrier protein (ACP) domain that serves as an anchor for the growing chain³. Additional β -keto processing domains for ketoreduction (KR), dehydration (DH), and enoyl reduction (ER) increase the structural diversity of the polyketide chain. The architecture of the megasynthase is typically arranged in a way that each module catalyzes a single chain elongation, ensuring that polyketide assembly progresses in a unidirectional manner until the full-length product is released by a thioesterase (TE) domain³. Consequently, in most bacterial type I PKSs the PKS assembly line directly corresponds to the chemical structure of the resulting polyketide.

The paradigm of successive assembly and co-linearity between assembly steps and final product sets the basis for the *in silico* prediction of polyketide structures from type I PKS genes, thus enabling the discovery of natural products by genome mining⁴. Likewise, the assembly line logic inspires the rational reprogramming of PKSs². In many cases, such engineering attempts lead to the production of natural product derivatives^{5,6} and provide insight into the mechanisms of polyketide assembly⁷. However, routine PKS engineering is challenging not only because of constraints in DNA manipulation and recombinering⁸ but also because of the intricacies and complex dynamics of the megasynthases. For a holistic view on the molecular machineries insight into catalytic dynamics, spatial constraints, potential incompatibilities, and substrate specificities of the catalytic domains is required. It is not surprising that many combinatorial biosynthesis approaches in a mix-and-match fashion are unsatisfactory, as domains were swapped or recombined neglecting the structural impact, including non-functional regions, on protein–protein interactions. Consequently, domains and/or modules of an engineered PKS are impaired or simply non-functional⁹, which results in dramatically reduced yields or complete loss of production¹⁰.

Regardless of current constraints in PKS engineering, nature provides us with successful strategies to diversify polyketide structures through evolutionary processes¹¹. Extensive bioinformatics analyses suggest that nature harnesses point mutations, gene duplication, gene loss, homologous recombination, and horizontal gene transfer in the evolution of type I PKSs^{12–15}. However, there is little experimental evidence for such mechanisms being at work in the evolution of modular PKSs. Emulating evolutionary processes could address some of the key questions in PKS research: What are the prerequisites for the morphing of one pathway into another? Which impacts have polyketide-tailoring enzymes that coevolved with the scaffold-assembly lines? Which sites are suited for recombinations and fusions and allow the reprogramming of polyketide assembly lines?

We reason that similar polyketide structures and the requisite homologous biosynthesis gene clusters would be a good starting point to study the evolutionary processes in modular PKSs and answer such general questions. Owing to their compact, yet densely functionalized polyketide products, the homologous biosynthetic pathways for the antibacterial and antifungal compounds aureothin (**1**)¹⁶ and neo-aureothin (**2**)¹⁷ are used as model in a proof-of-concept study. Both the *aur*

PKS and *nor* PKS (Fig. 1) employ CoA-activated *p*-nitrobenzoic acid (PNBA) as starter unit¹⁸, and in both pathways homologous PKS modules generate polyene-pyrone backbones^{19,20}, which are then subjected to further enzymatic tailoring by *O*-methyltransferases (*AurI/NorI*) and cytochrome P450 monooxygenases (*AurH/NorH*)²¹. A remarkable feature of the *aur* and *nor* PKSs is that both assembly lines contain iteratively acting domains and modules^{22–24}. The only difference between the two pathways is the size of the products' polyene backbones; compared to **1**, compound **2** is composed of two additional propionate units, which are introduced by means of two additional modules in the *nor* PKS (Fig. 1)²⁵. The close relationship of the two systems suggests that one has emerged from the other. Yet, the evolutionary scenarios involved have remained a riddle.

Here we report the successful, bidirectional morphing of modular PKSs from two distinct biosynthetic pathways, employing evolutionary models deduced from *in silico* analyses. Surprisingly, tailoring enzymes give important clues about the direction of evolution. We also show that the KS-AT linker is a suitable engineering site, which points to alternative module boundaries in these assembly lines and may thus facilitate future engineering approaches in related PKS systems.

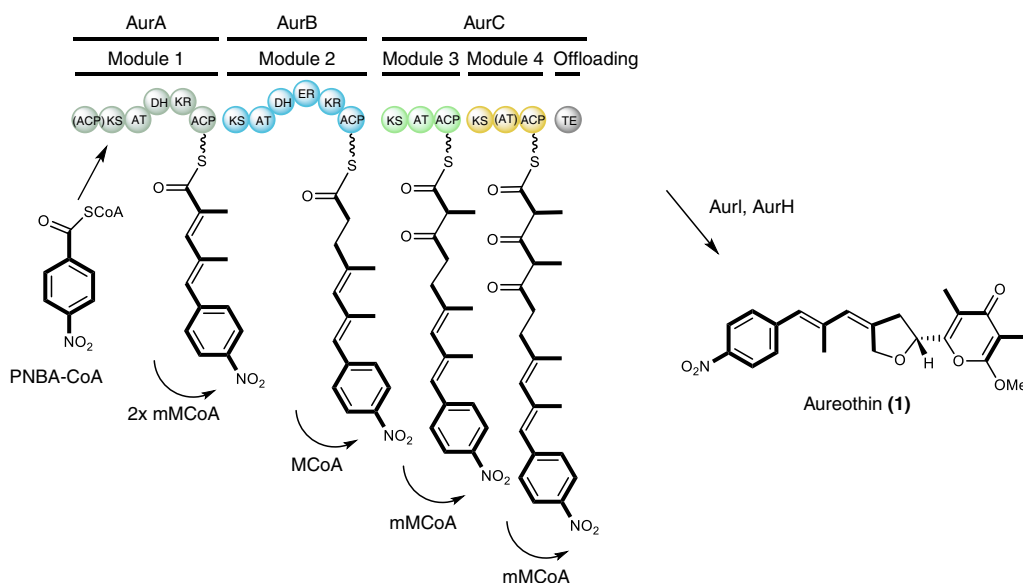
Results

Genome mining and phylogenetic analyses of gene clusters. To gain insight into potential gene evolution scenarios we searched for gene loci coding for assembly lines related to the *aur* and *nor* PKSs. Genome mining using BLAST (Basic Local Alignment Search Tool) in the NCBI (National Center for Biotechnology Information) database identified several gene loci with high homology to the *aur* and *nor* gene clusters (Fig. 2a). These gene clusters share the genes for PNBA starter unit biosynthesis, polyketide chain elongation, post-PKS modification and regulation (Supplementary Table 1).

Although eight of the gene clusters can be grouped either into *aur*- or *nor*-type, one orphan gene cluster from *Streptomyces scabrissporus* (DSM41855) deviates from the others as it lacks a *norB/aurB* homolog. The absence of this gene and the correct assembly of the contigs were confirmed by four pairs of PCR primers (Supplementary Fig. 1). To elucidate the product of the encoded cryptic assembly line, we cultured *S. scabrissporus* and monitored its metabolic profile. Unexpectedly, we found that this strain produces neo-aureothin (Fig. 2b and Supplementary Fig. 11) despite the absence of a *norB* homolog in the identified gene cluster (Fig. 2a). It is conceivable that a *NorB* homolog is encoded elsewhere in the genome. Indeed, we identified a candidate for a freestanding *NorB* gene (WP_078978330.1) in the yet incomplete genome sequence of *S. scabrissporus*. This finding of a split *nor* gene cluster is intriguing as it shows that gene rearrangements take place in *aur/nor*-type gene clusters. Such rearrangements could in fact drive the evolution of metabolic diversity in the aureothin family.

With the enlarged set of gene clusters at hand, we aimed at gaining insight into their phylogenetic relationship, which could give clues about their evolution. Therefore, the amino acid sequences of the KS domains from those homologous gene clusters were aligned with the KS domain sequences from other actinobacterial PKS gene clusters by the GUIDANCE2 Server²⁶. The aligned sequences were subjected to phylogenetic analyses, and the evolutionary tree was constructed by Bayesian inference with the MrBayes software²⁷ (Fig. 2c, Supplementary Fig. 2, Supplementary Table 2). For simplification, some KS domains from *aur*- and *nor*-type gene clusters were excluded as they showed the exact same sequences.

aur PKS



nor PKS

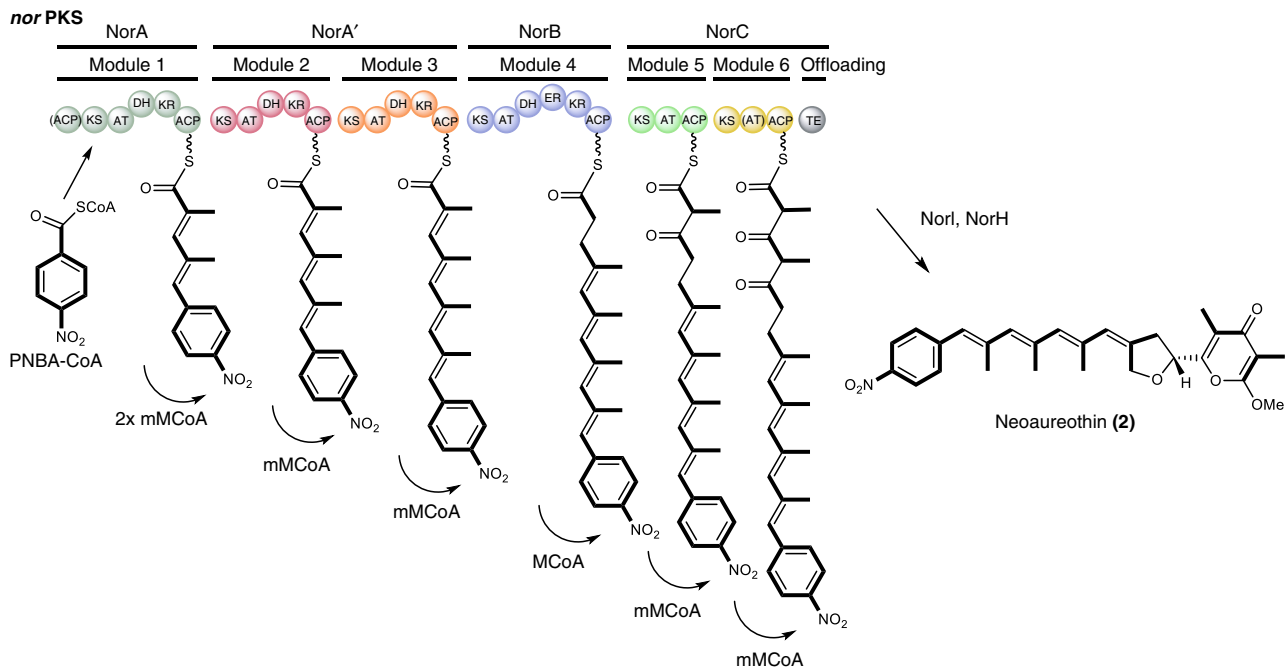


Fig. 1 Non-colinear polyketide assembly lines for aureothin and neoareothin biosynthesis. AurA and NorA catalyze two chain elongations. PNBA: *p*-nitrobenzoic acid, mM-CoA: methylmalonyl-CoA, M-CoA: malonyl-CoA, KS: ketosynthase, AT: acyltransferase, DH: dehydratase, ER: enoylreductase, KR: ketoreductase, ACP: acyl carrier protein, TE: thioesterase. The inactive domains in (AurA/NorA) and (AurC/NorC) are indicated with brackets. Color code in module indicates homologous proteins (except for AurB (cyan) and NorB (purple), for which different colors are used to distinguish the engineered site in Figs. 3d and 5.) These color codes correspond with those shown in Fig. 2c

Each module of the *aur* and *nor* clusters is monophyletic, i.e. the sequences form their own branches without intermingling of sequences from other modules. The only apparent exception is the NorC KS-1 sequence of *S. scabrisporus*, which clusters together with NorB KS sequences. An alignment of the *S. scabrisporus* NorC KS-1 with selected AurB KS/NorB KS and AurC KS-1/NorC KS-1 sequences (Supplementary Fig. 3) revealed its hybrid nature. When comparing positions that have a characteristic amino acid residue or indel pattern in the NorB KS or NorC KS-1 group, it becomes clear that the *S. scabrisporus* sequence resembles more the module B type in the amino-terminal region whereas the carboxy-terminal stretch shows higher similarity to the module C KS-1 type. This is probably due

to a recombination event. Since the similarity with the module B type prevails, the sequence is located in the module B branch of the tree.

AurA/NorA as well as NorA'-2 sequences form separate branches. The other modules, however, originated from a common ancestor. It is important to note that the NorA'-1 cluster, which comprises sequences exclusive to the *nor* cluster, forms an early branch within that big monophyletic group. Therefore, from a phylogenetic perspective, it is reasonable to propose that *aur*-type PKSs emerged from *nor*-type PKSs, possibly through gene deletion (Fig. 2d). This result is in line with the previous analysis, which suggested that a *nor*-to-*aur* PKS evolution would be the most parsimonious scenario²⁰.

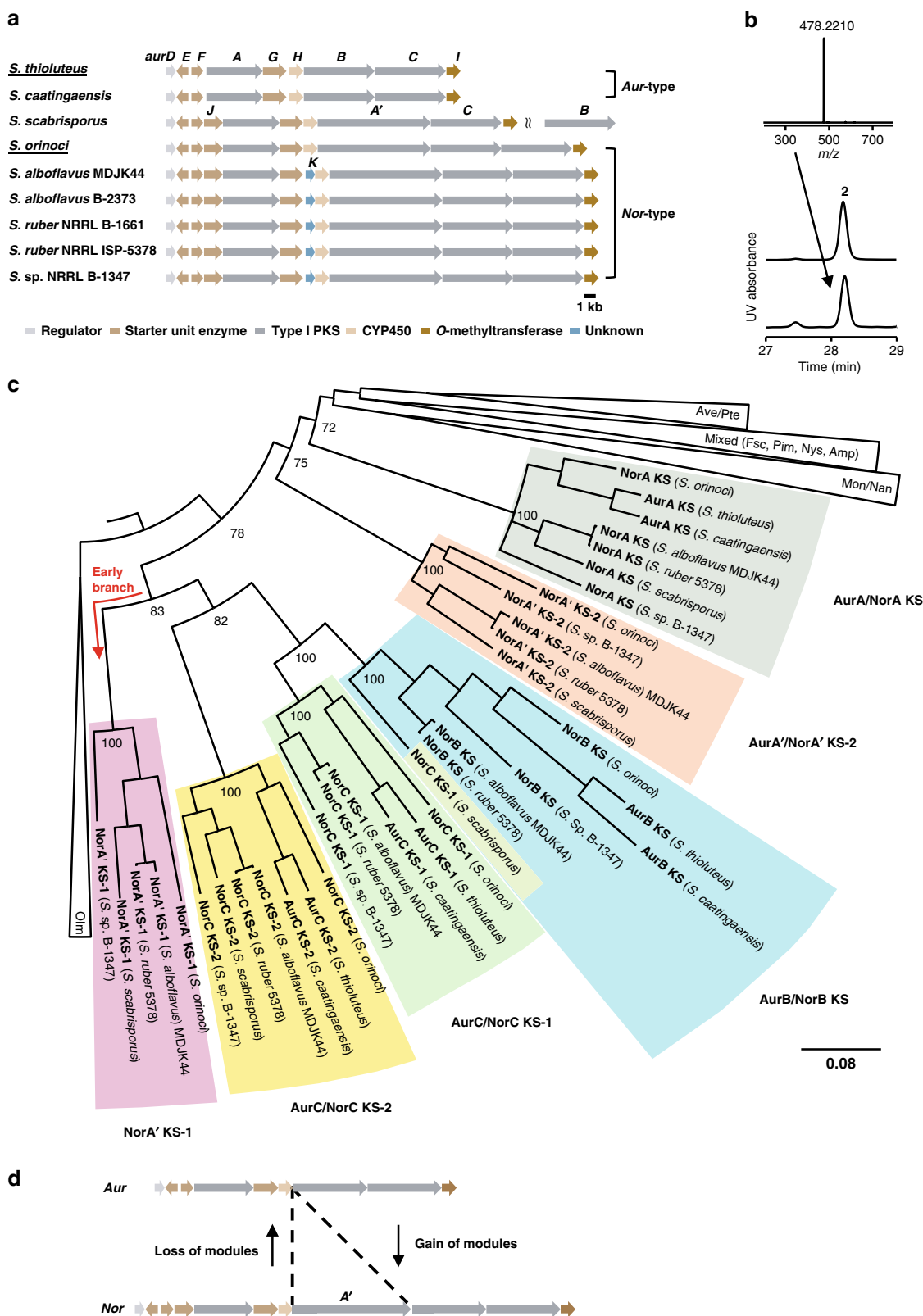


Fig. 2 Phylogenetic analysis of the *aur* PKS and *nor* PKS. **a** The *aur*-type and *nor*-type gene clusters. The original producers of aureothin and neoaureothin are underlined. **b** The HPLC profile of authentic reference of neoaureothin (**2**, middle panel), and *S. scabrisporus* culture (bottom panel). HR-MS spectrum of neoaureothin in top panel. UV detection is at 350 nm. **c** Phylogenetic tree of KS domain amino acid sequences of the *aur*-type and *nor*-type PKSs and other actinobacterial PKSs. The tree was reconstructed by Bayesian inference. Numbers at nodes indicate clade credibility values. Branch length represents the number of amino acid changes per position. Olm, oligomycin; Tyl, tylacton; Nan, nanchangmycin; Mon, monensin; Fsc, candidin; Pim, pimaricin; Nys, nystatin; Amp, amphotericin; Ave, avermectin; Pte, polyene macrolide from *S. avermitilis*. Color code for KS clades corresponds with those shown in Fig. 1 **d**. Possible evolutionary scenarios between *aur* PKS and *nor* PKS. The source data underlying Fig. 2 **c** are provided as a Source Data file

Morphing the *nor* PKS into an *aur* PKS. For functional analyses and PKS engineering approaches, we needed to establish a robust expression system. Initially, the heterologous expression of the *nor* gene cluster was achieved in *S. lividans* by coexpression of the transcriptional regulator AurD from the *aur* gene cluster. The titer of neo-aureothin was, however, unsatisfactory (15 mg L^{-1})²⁸. This low yield might result from the non-concerted expression of the *nor* biosynthesis genes using a three-plasmid system. To increase neo-aureothin production we optimized the heterologous expression system to reassemble the *nor* gene cluster in a continuous gene region (Fig. 3a). First, a part of the *nor* gene cluster (pNT42) was integrated into the genome of the heterologous expression host *S. albus* by site-specific recombination. The complete *nor* gene cluster was then obtained by a homologous recombination using a suicide vector (pYU93) harboring the left part of the gene cluster. The resulting strain (*S. albus*::pNT42/pYU93; *S. albus_nor* PKS) produced threefold higher titers (45 mg L^{-1}) of neo-aureothin compared to the previous construct (Fig. 3b). Thus, *S. albus_nor* PKS was used as a platform for PKS engineering.

To emulate the presumed evolutionary processes involved in the *nor*-to-*aur* PKS transformation (Fig. 2d) we attempted to morph the *nor* PKS into an assembly line producing aureothin. Since the polyketide backbone of aureothin lacks two methylmalonyl-derived C2 units, two requisite modules in the *nor* PKS needed to be deleted. As (AurA/NorA), (AurB/NorB), and (AurC/NorC) share high identities on the DNA and amino acid levels, we excised the gene regions for modules 2 and 3 in the *nor* gene cluster. Thereby, it was essential to consider the compatibility of the docking domains between the individual PKS proteins^{29–33}, because the sequence elements at the extreme C- and N-termini of PKS subunits help mediate their interactions. The amino acid sequence alignment among PKS proteins indicated that the ^CDD (C-terminal docking domain) of AurA/NorA and the ^NDD (N-terminal docking domain) of AurB/NorA' is a class 1a docking domain system (Supplementary Fig. 4)^{32,33}. On the other hand, the ^CDD of NorA' and the ^NDD of NorB are a class 1b system (Supplementary Fig. 4)^{32,33}. In a control experiment neglecting ^CDD/^NDD compatibility we generated a mutant (*S. albus_aur^S* PKS) lacking *norA'* (Fig. 3c). Surprisingly, this strain produced intermediate (5) and trace amounts of 7-hydroxydeoxyaureothin (3), but not aureothin (1) (Fig. 3e and Supplementary Fig. 12). Apparently, the different types of docking domains between NorA and NorB can communicate, albeit only weakly. To achieve higher compatibility between ^CDD of NorA and ^NDD of NorB, we constructed two recombinant PKS variants with different fusion sites. Initially, we have employed *S. lividans* as heterologous host²⁸, but to increase neo-aureothin production we reconstructed an *S. albus_nor* PKS expression system. Whereas the first system had a fusion site at the docking domain region (*aur[#]* PKS), the second one has a fusion site at the KS-AT linker region (*aur^{*}* PKS) (Fig. 3c, d and Supplementary Fig. 5). In both cases, we have swapped the ^NDD of NorB for that of NorA', the natural partner of the ^CDD of NorA. In modular PKS, two hot spots for evolutionary recombination events have been suggested, KS-AT linker and post AT linker^{34,35}. We initially chose one fusion site upstream of the conserved KS-AT linker, as we have already succeeded in engineering aureothin congeners using this site for recombinations (Supplementary Fig. 5)^{28,36}. The constructs were introduced into *S. albus* via triparental conjugation to generate *S. albus_aur[#]* PKS and *S. albus_aur^{*}* PKS.

The verified recombinant strains were fermented, and the ethyl acetate extracts of the cultures were monitored by HPLC. In both recombinant strains the production of intermediate (5, Supplementary Figs. 19–23, Supplementary Table 8) and a trace amount

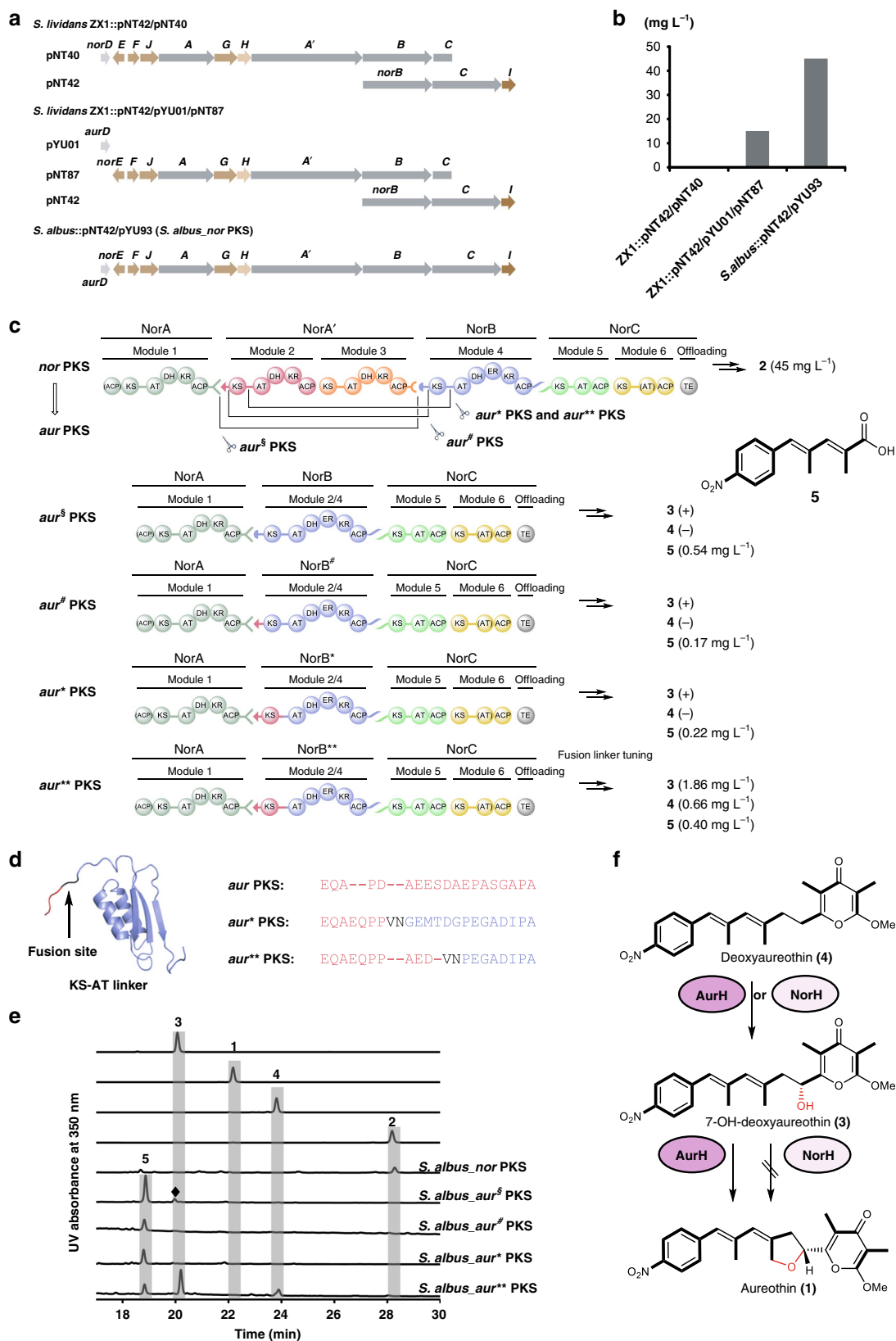
of 7-hydroxydeoxyaureothin (3) but no aureothin (1) could be detected (Fig. 3e, and Supplementary Fig. 12). All recombinant strains *aur^S_PKS*, *aur[#]_PKS*, and *aur^{*}_PKS* produce elevated amounts of intermediate (5) compared to AurA alone²⁴, indicating that the docking domains promote protein interactions. However, the low production of the full-length polyketide pointed to incorrect protein folding, likely because of suboptimal fusion sites. Thus, we revisited the fusion site in the *aur^{*}* PKS and found that the recombinant KS-AT linker in the *aur^{*}* PKS bore four amino acids more than that in the genuine *aur* PKS (Fig. 3d). Since this difference might have an impact on protein configuration and interactions in the *aur^{*}* PKS, we shortened the recombinant KS-AT linker region by λ Red-mediated recombination, yielding, in part serendipitously, the *aur^{**}* PKS gene cluster (Supplementary Figs. 5 and 6).

In the extract of the culture broth of *S. albus_aur^{**}* PKS, the target molecule (1) could still not be detected by HPLC. Instead, we noted the formation of other metabolites (Fig. 3e and Supplementary Fig. 12). Through HPLC-HRMS analysis and by comparison with an authentic reference, the compounds were identified as 7-OH-deoxyaureothin (3) and 7-deoxyaureothin (4), which differs only from 1 in that they do not form the tetrahydrofuran ring (Supplementary Fig. 12). Notably, the polyketide backbones of 1–3 are identical, which indicated that the *nor* PKS has been successfully morphed into an aureothin assembly line. Yet, the enzymatic tailoring of the polyketide scaffold proved to be erratic.

From in vivo and in vitro studies we know that the formation of the tetrahydrofuran ring is the last step in aureothin biosynthesis and that its installation involves two sequential C–O-bond formations catalyzed by a single cytochrome P450 monooxygenase, AurH^{37,38}. Furthermore, the AurH-mediated oxygenation processes are highly fine-tuned, and changes in the enzyme or in the size of the substrate result in incomplete transformations or alternative reaction channels^{39,40}. By analogy, the short aureothin backbone does not appear to be the preferred substrate of the homologous oxygenase (NorH) from the *nor* pathway. NorH is only able to convert deoxyaureothin (4) into the hydroxylated congener 3, whereas the second oxidation and thus also heterocyclization do not take place (Fig. 3f). In case of the tentative *nor*-to-*aur* PKS evolution, not only the PKS needed to morph, but also the tailoring enzyme (AurH) needed to adjust.

Mutagenesis and cross-complementation. To gain insight into possible changes in the CYP450 we compared NorH and AurH. Both enzymes share similar amino acid sequences (64% identity, 75% positives), and recognize similar substrates. Thus, the overall structure of NorH is likely similar to AurH³⁹. Threading of the NorH amino acid sequence onto the AurH crystal structure revealed a conserved hydrophobic pocket for binding the pyrone ring of deoxyaureothin. However, the modeling indicated that NorH has a wider cavity around the active center, as some amino acid residues possess smaller or more flexible side chains than those found in AurH (Fig. 4a). Based on this information, a number of point mutations (I19F, V71L, T291L, T292P, W317F, and T392L) could, in principle, reconfigure the active site of NorH to recognize deoxyaureothin³⁹.

To test this hypothesis, we altered the active site of NorH by site-directed mutagenesis. Thus, we constructed a range of *norH* variants, including the I19F, V71L, T291L, T292P, W317F, and T392L mutants, and cloned them individually into expression vectors for complementation of an *aurH* knock-out mutant (Δ *aurH*)³⁷ (Fig. 4b and Supplementary Table 3). Expression vectors containing wild-type *norH* and *aurH* served as negative and positive controls, respectively. All plasmids were introduced



into the $\Delta aurH$ mutant by triparental conjugation, and the metabolic profiles of the individual transformants were monitored by HPLC-MS (Fig. 4b and Supplementary Fig. 13).

The $\Delta aurH$ mutant produces exclusively 4; when complementing the mutant with native AurH, 4 is readily converted into 1 (positive control). In contrast, the mutant complemented with

NorH partly transformed 4 into 3, and it was not capable of forming 1 (negative control). All $\Delta aurH$ mutant strains complemented with point-mutated NorH variants showed the same chemotype. The only difference was a slightly increased 3-to-4 ratio for NorH-V71L and NorH-T292P (Fig. 4b and Supplementary Fig. 13). These results indicate that these

Fig. 3 Overview of the modified *nor* PKS variants. **a** The optimization of the *nor* gene cluster heterologous expression system. The repressor gene *norD* is replaced by the activator gene *aurD*, and the intact *nor* gene cluster is assembled. **b** The titers of neo-aureothin in different mutants are presented in the bar chart. **c** Domain architectures of the *aur^S* PKS, *aur[#]* PKS, *aur^{*}* PKS, and *aur^{**}* PKS. Different docking domain pairs are in different shapes. The titers of polyketides in different mutants are given. Plus (+) indicates minute production only detected by LC-MS. Minus (–) indicates no production. **d** The fusion site VN is within the KS-AT linker and induced by the restriction enzyme *HpaI*. The amino acid sequences of the KS-AT linker region among the *aur* PKS, *aur^{*}* PKS and *aur^{**}* PKS are aligned. **e** HPLC profile of authentic reference of 7-OH-deoxyaureothin (**3**), aureothin (**1**), deoxyaureothin (**4**), neo-aureothin (**2**), *S. albus* mutant strains. The peak (marked with a diamond) with the similar retention time of **3** corresponds to a pigment produced by the host (*S. albus*). UV detection is at 350 nm. **f** NorH converts **4** to **3**. It cannot catalyze the tetrahydrofuran ring formation. The source data underlying Fig. 3b are provided as a Source Data file

individual point mutations of NorH are not sufficient to reconfigure its active site to generate the THF ring of aureothin.

Therefore, we generated an expression plasmid for a NorH variant containing all six point mutations (I19F-V71L-T291L-T292P-W317F-T392L). Yet, in the metabolic profile of the Δ *aurH* strain complemented with the multiple point-mutated NorH variant, **1** could not be detected, either (Fig. 4b and Supplementary Fig. 13). Thus, we scrutinized the highly similar P450 monooxygenases NorH and AurH (Supplementary Fig. 7 and Supplementary Table 4) and attempted to target the protein domains that are relevant for THF-ring formation by constructing chimeric NorH variants. AurH variants adopt different conformations mainly at the B2 and B2' two-helix-bundle, FG-loop and β 2-loop that surround the active center and approach the center after binding to substrate³⁹. These residues likely generate steric pressure to bend the intermediate and push it towards the reaction center, facilitating THF-ring formation. In order to test whether these residues are important for THF-ring formation, five gene regions around these residues from AurH were amplified and used to replace each corresponding region in the NorH gene (Supplementary Fig. 8, Supplementary Table 5). We noted, however, that these chimeras also produce exclusively 7-OH deoxyaureothin (**3**) (Supplementary Figs. 14 and 15). We also created AurH/NorH hybrids differing at the N-terminal end of the α helix (Supplementary Fig. 8, Supplementary Table 6). These head/tail exchange hybrids showed reduced catalytic activity (Fig. 4c and Supplementary Fig. 16). Thus, the exchanged region was narrowed down to avoid possible deleterious effects on the overall structure. The fusion sites were placed within the C helix and the K helix (Supplementary Fig. 8). Thus, NorH was divided into three areas, part A, B, and C. Correspondingly, AurH was dissected into parts a, b, and c (Fig. 4c). The HPLC profiles of the obtained hybrids, NorH/AurH ABC, AbC, aBC, Abc, aBc, and abC variants, indicated that only hybrid Abc variant could transform 7-deoxyaureothin (**4**) to aureothin (**1**), albeit only incompletely (Fig. 4c, Supplementary Fig. 16).

Taken together, structure-based, rational mutations and domain swapping of NorH are not sufficient to reconfigure the active site of NorH to function like AurH. The 70% C-terminal AurH hybrid NorH (Abc) shows only substantially reduced activity. These results indicate that complex evolutionary processes would be required to maintain THF ring formation activity when mutating NorH to AurH. Consequently, we also considered the reverse scenario and aimed at emulating a potential *aur*-to-*nor* PKS evolution.

Therefore, we interrogated the substrate specificity of AurH with respect to the enzyme's ability to transform deoxyneo-aureothin (**6**) into **2**. To this end, *norH* was deleted in *S. albus_{nor}* PKS using the λ Red system. The resulting Δ *norH* mutant was fermented, and the culture extract was analyzed by HPLC. The HPLC profile showed that Δ *norH* lost the ability to produce **2**. Instead, a different metabolite was detected (Fig. 4d and Supplementary Fig. 17). The structure of this metabolite was determined as **6** by ¹H and ¹³C

NMR, ¹H-¹H COSY, HSQC and HMBC (Supplementary Figs. 17 and 24–29, Supplementary Table 9).

The NorH and AurH expression plasmids were introduced into the Δ *norH* mutant by triparental conjugation. As expected, NorH restored the production of **2** (Fig. 4d and Supplementary Fig. 17). Surprisingly, complementing the Δ *norH* mutant with AurH also restored the production of **2** (Fig. 4d and Supplementary Fig. 17). This result demonstrates that AurH is relatively flexible in substrate specificity and highly efficient in converting **6** to **2** (Fig. 4e and Supplementary Table 7). The broader substrate specificity of AurH indicates that an evolution from AurH to NorH is a more probable scenario according to the generalist-to-specialist model in enzyme evolution, where ancestral enzymes show higher promiscuity and the more specialized enzymes are evolved to catalyze specific reactions^{41,42}.

Morphing the *aur* PKS into a *nor* PKS. Since AurH catalyzes the transformation of **6** into **2**, we investigated the possibility of an *aur*-to-*nor* gene cluster evolution. Therefore, we aimed at integrating modules 2 and 3 of the *nor* PKS between modules 1 and 2 of the *aur* PKS. To achieve this goal, two chimeras with different recombination sites were constructed: one fusion site is in the docking domain region (*nor[#]* PKS), and the other one in the KS-AT linker region (*nor^{*}* PKS) (Fig. 5a).

For the first construct (*nor[#]* PKS), we generated a fusion site at the N-terminal docking domain region of *aurB* and PCR-amplified intact *norA'* including its N- and C-terminal docking domains (*norA'[#]*) (Supplementary Fig. 9). The alignment of amino acid sequences of the docking domains showed that the interaction between AurA/AurB and NorA/NorA' are of class 1a, and the interaction between NorA'/NorB is of class 1b (Supplementary Fig. 4)^{29,30}. To facilitate the interaction between NorA' and AurB, the N-terminal docking domain of AurB was swapped with that of NorB to generate the *nor[#]* PKS variant. For the second construct (*nor^{*}* PKS), we chose a fusion site within the KS-AT linker region of AurB, and amplified the gene fragments for the region between the NorA'-AT2 domain and the NorB-KS4 domain (NorA'*-NorB*) for recombination (Supplementary Figs. 5 and 10). To maintain the overall conformation of proteins in the *nor^{*}* PKS, the length of the recombinant KS-AT linker was adjusted to match the size of the genuine *nor* PKS (Fig. 5b and Supplementary Fig. 5).

The verified constructs were introduced into *S. albus* to generate *S. albus_{nor[#]}* PKS and *S. albus_{nor^{*}}* PKS. In the metabolic profile of *S. albus_{nor[#]}* PKS compound **2** could not be detected. Yet, PNBA and **1** were produced, indicating that gene expression and polyketide production were functional (Fig. 5c). In the case of *S. albus_{nor^{*}}* PKS, we detected **1** and **2** (Fig. 5c and Supplementary Fig. 18), which indicated that we successfully modified the *aur* PKS to produce the homologous polyketide **2**. Thus, we showed that it is possible to modify the *aur* gene cluster to the *nor* gene cluster through PKS engineering in a manner that emulates natural evolutionary processes. The strain expressing

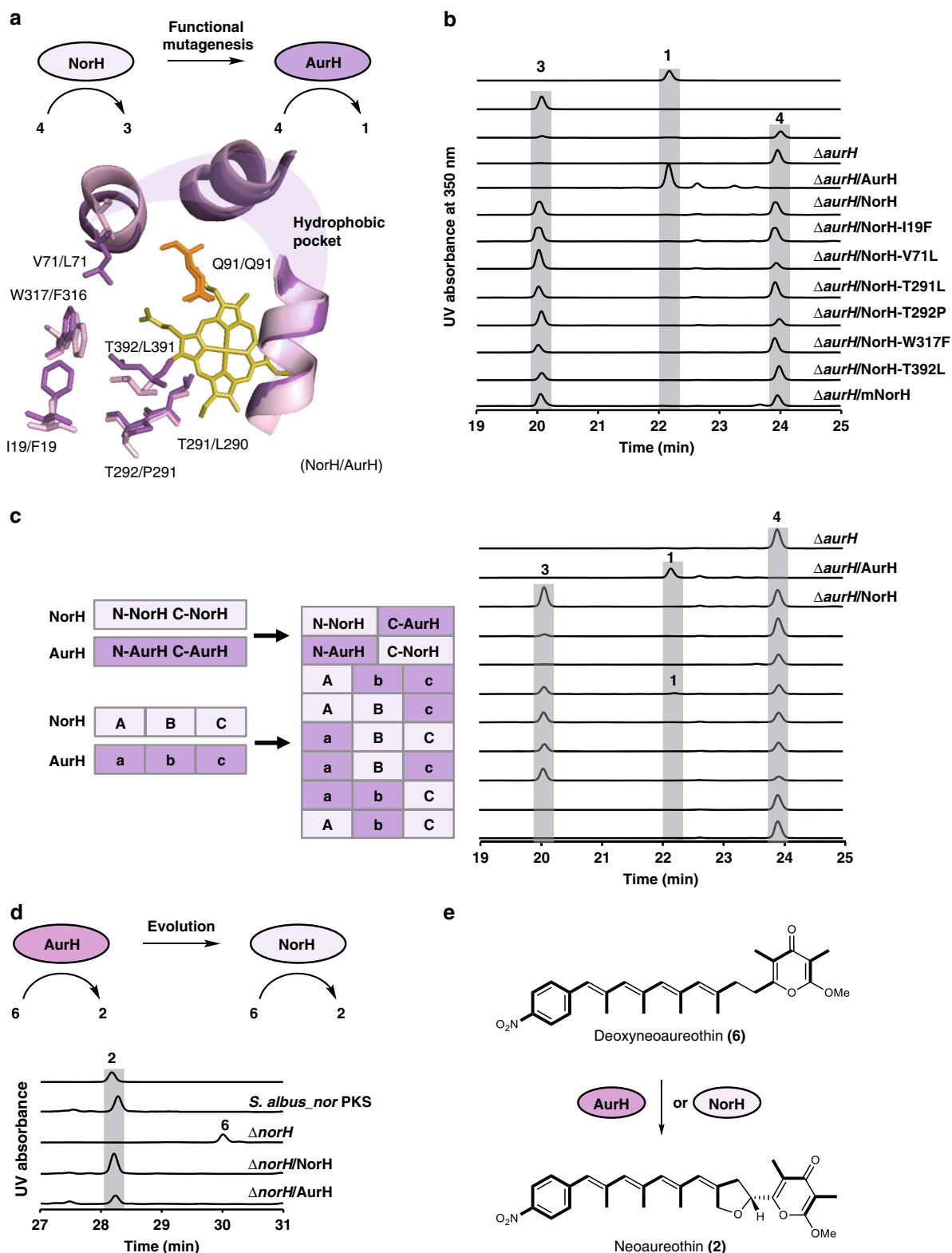


Fig. 4 Mutagenesis and comparison of oxygenase functions. **a** Site-directed mutagenesis of NorH. The active center of NorH/AurH model shows a hydrophobic pocket. Residues that may contribute to the different sizes of the cavity are presented in stick. The catalytic residue glutamine (orange) and the cofactor heme (yellow) are indicated. **b** The HPLC profile of authentic standard of aureothin (**1**), 7-OH-deoxyaureothin (**3**), deoxyaureothin (**4**), and NorH variants. UV detection is at 350 nm. **c** Schematic presentation of strategies to prepare hybrid NorH/AurH variants, and profiling of biotransformations. The hybrid NorH/AurH variants (left panel) correspond to HPLC profiles of extracts (right panel) of those variants. UV detection at 350 nm. **d** The HPLC profile of authentic standard of **2** and *S. albus* mutant strains. **e** AurH and NorH catalyze the biotransformation from **6** to **2**

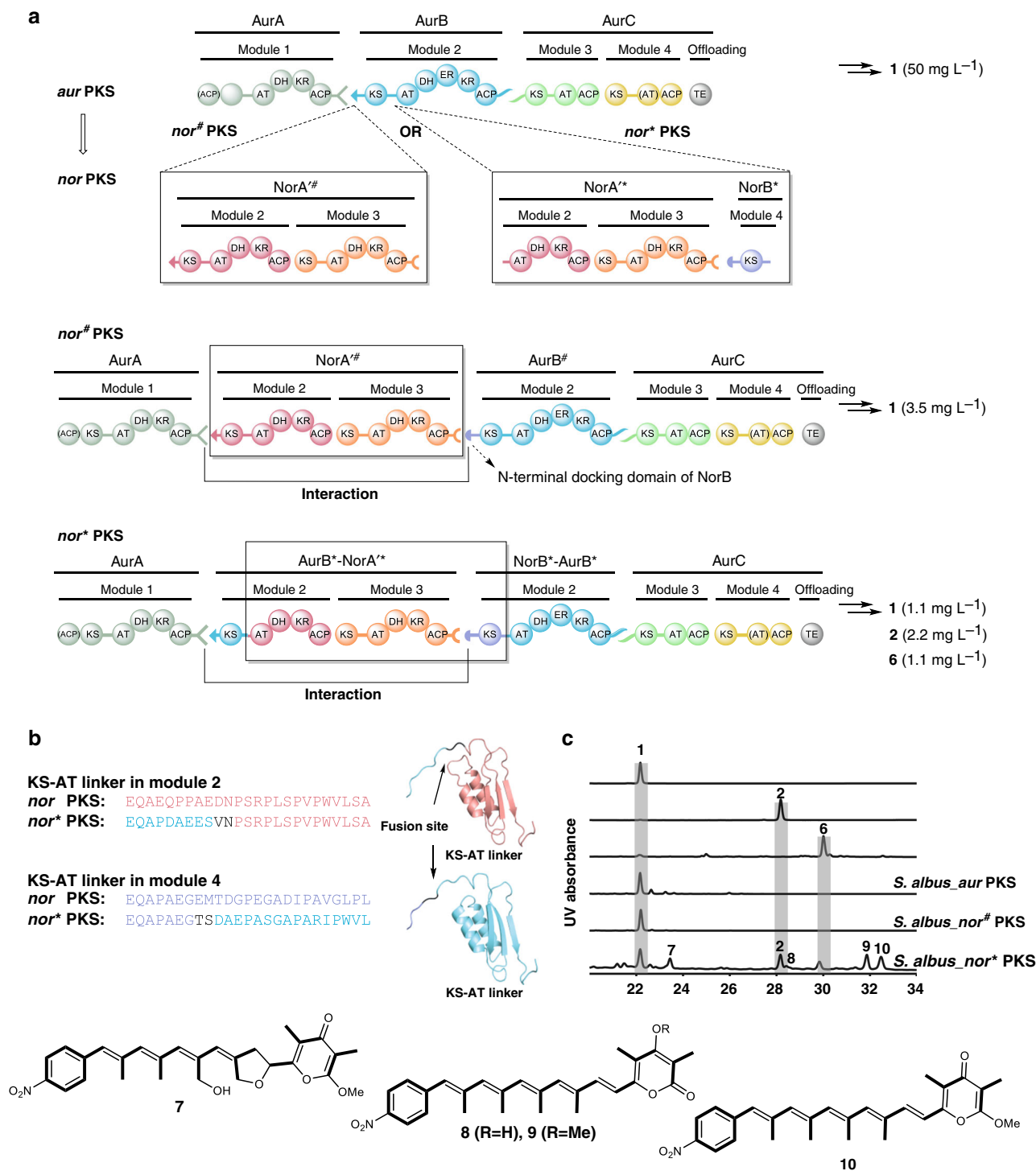


Fig. 5 Overview and metabolic profile of modified *aur* PKS variants *nor[#]* PKS and *nor** PKS. **a** Domain architectures of the *nor[#]* PKS and *nor** PKS. The N-terminal docking domain of *aurB* is swapped to that of *norB* in the *nor[#]* PKS. In the *nor** PKS, the length of recombinant KS-AT linker is manipulated to match that of the *nor* PKS. **b** The fusion sites are within the KS-AT linker and induced by the restriction enzyme *HpaI* and *SpeI*. The amino acid sequences at the fusion regions of *nor* PKS and *nor** PKS are aligned. **c** HPLC profile of authentic standard of aureothin (**1**), neo-aureothin (**2**), deoxyneoaureothin (**6**), and *S. albus* mutant strains. UV detection is at 350 nm

the *nor** PKS genes also produces a series of congeners, non-oxidized 7-deoxyneoaureothin (**6**), over-oxidized 11a-hydroxyneoaureothin (**7**), non-reduced 7-deoxy-7-dehydroneoaureothin (**10**), non-reduced 2-pyrone-7-deoxy-7-dehydroneoaureothin (**9**), and 2-pyrone-4-desmethyl-7-deoxy-7-dehydroneoaureothin (**8**). The structures of all compounds were elucidated by NMR analyses (7, Supplementary Figs. 30–34, Supplementary Tables 10

and 8, Supplementary Figs. 35–40, Supplementary Tables 11 and 9, Supplementary Figs. 41–46, Supplementary Tables 12 and 10, Supplementary Figs. 47–52, Supplementary Table 13). The presence of three highly instable, non-reduced congeners indicates that the enoylreductase domain of *AurB* cannot process the polyketide intermediate accurately, likely because the length of the intermediate is different to the original substrate. This

observation is in line with the remarkable finding that the *nor** PKS can still produce **1**. A plausible explanation could be that two modules are skipped during polyketide chain elongation. The phenomenon of PKS module skipping has previously been reported in a PKS engineering study where module 2 from the rapamycin PKS was inserted between module 1 and 2 in DEBS1-TE⁴³. In the hybrid PKS the polyketide chain underwent direct ACP-to-ACP transfer to pass through rapamycin module 2⁴⁴. In the case of *nor** PKS, the formation of the shorter chain may be rationalized by a similar ACP-to-ACP transfer or by the interaction between the C-terminal docking domain of AurA and the N-terminal docking domain of NorB, which is also present in the *nor[#]* PKS.

To corroborate this model, we generated another strain (*S. albus*:pHY129) in which the N-terminal docking domain of AurB was swapped to that of NorB without inserting NorA'[#] (Supplementary Fig. 9). By LC-MS monitoring we found that the recombinant strain produces aureothin. This experiment confirms that the C-terminal docking domain from AurA in fact recognizes the N-terminal docking domain from NorB. Thus, the shortcut in the *nor** PKS can be rationalized. It remains unclear why the class 1a C-terminal docking domain of AurA cannot interact with the class 1a N-terminal docking domain of NorA' in *S. albus_nor#*PKS. Weak interactions between different docking domain types have also been observed in other PKS systems⁴⁵.

Discussion

With the aim of rationally designing therapeutics by synthetic biology approaches⁴⁶, a giant body of knowledge on modular polyketide assembly lines has been built in the past three decades. Inspired by pioneering works on the erythromycin PKS, many natural product derivatives have been created by mutations or replacements of PKS domains and swaps of entire PKS modules or subunits⁷. Yet, information on the structures and dynamics of modular PKSs has been limited⁴⁷, and many trial-and-error engineering approaches gave unsatisfactory results because of incompatibilities, non-functional constructs and low yields. Only recently, cryo-electron microscopy provided the first complete type I PKS module structure, which may grant insights that could facilitate rational PKS engineering^{47–49}. Complementary to structure-guided engineering approaches, studying nature's strategies to evolve structural diversity may give important clues for synthetic biology. In silico analyses of PKS domains, modules, and gene clusters have suggested the impressive wealth of polyketide structures has evolved through point mutations⁵⁰, gene duplications⁵¹, homologous recombinations⁵², and horizontal gene transfer⁵³. An evolution-guided reprogramming of PKSs appears to be a promising strategy to generate desired polyketide metabolites^{54–56}.

In this study we used the related *aur* and *nor* gene clusters as model system to study and emulate potential evolutionary processes that lead to the structural diversity of complex polyketides. Through genome mining we identified nine homologous *aur*- and *nor*-type gene clusters, including one split *nor*-type gene cluster that already hinted to gene rearrangement events implicated in pathway evolution. The phylogenetic analyses of the *aur*- and *nor*-type PKSs predicted that the *aur* and *nor* gene clusters have a common evolutionary ancestor, and that the *nor* gene cluster likely represents the older type. Thus, we established a robust expression system for the expression of the *nor* biosynthetic gene cluster and emulated the possible natural *nor*-to-*aur* PKS evolution. By deleting two modules in the *nor* PKS we succeeded in the PKS morphing to generate the aureothin backbone. Yet, we unexpectedly found that the narrow substrate tolerance of the polyketide-tailoring enzyme NorH hampers *O*-heterocyclization,

the final biosynthetic step towards aureothin. Notably, multiple point mutations and enzyme domain swapping are not sufficient to adjust the enzyme function. Eventually, we succeeded to create a chimera that partially converts deoxyaureothin (**4**) into aureothin (**1**). The homolog AurH, however, was found to have broader substrate specificity than NorH and readily transforms the deoxy precursor of neo-aureothin into the final product.

According to the model of generalist-to-specialist enzyme evolution^{41,57}, the substrate specificities of the oxygenases may indicate that AurH is more ancient than the more specialized NorH (Fig. 6a). Furthermore, since the presence of the tetrahydrofuran ring is crucial for the antifungal activity of aureothin⁵⁸, the intermediate stage of an *aur* PKS and an oxygenase yielding the less active 7-OH-deoxyaureothin would not be preferred from a functionalistic point of view¹². Presuming that the post-PKS gene coevolved with the PKS genes in the same gene locus, the *aur* gene cluster would represent the ancestor. The successful reverse (*aur*-to-*nor*) PKS morphing by insertion of two modules corroborated the revised model, although it contrasts with the conclusion drawn from the phylogeny. This is a rare example of successfully engineering two additional modules into a PKS and shows that the *aur* and *nor* gene cluster share the same evolutionary origin.

The general lessons learned from the *aur/nor* engineering experiments are that (1) evolutionary models guide the morphing of modular PKSs from two distinct biosynthetic pathways into another, and that (2) the specificity of tailoring enzymes may give additional clues on the direction of evolution. Phylogenies alone should be interpreted with care. The third important lesson relates to the design of functional PKS constructs; (3) the KS-AT linker is particularly well suited as engineering site for both excision and insertion of modules. As in a plethora of other, mainly unpublished, studies, engineering the *aur* and *nor* PKSs has been a tour de force because of suboptimal recombination sites that led to less active or fully non-functional constructs. In our efforts to alter the *aur* and *nor* PKSs we found that best results for excising or inserting modules were obtained when the recombination sites were located in the KS-AT linker region. Only in this way it has been possible to create a *nor** PKS that produces the aureothin backbone, and an *aur*** PKS assembling the neo-aureothin scaffold. Still, the product yields of the engineered *aur***_PKS (aureothin backbone) and the *nor**_PKS (neo-aureothin backbone) are substantially lower (10–20-fold) than those of the corresponding native assembly lines, and could be optimized by further adjusting the KS-AT linker region⁵⁹, for example. It should be pointed out that the constructs with recombination sites within the docking domains (*aur[#]* PKS and *nor[#]* PKS) were non-functional.

Our results of the engineering experiments, specifically the location of the fusion sites, are particularly intriguing in light of a recently proposed revision of the PKS module boundaries. Contrary to the textbook architecture of modules (KS-AT-X-ACP), in the phylogeny-based model a PKS module would start with an AT domain and end with the KS domain that was previously assigned to the downstream module (AT-X-ACP-KS)^{52,60}. Notably, in the alternative model the boundary for each module is the KS-AT linker, which in our case proved to be the best site for adding or deleting modules (Fig. 6b). Therefore, our results strongly support this alternative PKS module definition, and suggest that the KS-AT linker is an alternative construction site for future PKS engineering. Although *trans*-AT PKSs⁶¹ likely evolved in a fundamentally different fashion from *cis*-AT systems like the ones investigated here, it should be noted that their mosaic structure¹⁴, the KS substrate specificities⁶², and ACP-KS relationship⁶³ are in accord with the revised module boundaries. It is remarkable that this scheme extends to the related modular

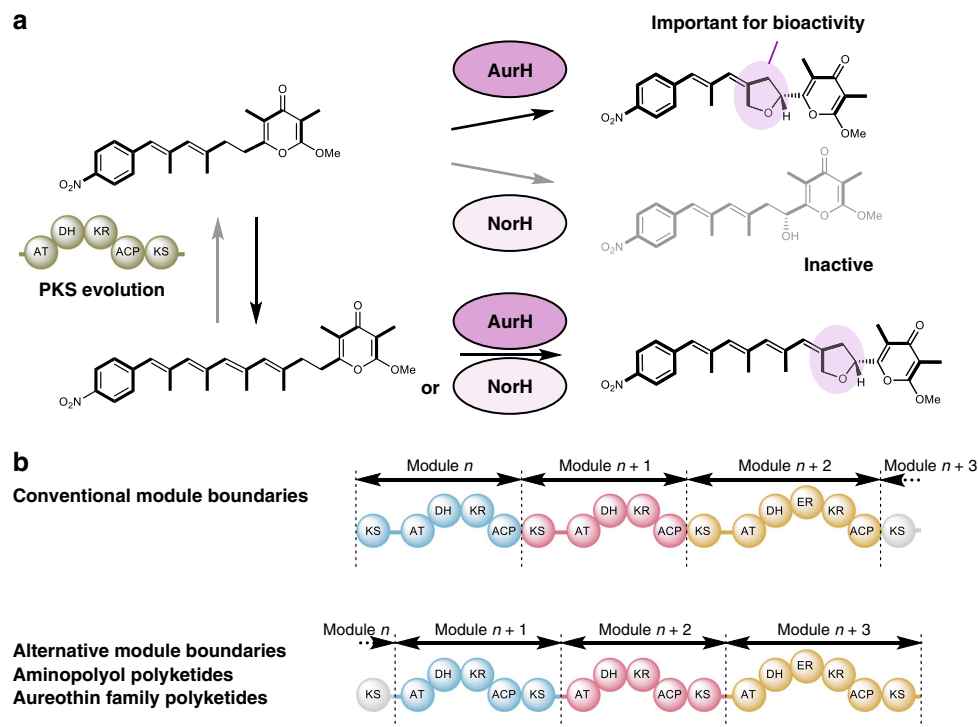


Fig. 6 Model of plausible evolutionary processes in the *aur/nor* modular type I PKSs and tailoring enzymes. **a** Insertion of two additional modules yields a functional *nor* assembly line in which bioactivity is maintained. **b** Scheme showing paradigm shift related to conventional and alternative PKS module boundaries that are in line with the proposed module insertions into the KS-AT linker regions

non-ribosomal peptide synthetases (NRPSs). Recently, a platform has been established that allows a swift recombination of NRPS modules to create artificial peptides^{64,65}. Interestingly, for the de novo design and engineering of NRPSs this strategy employs the C-A (condensation-adenylation) linker region as fusion point⁶⁶, which corresponds to the KS-AT linker in PKSs.

In conclusion, our work is a proof-of-concept for a successful evolution-guided reprogramming of modular PKSs. Beyond insight into nature's chemical diversification by evolving PKS and post-PKS enzymes, the knowledge gained supports a paradigm for module boundaries for at least some assembly lines, which may advance the field of rational PKS engineering and synthetic biology.

Methods

Strains and media. For routine subcloning, *E. coli* strains TOP10 (Invitrogen) and XL1-Blue (Agilent) were used, *E. coli* HB101/pRK2013 for triparental conjugation⁶⁷, and BW25113/pIJ790 for PCR-targeting procedures⁶⁸. For plasmid selection, *E. coli* strains were cultured in Luria-Bertani (LB) medium supplemented with spectinomycin (100 µg mL⁻¹, Sigma-Aldrich), apramycin (50 µg mL⁻¹, Sigma-Aldrich), kanamycin (25 µg mL⁻¹, Sigma-Aldrich) or chloramphenicol (25 µg mL⁻¹, Carl Roth). *S. albus* (kindly provided by Prof. Dr. Jose A. Salas, University of Oviedo, Spain) was used as host strain for heterologous expression experiments. For metabolite production, *Streptomyces* strains were grown at 28 or 30 °C with 150 or 180 rpm orbital shaking in TSB, R5 (103 g L⁻¹ sucrose, 10 g L⁻¹ glucose, 5 g L⁻¹ yeast extract, 5.73 g L⁻¹ TES, 1 g L⁻¹ casamino acids, 0.05 g L⁻¹ KH₂PO₄, 0.25 g L⁻¹ K₂SO₄, 10.12 g L⁻¹ MgCl₂·6H₂O, 2.94 g L⁻¹ CaCl₂·2H₂O, 0.28 g L⁻¹ NaOH, 1.5 ml 20% L-proline, 2 mL trace element solution) or J-medium (100 g L⁻¹ sucrose, 30 g L⁻¹ tryptic soy broth, 10 g L⁻¹ yeast extract, 4.7 g L⁻¹ MgCl₂) for 2 days. Seed cultures (10 mL) were used to inoculate 100 mL of the up-scaled culture in R5, modified R5 (R5 medium without sucrose and 1 g L⁻¹ casamino acids) or M10 medium (10 g L⁻¹ malt extract, 4 g L⁻¹ yeast extract, 4 g L⁻¹ glucose, pH 7.3). The fermentation was stopped after 5 days cultivation (orbital shaking at 28 °C). For sporulation, *Streptomyces* strains were grown on MS (mannitol soya flour) agar plates for 7 days. For conjugation MS agar plates supplemented with MgCl₂ (10 mM) were used, which were overlaid after one-day incubation with nalidixic acid (20 µg mL⁻¹, Carl Roth) and appropriate antibiotics.

General DNA procedures. All routine experiments such as PCR, DNA isolation, plasmid preparation, restriction digests, gel electrophoresis, ligation, and

transformation were performed according to standardized methods for *E. coli*⁶⁹. Restriction enzyme digested and PCR-amplified DNA fragments were purified from agarose gel using the Monarch DNA Gel Extraction Kit (New England Biolabs).

Phylogenetic tree construction. The amino acid sequences of the KS domains of the aureothin and neo-aureothin gene clusters and other selected actinobacterial PKS clusters (oligomycin, olm; tylacton, tyl; nanchangmycin, nan; monensin, mon; candicidin, fsc; pimaricin, pim; nystatin, nys; amphoterin, amp; polyene macrolide from *Streptomyces avermitilis*, pte; avermectin, ave) were aligned using the GUIDANCE2 Server²⁶. Alignment columns with a score below 0.93 were removed from the alignment. Tree reconstruction was done by the method of Bayesian inference using version 3.2.6 of the MrBayes software²⁷. The calculation used the model jumping option and was run with four independent chains for 1.5 million generations sampling one tree per 1000 generations. Trees were summarized with a burnin of 300 trees.

Construction of *S. albus_nor* PKS. The 0.8 kb *Bsa*I-*Sap*I DNA fragment from pBluescript II SK (+) (Stratagene) was blunted and ligated into *Afe*I-digested pYU01²⁸, in which *aurD* was cloned downstream of constitutive promoter, *actII-ORF470*, to yield the plasmid pYU91. The 34.5 kb *Xba*I DNA fragment from pNT87²⁸ was cloned into *Spe*I-digested pYU91. The resulting suicide vector pYU93 was introduced into *S. albus*::pNT42 containing *norB*, *C*, and *I* by triparental conjugation to give the recombinant strain *S. albus*::pNT42/pYU93 (*S. albus_nor* PKS).

Construction of *S. albus_aur*^S PKS. To remove the *norA'* gene, primer pair NorA'-HpaI-fw (5'-CCGGACGTCCTGCCACTGCGCTTCGGCGCCTGAGCGAGCGTTAACGAACAGTAGAAAGCCAGTC-3') and NorA'-HpaI-rv (5'-GCAGCCCTTCGGCGTCAGTCGGTTCGATGGAGCCGGGTTAACTTGGTCGGTCATTTTCGAACC-3') were used to amplify the kanamycin resistance cassette from pK19⁷¹. The obtained amplicon was introduced into *E. coli* BW25113/pIJ790 harboring pYU93. The targeted *norA'* gene was replaced by the kanamycin resistance cassette through the λRED system to generate pYU93 Δ*norA'*-Kan. This plasmid was restricted by *Hpa*I and then self-ligated to generate pYU93 Δ*norA'*. This plasmid was introduced via triparental mating into *S. albus*::pNT42 to generate *S. albus*::pNT42/pYU93 Δ*norA'* (*S. albus_aur*^S PKS).

Construction of *S. albus_aur*^{*} PKS and *S. albus_aur*[#] PKS. The 23.7-kb *Xba*I DNA fragments from pYU48²⁸ and pYU69²⁸ were individually ligated into *Spe*I-digested pYU91. The resulting suicide vectors pYU95 and pYU98 were introduced

into *S. albus*::pNT42 to generate *S. albus*::pNT42/pYU95 (*S. albus_aur*^{*} PKS) and *S. albus*::pNT42/pYU98 (*S. albus_aur*[#] PKS), respectively.

Construction of *S. albus_aur*^{} PKS.** The *S. albus_aur*^{**} PKS mutant was constructed by a PCR-targeting approach. To adjust the length of the KS-AT linker region in *norB*^{**}, primer pair KS2_fw_HpaI_2 (5'-GTCATCCTGGACAGCCGAGCAGCCCCGGCCGAGGACGTTAACGGAACCTTCAAGTTCCTCCGC-3') and KS4_rv_HpaI_2 (5'-CAGGGCAGGCCACGGCCGGGATGTCGCGCCCTCCGGGTTAACGGAATAGGAACCTTCATGAGC-3') were designed and used to amplify the spectinomycin resistance cassette *aadA* from pIJ778. The 1.4-kb PCR product was introduced into *E. coli* BW25113/pIJ790 harboring pNT87. Mediated by the λRED system, the targeted DNA region was replaced by the *aadA* cassette to generate pYU71. pYU71 was digested with *HpaI* to remove the *aadA* cassette and then self-ligated to yield pYU72. The 23.6 kb *XbaI* DNA fragment was cut out from pYU72 and cloned into *SpeI*-digested pYU91 to obtain suicide plasmid pYU99. Finally, pYU99 was introduced into *S. albus*::pNT42 to generate *S. albus*::pNT42/pYU99 (*S. albus_aur*^{**} PKS).

Site-directed mutagenesis of NorH. The Zero Blunt PCR Cloning Kit (Invitrogen) was used for *norH* subcloning. The QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies) was used for site-directed mutagenesis. The *E. coli-Streptomyces* shuttle vector pMR17³⁸ was used for *norH* expression in *S. albus*::pHJ68 (*DaurH*)³⁷ and as the template for site-directed mutagenesis. The primers for site-directed mutagenesis are listed in Supplementary Table 3.

Chimeric NorH variants with domain swaps. Five pairs of partially overlapping primers (Supplementary Table 5) were designed using the same strategy as for site-directed mutagenesis, except that the overlapping regions were the corresponding sequences from AurH.

Hybrid NorH/AurH variants. Two hybrid NorH/AurH variants were constructed with a fusion site (*FspI* in the corresponding genes) located at the N-terminal end of I helix (Supplementary Fig. 8). The N-terminal region of *norH* was amplified with primer pair N-NorH-fw/rv, while the C-terminal region of *aurH* was amplified with C-AurH-fw/rv. Then, the two fragments were assembled with pMR17/*SpeI* + *SspI* using the NEBuilder[®] HiFi DNA Assembly Cloning Kit to generate the N-NorH-C-AurH hybrid. The same procedure was used to generate the N-AurH-C-NorH hybrid. Primer pairs N-AurH-fw/rv and C-NorH-fw/rv-1 were employed to amplify the N-terminal region of *aurH* and C-terminal region of *norH*, respectively. The primers used in these hybrid preparations are listed in Supplementary Table 6.

Construction of Δ*norH* mutant. The PCR-targeting approach was used to replace the gene *norH* in pYU93 with the kanamycin resistance cassette. The kanamycin resistance cassette was amplified from pCR-Blunt vector (Invitrogen) with primers Kan-NorH-fw-1 (5'-GCGGCAGCAGCCGCGCACCTTGAGCGAAGACCGCTGTTACTGGGCTATCTGGACAAG-3') and Kan-NorH-rv-1 (5'-GCCCTCAA CCGGAGTTGAGATATCCGATGGCTCGCCCTGATGCGGTGTGAAATAC-3'). The PCR product was introduced into *E. coli* BW25113/pIJ790 harboring pYU93 to generate recombinant and suicide plasmid pHY42. pHY42 was introduced into *S. albus*::pNT42 to yield *norH* knock-out mutant *S. albus*::pNT42/pHY42 (Δ*norH*).

Construction of *S. albus_nor*[#] PKS. *E. coli-Streptomyces* shuttle plasmid pHJ48¹⁹ containing the *aur* gene cluster was digested with *KpnI*. The 15 kb DNA fragment was inserted into the pCR-Blunt vector to yield pHY115, which was used for further gene cloning steps. The PCR-targeting approach was used to insert the spectinomycin resistance cassette *aadA* into the point mutation site. The *aadA* cassette was amplified from pIJ778 with primer pair Spec-*norA*^{*}-fw (5'-GGGC CGGACGCCCTGCGCTGCGCTCGGCGCGGCTGAAATATTATCCGG GGATCCGTCGACC-3') and Spec-*norA*^{*}-rv (5'-GGGGTAGCGACATCCCATT GACACCACCGCAGCGGGCTCCTTAAGTGTAGGCTGGAGCTGCTTC-3'), which also introduced *SspI* and *AflIII* restriction sites into the amplicon. The 1.5 kb *aadA* cassette was introduced into *E. coli* BW25113/pIJ790 harboring pHY115. Then, this cassette was inserted into pHY115 by homologous recombination, yielding pHY119. The 16 kb *KpnI*-digested pHY119 fragment was inserted into *KpnI*-digested pHJ48-oriT fragment to obtain pHY124. The gene region for the N-terminal docking domain region of AurB was replaced by that of NorB in pHY124 to generate pHY129. A DNA fragment *norA*^{*} including the gene regions for the N- and C-terminal docking domains was constructed as follows: The main part of *norA*^{*} was cut out from pYU93 by *EcoRI* and *EcoRV*, and sequences upstream and downstream of the main part of *norA*^{*} were PCR-amplified. The three DNA fragments were ligated into the pCR-Blunt vector to obtain pHY125. The 11 kb DNA fragment was cut out from pHY125 by *XbaI* and *EcoRI* and ligated to pBU12²³/*XbaI* + *EcoRI* fragment to yield pHY132. pHY129 and pHY132 were co-introduced into *S. albus* by triparental conjugation to generate *S. albus*::pHY129/pHY132 (*S. albus_nor*[#] PKS).

Construction of *S. albus_nor*^{*} PKS. The *norA*^{*} fragment harboring the gene region coding for the *norA*^{*}-AT2 domain to *norB*-KS4 domain, and the KS-AT linker in *aurB* was chosen as insertion site. The strategy to induce the *HpaI* and *SpeI* restriction site in pHJ48 was the same as above, with primer pair Spec-*aurB*-fw (5'-CCACCTGATCCTCGAACAGGCCCGGATGCGGAGGAGTCCGGTTAAC AT TCCGGGGATCCGTCGACC-3') and Spec-*aurB*-rv (5'-GGATCCGTCGCGG CGGCGCGGACGCGGGCTCGGCTACTAGTGTAGGCTGGAGCT GCT TC-3'), and the resulting *aadA* cassette was inserted into pHJ48-oriT to yield pHY133. A DNA fragment coding for the *norA*^{*}-AT2 domain to *norB*-KS4 domain was cloned and inserted into the pCR-Blunt vector to produce pHY121. The 11-kb *HpaI/SpeI*-digested DNA fragment of pHY121 was ligated into the *HpaI/SpeI* fragment of pHY133. The resulting plasmid (pHY134) was introduced into *S. albus* by triparental conjugation to generate *S. albus*::pHY134 (*S. albus_nor*^{*} PKS).

Analysis of *nor* gene cluster in *S. scabrisporus* by PCR. Strain *S. scabrisporus* (DSM41855) was cultured in TSB medium (25 mL) with orbital shaking at 30 °C for 3 days. The genomic DNA was purified by the Wizard[®] Genomic DNA Purification Kit (Promega). The joint region between *norA*^{*} and *norC* was PCR-amplified using four primer pairs: NorAAfw-Ss (5'-TGCGATGGCGCTGCAGC ACG-3')/NorCrv-Ss (5'-GACAACACACTGGCGCGGC-3'), NorAAfw-Ss/NorAArv2-Ss (5'-CAACGACGCCGAACA CGCCG-3'), NorAAfw2-Ss (5'-GCCTG CTGGACTCGCTGTG-3')/NorCrv-Ss, and NorAAfw2-Ss/NorCrv2-Ss. NorAAfw-Ss and NorAAfw2-Ss are derived from *norA*^{*}, while NorCrv-Ss and NorCrv2-Ss are derived from *norC*. The PCR reaction was performed by KAPA 2 G fast HotStart DNA polymerase (KAPABIOSYSTEMS) with 2.5% DMSO.

General analytical procedures. NMR spectra of compounds dissolved in CDCl₃ or CDCl₃/CD₃OD were measured on Bruker Avance DRX 500 or 600 MHz spectrometers equipped with a cryo probe at 300 K. The residual solvent peak ($\delta_{\text{H}} = 7.24$, $\delta_{\text{C}} = 77.7$) was used as reference. LC-HRMS measurements were carried out on a Thermo Fisher Scientific Exactive Orbitrap equipped with an electrospray ion source. Column: Betasil 100-3 C18 (2.1 × 150 mm); elution gradient: solvent A: water + 0.1% formic acid, solvent B: acetonitrile, gradient: 5% B for 1 min, 5–98% B in 15 min, 98% B for 3 min, flow rate: 0.2 mL min⁻¹.

HPLC analysis of *S. albus* mutant strains. To the bacterial culture 1 V ethyl acetate was added, and the mixture was stirred at room temperature overnight. Then, the organic phase was dried over sodium sulfate and concentrated to dryness under reduced pressure. The residue was dissolved in 1 mL methanol. The obtained extract (10 μL) was analyzed by reversed-phase column HPLC (Symmetry C18 5 μm, 4.6 × 150 mm, Waters) using a gradient program (solvent A (water + 0.1% trifluoroacetic acid) and solvent B (acetonitrile), 20% B for 5 min, to 99% B in 30 min and kept for 5 min at a flow rate of 1 mL min⁻¹). Aureothin (350 nm), 7-deoxyaureothin (350 nm), 7-hydroxydeoxyaureothin (350 nm), neo-aureothin (378 nm), intermediate (338 nm), and 7-deoxyneo-aureothin (378 nm) productions were quantified by calculating peak area sizes compared to those of references.

HPLC analysis of *S. scabrisporus*. *S. scabrisporus* DSM41855 was cultured on five PDA (potato dextrose agar) plates for 7 days. Agar was chopped and stirred in 100 mL ethyl acetate overnight. The extracts were dried over sodium sulfate and concentrated under reduced pressure. The residue was dissolved in 0.5 mL methanol. The reversed-phase HPLC analysis was performed as described above.

Isolation of intermediate compound 5 from *S. albus_aur*[#] PKS. The ethyl acetate extract (1.5 L) of a *S. albus*::pNT42/pYU93 Δ*norA*^{*} (*aur*[#] PKS) culture (1.5 L) was filtered through Celite 545. The water phase was further extracted with ethyl acetate (2 × 1.5 L). The combined extracts were dried over anhydrous sodium sulfate and concentrated under reduced pressure. The resulting residue was subjected to silica gel column chromatography (3 × 13 cm, silica gel 60, 0.040–0.063 mm) using a dichloromethane/methanol gradient. The dichloromethane-eluted fractions were subjected to preparative RP-HPLC (column, Nucleosil 100-7C18, 20 × 250 mm; flow rate, 16 mL min⁻¹; gradient, solvent A (water containing 0.1% trifluoroacetic acid), B (83% acetonitrile); 20% B in 10 min to 100% B in 30 min and kept for 15 min) to yield 5 (1.0 mg). Pale yellow solid; HRMS: *m/z* [M-H]⁻ = 246.0770 (calculated for C₁₃H₁₂NO₄, 246.0761); NMR data see Supplementary Table 8 and Supplementary Figs. 19–23.

Isolation of compound 6 deoxyneo-aureothin. A 2-L culture of *S. albus*::pNT42/pHY42 was extracted with ethyl acetate (3 × 2 L). The organic phases were combined, dried over sodium sulfate, and concentrated under reduced pressure. The extract was subjected to silica gel column chromatography (3 × 15 cm, Silica gel 60, 0.040–0.063 mm) using a dichloromethane/methanol gradient. Three fractions containing 6 were combined and subjected to preparative RP-HPLC (column, Supercosil 100-5C18, 21.2 × 250 mm; flow rate, 12 mL min⁻¹; gradient, solvent A (40% acetonitrile), B (100% acetonitrile); 1% B in 10 min to 100% B in 30 min and kept for 10 min) to yield crude 6 (18.9 mg). The crude product was subjected to semipreparative RP-HPLC (column, Phenomenex Phenyl-Hexyl, particle size 5 μm, pore size 100 Å, 10 × 250 mm; flow rate, 6 mL min⁻¹; gradient, solvent A

(40% acetonitrile), B (100% acetonitrile), 1% B in 10 min to 100% B in 30 min and kept for 10 min) yielding deoxyneoureothin (**6**) (7.7 mg). Yellow solid; HRMS: m/z $[M + H]^+ = 464.2426$ (calculated for $C_{28}H_{34}NO_5$, 464.2431); NMR data see Supplementary Table 4 and Supplementary Figs. 24–29.

Isolation of congeners from *S. albus_nor*[#] PKS. All isolation steps were carried out in the dark to avoid the degradation of polyketides. The (emulsion-forming) ethyl acetate extract (4.5 L) of a *S. albus* pHY134 (*nor*[#] PKS) culture was filtered through Celite 545, and the water phase was again extracted with ethyl acetate (2 × 4.5 L). The combined extracts were dried over anhydrous sodium sulfate, concentrated under reduced pressure and subjected to chromatography on silica gel (column, 5 × 16 cm, Silica gel 60, 0.040–0.063 mm; gradient, dichloromethane/methanol). Further purification was achieved by preparative RP-HPLC (column, Nucleosil 100-7C18, 20 × 250 mm; flow rate, 16 mL min⁻¹; gradient, solvent A (water containing 0.1% trifluoroacetic acid), B (83% acetonitrile); 40% B in 5 min to 100% B in 25 min and kept for 15 min) to yield crude **9**. This fraction was further subjected to semipreparative RP-HPLC (column, Nucleodur C18 HTec, 10 × 250 mm; flow rate, 6 mL min⁻¹) using an isocratic system 83% acetonitrile/water containing 0.1% trifluoroacetic acid (80:20) to yield **9** (3.5 mg). NMR and MS data see Supplementary Figs. 18 and 41–46, Supplementary Table 12. A fraction containing **8** and **10** was subjected to preparative RP-HPLC (column, Nucleosil 100-7C18, 20 × 250 mm; flow rate, 12 mL min⁻¹; gradient, solvent A (water containing 0.1% trifluoroacetic acid), B (83% acetonitrile); 20% B in 10 min to 100% B in 30 min and kept for 15 min) to yield crude **8** and **10**, respectively. Crude **10** was further purified by semipreparative RP-HPLC (column, Nucleodur C18 HTec, 10 × 250 mm; flow rate, 6 mL min⁻¹; eluent, isocratic system 83% acetonitrile/water containing 0.1% trifluoroacetic acid, *v/v* 90:10) to yield **10** (5.5 mg). NMR and MS data see Supplementary Figs. 18 and 47–52, Supplementary Table 13. Crude **8** was further purified by semipreparative RP-HPLC (column, Nucleodur C18 HTec, 10 × 250 mm; flow rate, 6 mL min⁻¹; eluent, isocratic system 83% acetonitrile/water containing 0.1% trifluoroacetic acid, *v/v* 80:20, followed by isocratic system composed of 83% acetonitrile/water containing 0.1% trifluoroacetic acid, *v/v* 70:30) to yield a mixture containing **8**. This fraction was further subjected to semipreparative RP-HPLC (column, Phenomenex Phenyl-Hexyl, 10 × 250 mm; flow rate, 6 mL min⁻¹; eluent, isocratic system methanol/water containing 0.1% trifluoroacetic acid, *v/v* 80:20) to yield **8** (0.8 mg). NMR and MS data see Supplementary Figs. 18 and 35–40, Supplementary Table 11. The fraction containing **7** was subjected to preparative RP-HPLC (column, Nucleosil 100-7C18, 20 × 250 mm; flow rate, 16 mL min⁻¹; gradient system, solvent A (water containing 0.1% trifluoroacetic acid), B (83% acetonitrile); 40% B in 5 min to 100% B in 25 min and kept for 15 min) to yield crude **7**. Final purification was achieved by another semipreparative RP-HPLC (column, Nucleodur C18 HTec, 10 × 250 mm; flow rate, 6 mL min⁻¹; eluent, isocratic system 83% acetonitrile/water containing 0.1% trifluoroacetic acid, *v/v* 55:45) to yield **7** (2.7 mg). NMR and MS data see Supplementary Figs. 18 and 30–34, and Supplementary Table 10. Due to instability of pure congeners **8** to **10**, these compounds rapidly degraded after NMR measurements.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Data supporting the findings of this work are available within the paper and its Supplementary Information files. A reporting summary for this Article is available as a Supplementary Information file. The datasets generated and analyzed during the current study are available from the corresponding author upon request. The source data underlying Figs. 2c and 3b, as well as Supplementary Figs. 1 and 2 are provided as a Source Data file.

Received: 28 September 2018 Accepted: 12 August 2019

Published online: 02 September 2019

References

- Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs from 1981 to 2014. *J. Nat. Prod.* **79**, 629–661 (2016).
- Hertweck, C. The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed.* **48**, 4688–4716 (2009).
- Keatinge-Clay, A. T. The structures of type I polyketide synthases. *Nat. Prod. Rep.* **29**, 1050–1073 (2012).
- Helfrich, E. J., Reiter, S. & Piel, J. Recent advances in genome-based polyketide discovery. *Curr. Opin. Biotechnol.* **29**, 107–115 (2014).
- Klaus, M. & Grninger, M. Engineering strategies for rational polyketide synthase design. *Nat. Prod. Rep.* **35**, 1070–1081 (2018).
- Barajas, J. F., Blake-Hedges, J. M., Bailey, C. B., Curran, S. & Keasling, J. D. Engineered polyketides: synergy between protein and host level engineering. *Synth. Syst. Biotechnol.* **2**, 147–166 (2017).
- Weissman, K. J. Genetic engineering of modular PKSs: From combinatorial biosynthesis to synthetic biology. *Nat. Prod. Rep.* **33**, 203–230 (2016).
- Wong, F. T. & Khosla, C. Combinatorial biosynthesis of polyketides - a perspective. *Curr. Opin. Chem. Biol.* **16**, 117–123 (2012).
- Xu, W., Qiao, K. & Tang, Y. Structural analysis of protein-protein interactions in type I polyketide synthases. *Crit. Rev. Biochem. Mol. Biol.* **48**, 98–122 (2013).
- Menzel, H. G. et al. Combinatorial polyketide biosynthesis by *de novo* design and rearrangement of modular polyketide synthase genes. *Nat. Biotechnol.* **23**, 1171–1176 (2005).
- Fischbach, M. A., Walsh, C. T. & Clardy, J. The evolution of gene collectives: How natural selection drives chemical innovation. *Proc. Natl Acad. Sci. USA* **105**, 4601–4608 (2008).
- Jenke-Kodama, H. & Dittmann, E. Evolution of metabolic diversity: Insights from microbial polyketide synthases. *Phytochemistry* **70**, 1858–1866 (2009).
- Ginolhac, A. et al. Type I polyketide synthases may have evolved through horizontal gene transfer. *J. Mol. Evol.* **60**, 716–725 (2005).
- Nguyen, T. et al. Exploiting the mosaic structure of *trans*-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nat. Biotechnol.* **26**, 225–233 (2008).
- Medema, M. H., Cimermanic, P., Sali, A., Takano, E. & Fischbach, M. A. A systematic computational analysis of biosynthetic gene cluster evolution: Lessons for engineering biosynthesis. *PLoS Comput. Biol.* **10**, e1004016 (2014).
- Hirata, Y., Nakata, H., Yamada, K., Okuhara, K. & Naito, T. The structure of aureothin, a nitro compound obtained from *Streptomyces thioluteus*. *Tetrahedron* **14**, 252–274 (1961).
- Cassinelli, G., Grein, A., Orezzi, P., Pennella, P. & Sanfilippo, A. New antibiotics produced by *Streptoverticillium orinoci*, n. sp. *Arch. Microbiol.* **55**, 358–368 (1967).
- He, J. & Hertweck, C. Biosynthetic origin of the rare nitro aryl moiety of the polyketide antibiotic aureothin: Involvement of an unprecedented *N*-oxygenase. *J. Am. Chem. Soc.* **126**, 3694–3695 (2004).
- He, J. & Hertweck, C. Iteration as programmed event during polyketide assembly; molecular analysis of the aureothin biosynthesis gene cluster. *Chem. Biol.* **10**, 1225–1232 (2003).
- Traitcheva, N., Jenke-Kodama, H., He, J., Dittmann, E. & Hertweck, C. Non-colinear polyketide biosynthesis in the aureothin and neo-aureothin pathways: An evolutionary perspective. *ChemBioChem* **8**, 1841–1849 (2007).
- Müller, M., He, J. & Hertweck, C. Dissection of the late steps in aureothin biosynthesis. *ChemBioChem* **7**, 37–39 (2006).
- He, J. & Hertweck, C. Functional analysis of the aureothin iterative type I polyketide synthase. *ChemBioChem* **6**, 908–912 (2005).
- Busch, B., Ueberschaar, N., Sugimoto, Y. & Hertweck, C. Interchain retrotransfer of aureothin intermediates in an iterative polyketide synthase module. *J. Am. Chem. Soc.* **134**, 12382–12385 (2012).
- Busch, B. et al. Multifactorial control of iteration events in a modular polyketide assembly line. *Angew. Chem. Int. Ed.* **52**, 5285–5289 (2013).
- Busch, B. & Hertweck, C. Evolution of metabolic diversity in polyketide-derived pyrones: using the non-colinear aureothin assembly line as a model system. *Phytochemistry* **70**, 1833–1840 (2009).
- Sela, I., Ashkenazy, H., Katoh, K. & Pupko, T. GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucleic Acids Res.* **43**, W7–W14 (2015).
- Ronquist, F. et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
- Sugimoto, Y. et al. Freedom and constraint in engineered noncolinear polyketide assembly lines. *Chem. Biol.* **22**, 229–240 (2015).
- Broadhurst, R. W., Nietlispach, D., Wheatcroft, M. P., Leadlay, P. F. & Weissman, K. J. The structure of docking domains in modular polyketide synthases. *Chem. Biol.* **10**, 723–731 (2003).
- Buchholz, T. J. et al. Structural basis for binding specificity between subclasses of modular polyketide synthase docking domains. *ACS Chem. Biol.* **4**, 41–52 (2009).
- Miyayama, A., Kudo, F. & Eguchi, T. Protein-protein interactions in polyketide synthase-nonribosomal peptide synthetase hybrid assembly lines. *Nat. Prod. Rep.* **35**, 1185–1209 (2018).
- Derival, J. et al. Characterization of intersubunit communication in the virginiamycin *trans*-Acyl transferase polyketide synthase. *J. Am. Chem. Soc.* **138**, 4155–4167 (2016).
- Whicher, J. R. et al. Cyanobacterial polyketide synthase docking domains: A tool for engineering natural product biosynthesis. *Chem. Biol.* **20**, 1340–1351 (2013).
- Ridley, C. P., Lee, H. Y. & Khosla, C. Evolution of polyketide synthases in bacteria. *Proc. Natl Acad. Sci. USA* **105**, 4595–4600 (2008).
- Jenke-Kodama, H., Börner, T. & Dittmann, E. Natural biocombinatorics in the polyketide synthase genes of the actinobacterium *Streptomyces avermitilis*. *PLoS Comput. Biol.* **2**, e132 (2006).

36. Sugimoto, Y., Ding, L., Ishida, K. & Hertweck, C. Rational design of modular polyketide synthases: morphing the aureothin pathway into a luteoreticulin assembly line. *Angew. Chem. Int. Ed.* **53**, 1560–1564 (2014).
37. He, J., Müller, M. & Hertweck, C. Formation of the aureothin tetrahydrofuran ring by a bifunctional cytochrome P450 monooxygenase. *J. Am. Chem. Soc.* **126**, 16742–16743 (2004).
38. Richter, M. E., Traitcheva, N., Knupfer, U. & Hertweck, C. Sequential asymmetric polyketide heterocyclization catalyzed by a single cytochrome P450 monooxygenase (AurH). *Angew. Chem. Int. Ed.* **47**, 8872–8875 (2008).
39. Zocher, G., Richter, M. E., Müller, U. & Hertweck, C. Structural fine-tuning of a multifunctional cytochrome P450 monooxygenase. *J. Am. Chem. Soc.* **133**, 2292–2302 (2011).
40. Richter, M. E. A., Busch, B., Ishida, K., Moore, B. S. & Hertweck, C. Pyran formation by an atypical CYP-mediated four-electron oxygenation-cyclization cascade in an engineered aureothin pathway. *ChemBioChem* **13**, 2196–2199 (2012).
41. Jensen, R. A. Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* **30**, 409–425 (1976).
42. Newton, M. S., Arcus, V. L., Gerth, M. L. & Patrick, W. M. Enzyme evolution: Innovation is easy, optimization is complicated. *Curr. Opin. Struct. Biol.* **48**, 110–116 (2018).
43. Rowe, C. J. et al. Engineering a polyketide with a longer chain by insertion of an extra module into the erythromycin-producing polyketide synthase. *Chem. Biol.* **8**, 475–485 (2001).
44. Thomas, I., Martin, C. J., Wilkinson, C. J., Staunton, J. & Leadlay, P. F. Skipping in a hybrid polyketide synthase: evidence for ACP-to-ACP chain transfer. *Chem. Biol.* **9**, 781–787 (2002).
45. Thattai, M., Burak, Y. & Shraiman, B. I. The origins of specificity in polyketide synthase protein interactions. *PLoS Comput. Biol.* **3**, 1827–1835 (2007).
46. Bayly, C. L. & Yadav, V. G. Towards precision engineering of canonical polyketide synthase domains: recent advances and future prospects. *Molecules* **22**, E235 (2017).
47. Weissman, K. J. Uncovering the structures of modular polyketide synthases. *Nat. Prod. Rep.* **32**, 436–453 (2015).
48. Dutta, S. et al. Structure of a modular polyketide synthase. *Nature* **510**, 512–517 (2014).
49. Whicher, J. R. et al. Structural rearrangements of a polyketide synthase module during its catalytic cycle. *Nature* **510**, 560–564 (2014).
50. Wu, K., Chung, L., Revill, W. P., Katz, L. & Reeves, C. D. The *FK520* gene cluster of *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. *Gene* **251**, 81–90 (2000).
51. Stinear, T. P. et al. Giant plasmid-encoded polyketide synthases produce the macrolide toxin of *Mycobacterium ulcerans*. *Proc. Natl Acad. Sci. USA* **101**, 1345–1349 (2004).
52. Zhang, L. et al. Characterization of giant modular PKSs provides insight into genetic mechanism for structural diversification of aminopolyl polyketides. *Angew. Chem. Int. Ed.* **56**, 1740–1745 (2017).
53. Anzai, Y. et al. Organization of the biosynthetic gene cluster for the polyketide macrolide mycinamicin in *Micromonospora griseorubida*. *FEMS Microbiol. Lett.* **218**, 135–141 (2003).
54. Wlodek, A. et al. Diversity oriented biosynthesis via accelerated evolution of modular gene clusters. *Nat. Commun.* **8**, 1206 (2017).
55. Awakawa, T. et al. Reprogramming of the antimycin NRPS-PKS assembly lines inspired by gene evolution. *Nat. Commun.* **9**, 3534 (2018).
56. Chemler, J. A. et al. Evolution of efficient modular polyketide synthases by homologous recombination. *J. Am. Chem. Soc.* **137**, 10603–10609 (2015).
57. Carbonell, P., Lecointre, G. & Faulon, J. L. Origins of specificity and promiscuity in metabolic networks. *J. Biol. Chem.* **286**, 43994–44004 (2011).
58. Werneburg, M. et al. Exploiting enzymatic promiscuity to engineer a focused library of highly selective antifungal and antiproliferative aureothin analogues. *J. Am. Chem. Soc.* **132**, 10407–10413 (2010).
59. Yuzawa, S. et al. Comprehensive *in vitro* analysis of acyltransferase domain exchanges in modular polyketide synthases and its application for short-chain ketone production. *ACS Synth. Biol.* **6**, 139–147 (2017).
60. Keatinge-Clay, A. T. Polyketide synthase modules redefined. *Angew. Chem. Int. Ed.* **56**, 4658–4660 (2017).
61. Helfrich, E. J. & Piel, J. Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat. Prod. Rep.* **33**, 231–316 (2016).
62. Jenner, M. et al. Substrate specificity in ketosynthase domains from *trans*-AT polyketide synthases. *Angew. Chem. Int. Ed.* **52**, 1143–1147 (2013).
63. Vander Wood, D. A. & Keatinge-Clay, A. T. The modules of *trans*-acyltransferase assembly lines redefined with a central acyl carrier protein. *Proteins* **86**, 664–675 (2018).
64. Cai, X. et al. Entomopathogenic bacteria use multiple mechanisms for bioactive peptide library design. *Nat. Chem.* **9**, 379–386 (2017).
65. Bozhüyük, K. A. J. et al. Modification and de novo design of non-ribosomal peptide synthetases using specific assembly points within condensation domains. *Nat. Chem.* **11**, 653–661 (2019).
66. Bozhüyük, K. A. J. et al. De novo design and engineering of non-ribosomal peptide synthetases. *Nat. Chem.* **10**, 275–281 (2018).
67. Ditta, G., Stanfield, S., Corbin, D. & Helinski, D. Broad host range DNA cloning system for Gram-negative bacteria: Construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl Acad. Sci. USA* **77**, 7347–7351 (1980).
68. Gust, B., Challis, G. L., Fowler, K., Kieser, T. & Chater, K. F. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl Acad. Sci. USA* **100**, 1541–1546 (2003).
69. Michael, R. G. & Joseph, S. *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor Laboratory Press, New York, 2012).
70. Arias, P., Fernandez-Moreno, M. A. & Malpartida, F. Characterization of the pathway-specific positive transcriptional regulator for actinorhodin biosynthesis in *Streptomyces coelicolor* A3(2) as a DNA-binding protein. *J. Bacteriol.* **181**, 6958–6968 (1999).
71. Pridmore, R. D. New and versatile cloning vectors with kanamycin-resistance marker. *Gene* **56**, 309–312 (1987).

Acknowledgements

This work was financially supported by the China Scholarship Council (CSC, fellowship to H.P.) and the German Science Foundation (DFG, Leibniz Award to C.H.). We thank A. Perner for HPLC-MS analysis, and H. Heinecke for NMR measurements.

Author contributions

All authors were involved in the design of the project. H.P., Y.S., and K.I. constructed plasmids, generated recombinant strains, and fermented strains. Y.S., H.P., and K.I. extracted cultures, and K.I. performed HPLC analyses and structural elucidation. H.J.-K. performed *in silico* analyses. All authors contributed to the early drafts of the manuscript. C.H. designed research and wrote the final version.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-019-11896-1>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Peer review information: *Nature Communications* thanks Kira Weissman and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019