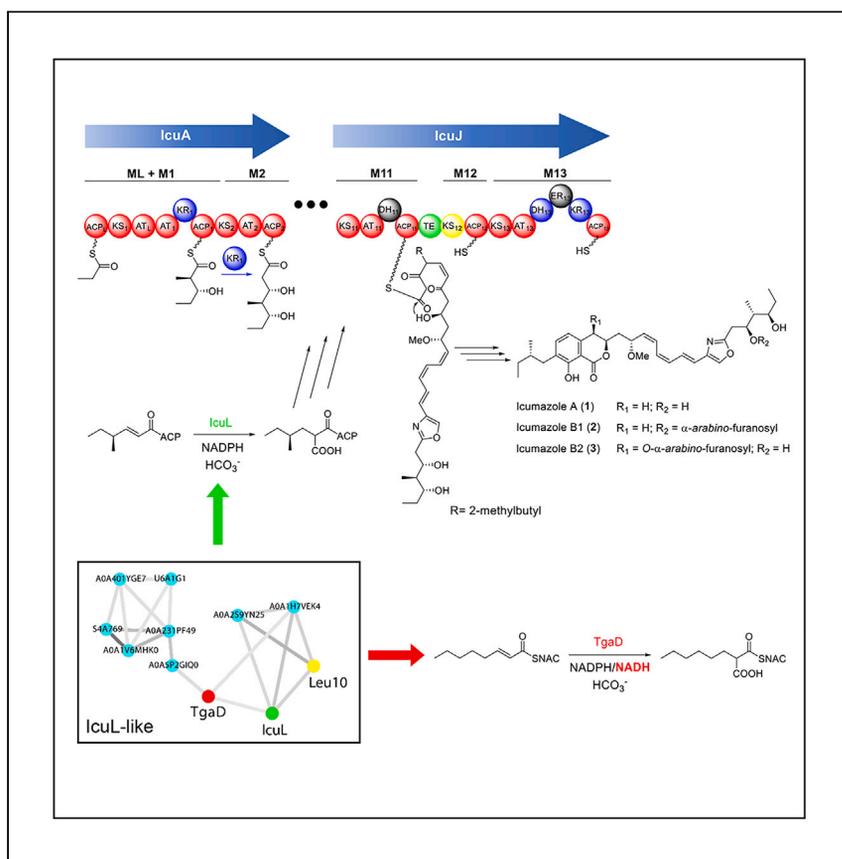


Report

Insights into the biosynthesis of icumazole, unveiling a distinctive family of crotonyl-CoA carboxylase/reductase



Feng Xie, Alexander F. Kiefer, Anna K.H. Hirsch, Olga V. Kalinina, Chengzhang Fu, Rolf Müller

chengzhang.fu@helmholtz-hips.de (C.F.)
rolf.mueller@helmholtz-hips.de (R.M.)

Highlights

Icumazoles are biosynthesized via a modular PKS/NRPS pathway

The side chain of the isochromanone derives from a trimodular PKS and an unusual CCR

The members from the new family CCR show strict substrate specificity

One unique member TgaD possesses the uncommon NADH-utilizing property

Xie et al. report the PKS/NRPS biosynthesis of the antifungal icumazole, including the biogenesis of an unusual PKS precursor generated via IcuL, which belongs to a distinguished CCR family featuring strict substrate specificity. One unique member, TgaD, utilizes NADH and NADPH, which differentiates the enzyme from all canonical NADPH-specific CCRs.

Report

Insights into the biosynthesis of icumazole, unveiling a distinctive family of crotonyl-CoA carboxylase/reductase

Feng Xie,¹ Alexander F. Kiefer,¹ Anna K.H. Hirsch,^{1,2,3} Olga V. Kalinina,^{1,4,5} Chengzhang Fu,^{1,2,*} and Rolf Müller^{1,2,3,6,*}

SUMMARY

Icumazoles are potent antifungal polyketides with intriguing structural features. Here, we present the polyketide synthase (PKS)/non-ribosomal peptide synthetase (NRPS) hybrid biosynthetic gene cluster of icumazoles. Surprisingly, an unusual nonterminal thioesterase domain divides the PKS/NRPS assembly line. The succeeding PKS modules potentially form a rare precursor 4-methyl-2-hexenoyl-ACP, thus deviating from the previously proposed polyoxypeptin pathway. The 4-methyl-2-hexenoyl-ACP is further reductively carboxylated to 2-methylbutylmalonyl-ACP, essential for icumazole biosynthesis by IcuL. We characterize IcuL and its homologs TgaD and Leu10 *in vitro*, suggesting a stricter substrate specificity of this new family of crotonyl carboxylases/reductases (CCRs) than found in canonical ones. Intriguingly, we also find that TgaD unprecedentedly utilizes both nicotinamide adenine dinucleotide phosphate (NADPH) and NADH as cofactors with similar efficiency, diverging from the NADPH-specific characteristic of canonical CCRs. Furthermore, a sequence similarity network-based and phylogenetic bioinformatic survey reveals that the IcuL-like CCRs are evolutionarily separated from canonical CCRs.

INTRODUCTION

Natural products are one of the most important sources for human and veterinary drug discovery as well as agrochemicals.¹ Gram-negative myxobacteria are prolific producers of novel bioactive small molecules. Numerous secondary metabolites with various bioactivities have been characterized,^{2,3} such as anticancer drug epothilones,^{4,5} broad-spectrum antibacterial cystobactamids,^{6,7} HIV inhibitor aetheramides,⁸ C1 metabolism-targeting biofilm inhibitor carolacton,^{9,10} and antifungal icumazoles.¹¹

Icumazole A (1) and its congeners B1 (2) and B2 (3) featuring isochromanone, triene, and oxazole moieties were isolated from *Sorangium cellulosum* So ce701.¹¹ However, the biosynthesis of icumazoles remained elusive. Icumazoles are polyketides containing an oxazole ring, as often observed in microbial natural products produced by hybrid biosynthetic gene clusters (BGCs) of polyketide synthases (PKSs) and nonribosomal peptides synthetases (NRPSs). Type I PKSs and NRPSs are multi-domain megaenzymes organized into catalytic modules.¹² Each module consists of multiple catalytic domains responsible for incorporating one building block to extend and modify the nascent intermediates tethered to the biosynthetic enzymes via acyl/peptidyl carrier protein (ACP/PCP) domains. Closer scrutiny of

¹Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research (HZI), and Department of Pharmacy, Saarland University, 66123 Saarbrücken, Germany

²Helmholtz International Lab for Anti-Infectives, Helmholtz Center for Infection Research, 38124 Braunschweig, Germany

³German Centre for Infection Research (DZIF), 38124 Braunschweig, Germany

⁴Medical Faculty, Saarland University, 66421 Homburg, Germany

⁵Center for Bioinformatics, Saarland Informatics Campus, 66123 Saarbrücken, Germany

⁶Lead contact

*Correspondence: chengzhang.fu@helmholtz-hips.de (C.F.), rolf.mueller@helmholtz-hips.de (R.M.)
<https://doi.org/10.1016/j.xcrp.2023.101394>



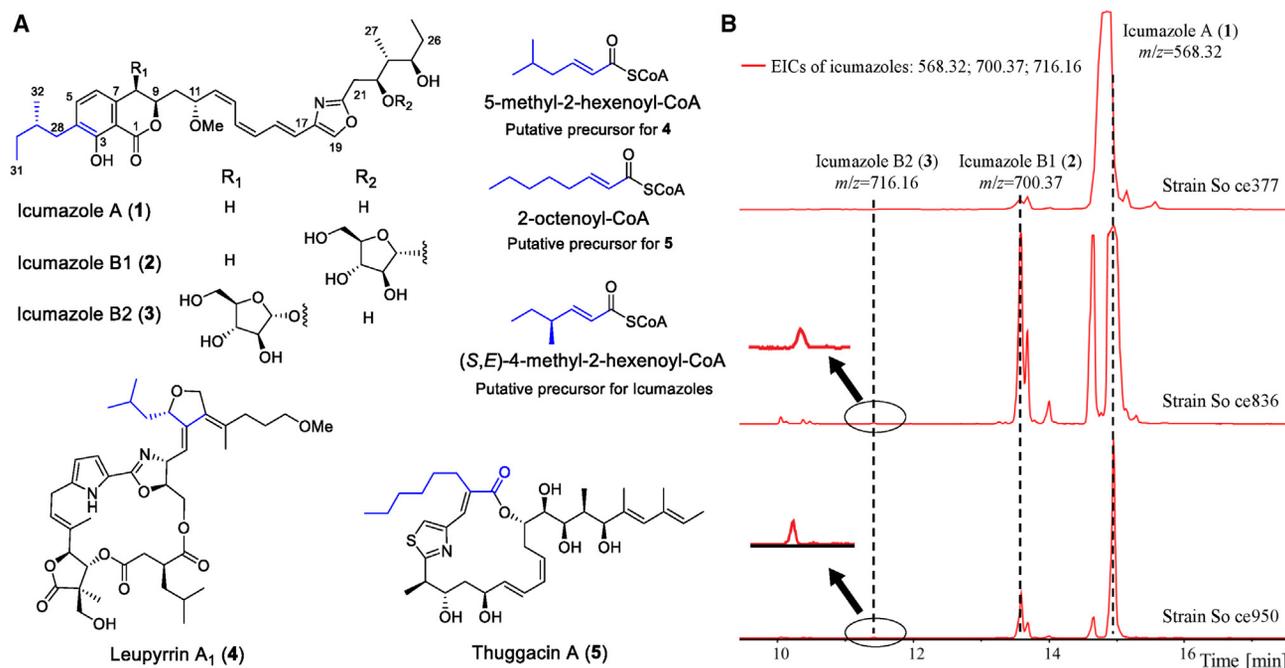


Figure 1. Myxobacterial compounds with side chains introduced by noncanonical CCRs

(A) Representative compound structures and their putative precursors. The carbon atoms derived from α -alkylmalonyl-CoAs are marked in blue. (B) Production of icumazoles in three *Sorangium* strains. Extracted ion chromatograms (EICs) of icumazoles from respective strains.

the chemical structures of icumazoles suggested that the biosynthesis of their 2-methylbutane side chain requires an unusual 2-(S)-2-(2-methylbutyl)malonyl-coenzyme A (MBM-CoA) as a PKS building block (Figure 1A). To date, different biosynthetic pathways of α -alkylmalonyl-CoAs are reported, which include the reductive carboxylation of α,β -unsaturated acyl-CoAs by crotonyl carboxylases/reductases (CCRs; also termed enoyl-CoA or enoyl-thioester carboxylases/reductases [ECRs]),^{13–18} the carboxylation of acyl-CoAs by acyl-CoA carboxylases (YCCs),¹⁹ and direct ligation of malonate and CoA by malonyl-CoA ligases.²⁰ Additionally, alcohol dehydrogenases (ADHs) with very low sequence similarity to the well-studied CCRs were proposed to carry out the formation of the corresponding alkylmalonyl-CoAs in pathways of myxobacterial compounds leupyrrin A₁ (4)²¹ and thuggacin A (5)²² (Figure 1A).

In this work, we discovered the icumazole BGC and proposed a biosynthetic pathway of icumazoles, which includes a PKS/NRPS hybrid assembly line and a CCR-based route to produce the precursor MBM-ACP. Further, *in silico* analysis of a CCR from this BGC, IcuL, unveiled a noncanonical class of CCRs, and it was found to exhibit significantly stricter substrate specificity, distinguishing it from biochemically earlier characterized CCRs. One member from the IcuL-like CCR family, TgaD, stands out due to its unique ability to utilize NADH in contrast to the strict nicotinamide adenine dinucleotide phosphate (NADPH) specificity of previously biochemically characterized CCRs.¹⁸

RESULT AND DISCUSSION

Identification of the icumazole BGC

Three *Sorangium cellulosum* strains from our in-house myxobacterial strain collection exhibited icumazole production capability, including two strains with available

genome data (complete genome sequence So ce836 and a draft genome for So ce377; Figure 1B). The BGC prediction in the So ce836 genome (Table S1) led to the identification of only one large modular PKS/NRPS BGC (BGC19) that has one analog in the genome of So ce377 (Table S2).

More hints toward the correct identification of the candidate BGC came from the correlation of the gene functions with the structural features of icumazoles. First, the isochromanone ring should be assembled by a PKS and a unique thioesterase (TE) domain based on the reported biosynthesis of ajudazol.²³ Second, we assumed the 2-methylbutyl side chain of the isochromanone to derive from MBM-CoA that was formed by a CCR as found encoded in the BGC. Third, the presence of an oxazole moiety in icumazoles indicates the requirement of an NRPS module featuring an adenylation (A) domain that activates Ser, a heterocyclization (HC) domain for the oxazoline ring formation, and an oxidation (OX) domain to eventually form the oxazole.²³ Furthermore, the oxazole module is supposed to be the only NRPS module, and three PKS modules should precede the oxazole module according to the so-called colinearity rule.¹² All these features are consistent with the gene function analysis of BGC19, consisting of 14 PKS modules and 1 NRPS module encoded by 10 PKS/NRPS genes and 1 noncanonical CCR (Table S2). Moreover, a transcriptome study of So ce836 showed that *icuA–L* were transcribed coincidentally with the production of icumazoles and also suggested the boundary of the *icu* BGC.²⁴ Consequently, BGC19 was considered the icumazole BGC (designated as *icu*).

Biosynthetic pathway of icumazoles

The *icu* BGC comprises more PKS modules than needed for building the carbon skeleton, indicating that the additional PKS modules are inactive or responsible for some noncanonical part of the carbon backbone (Figure 2A; Table S2). To determine the potential function of each domain, they were analyzed for the known critical residues (Data S1), including those for acyltransferase (AT), ketosynthase (KS), ketoreductase (KR), dehydratase (DH), enoylreductase (ER), ACP, and TE. The shorter AT₈ and AT₁₄ lack conserved catalytic residues, indicating their possible disability of loading building blocks to their cognate ACPs. KS₁₂ has a C-to-F point mutation in the catalytic site, similar to the C-to-Q mutation in KS_Q,²⁵ suggesting loss of condensation function. We also observed potential nonfunctional optional domains, including DH₁₀, DH₁₁, DH₁₄, and ER₁₃, which do not affect the PKS chain extension by the corresponding modules.

The icumazole megasynthetase adopts a unique architecture as found in other myxobacterial PKS, where an additional AT domain (AT_L) is embedded in the first extending module, forming a mixed module exhibiting ACP_L-KS₁-AT_L-AT₁-KR₁-ACP₁ organization (ML + M1; Figure 2A). We proposed that AT_L activated and transferred the starter unit onto ACP_L, in line with the reported biosynthesis of 5,²² ajudazol,²³ and stigmatellin.²⁶ However, unlike in other cases, AT_L from the *icu* pathway was predicted to recognize methylmalonyl-CoA as substrate (Table S2), which is supported by the critical R residue for selecting the dicarboxylic units in AT_L (Data S1A).^{27–29} KS₁ was proposed to bifunctionally catalyze the KS_Q-like decarboxylation in the loading step and the condensation in the first extension step. KR₁ was a B-type KR that generates a D-configuration hydroxyl group according to the sequence alignment (Data S1C), which is in agreement with the R configuration at C24 in icumazoles as confirmed by total synthesis (Figure 1A).¹¹ The second extending module (M2) lacks a KR domain, which is inconsistent with the corresponding hydroxyl group at C22 of icumazoles. We hypothesized that the ketoreduction in M2 may be catalyzed in *trans* by KR₁, while a similar case was observed in the azalomycin biosynthesis where a *cis*

Figure 2. Proposed biosynthetic pathway of icumazoles

(A) Scheme of the PKS/NRPS assembly line of icumazoles. Functional domains of PKS and NRPS are shown in red and orange, respectively; the optional domains are shown in blue; proposed inactive domains are shown in gray. Blue arrows indicate possible cross-modular catalyses. For details on the abbreviations, see the main text. For the sequence alignment analysis of each domain, see [Data S1](#). (B) Proposed pathway of the precursor 2-methylbutyrylmalonyl-ACP in the icumazole biosynthesis.

ER in *trans* is assumed to catalyze the enoyl reduction in a neighboring module.³⁰ Module 3 consists of an HC domain, an A domain, an OX domain, and a PCP domain (Figure 2A), forming a typical oxazole or thiazole biosynthetic module as observed in ajudazol²³ and epothilone³¹ biosyntheses. The following modules, M4 to M6, are assumed to further extend the PKS chain, giving rise to the triene moiety in icumazoles. However, these three modules only have two active DHs (Table S2; Data S1D), DH₄ and DH₆, implying that one of them might act twice during the biosynthesis. The dehydration in M5 that lacks a DH domain might be cross-modularly catalyzed by DH₆ because the double bonds supposed to be formed by M5 and M6 are both in Z configuration. The presence of KR₇ and MT₇ in module 7 is in line with the O-methyl group at C11 in icumazoles.

Subsequently, the *icu* biosynthesis enters the formation of the isochromanone ring resembling the TE-mediated cyclization mechanism in ajudazol biosynthesis.²³ Unlike the terminal TE from the ajudazol pathway, the icumazole TE is situated between PKS modules 11 and 12 in *IcuJ*. A PKS or NRPS with a nonterminal TE is highly unusual.³² Moreover, the actual function of this kind of TEs is still unclear except for the TE of Fr9C catalyzing the formation of a *cis*-double bond in FR901464 biosynthesis³³ and the TE-B enabling chain release in the malleicyprol pathway.³⁴ The TE domain in *IcuJ* possesses the conserved catalytic triad (Data S1G), suggesting its role in releasing **1**. This function was further supported by phylogenetic analysis, in which the icumazole TE clustered with the ajudazol TE (Data S1H). Intriguingly, modules M8 to M11 located before the TE seem almost collinear with the PKS chain extension required before isochromanone formation. Although AT₈ seems inactive, M8 may still be functional, aided by an adjacent AT acting in *trans*. An alternative would be that M8 is dysfunctional, whereas M9 works iteratively in this pathway. M10 is likely to elongate the nascent PKS backbone, utilizing the unique MBM building block (see below). Module 11 finalizes chain extension by using a methylmalonyl-CoA extender prior to chain release/ring formation by the TE (Figure 2A). The inactive DH₁₀ and DH₁₁ (Data S1D) indicate the involvement of a DH acting in *trans* in M10. It is thus proposed that *cis*-domains act in *trans* multiple times in icumazole biosynthesis. However, the exact mechanism is yet to be characterized. After the chain release and cyclization, remote unspecific glycosyltransferases are expected to mediate the glycosylation of **1** because no glycosyltransferase gene is found nearby the *icu* BGC (Figure 1B).

An unusual modular PKS-dependent biosynthesis of the MBM moiety

The activated MBM is a scarce PKS building block that was only reported once in the biosynthetic pathway of polyoxypeptin A.³⁵ A feeding study suggested that the MBM moiety in polyoxypeptin A is derived from isoleucine. A follow-up *in silico* analysis of the polyoxypeptin A BGC proposed a biosynthetic pathway to MBM, starting from isoleucine to form the corresponding α -keto acid.³⁶ The α -keto acid was thought to be extended by a FabH-like KS while tethered to a discrete ACP. Genes encoding α -keto acid dehydrogenases were also proposed to be involved in the following steps for producing the (*S,E*)-4-methyl-2-hexenoyl-CoA (MHE-CoA) as the substrate for the corresponding CCR. However, no genes encoding

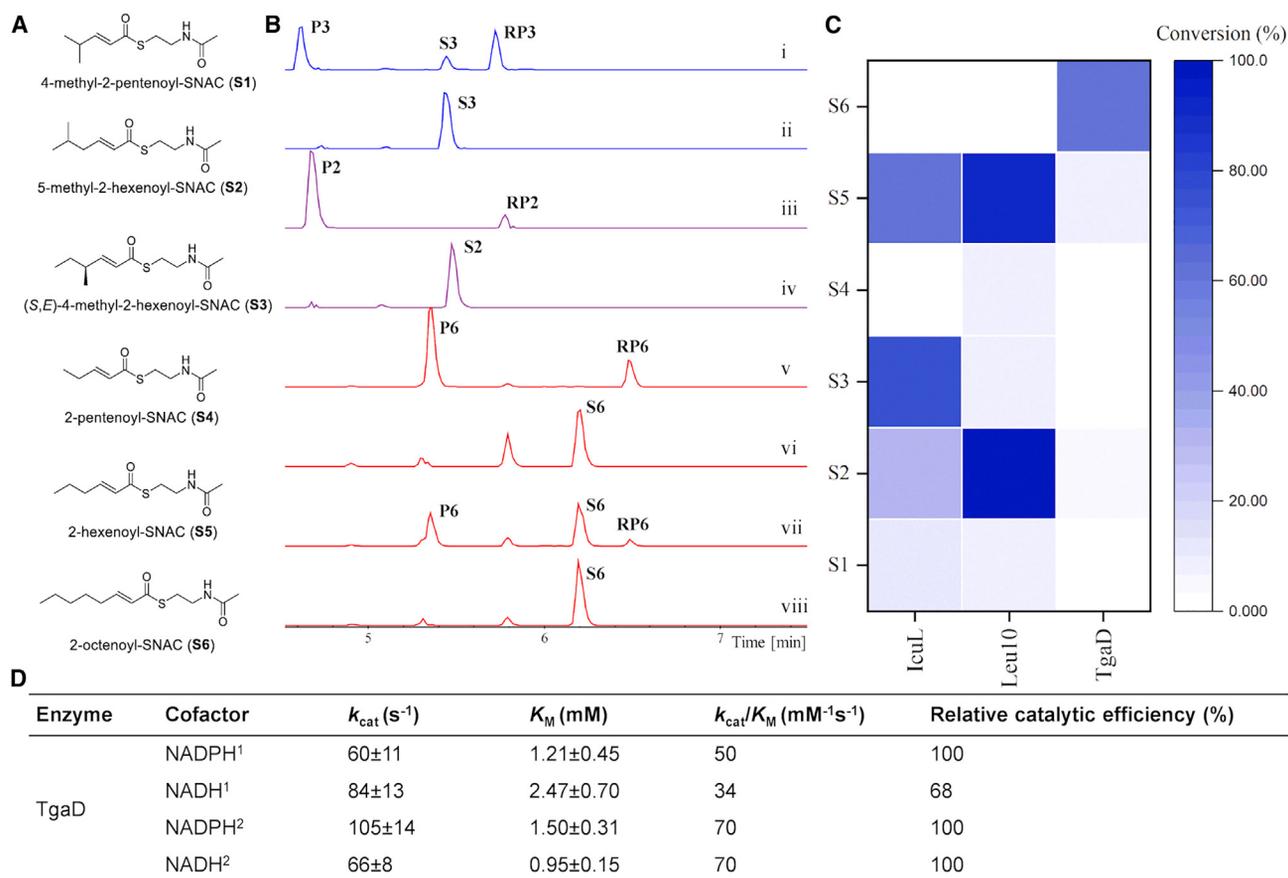
such enzymes could be discovered nearby the *icu* BGC except for the proposed CCR-encoding gene *icuC*. Moreover, these genes do not form an operon located elsewhere in the *So ce836* genome. Thus, all available evidence suggests a distinctive biosynthetic route to form MBM for icumazole biosynthesis.

As discussed above, three PKS modules of unassigned function follow the unusual nonterminal TE domain in *IcuJK*. Modules 12 and 13 were found in *IcuJ*, with M11 at the N terminus and the TE domain located before M12. *IcuK* comprising M14 is a single-module protein. We hypothesized that ACP₁₂ was self-malonylated and decarboxylated by the KS_Q-like KS₁₂. M13 next extends the acetyl-S-ACP species with one methylmalonate and reduces the β -keto to form a 2-methylbutyryl moiety attached to ACP₁₃. The nascent chain is further elongated by M14 with one malonyl-CoA (Figure 2B), and subsequently, KR₁₄ and DH₁₄ reduce the β -keto species to form an MHE moiety tethered to ACP₁₄, exactly representing the required substrate for a CCR reaction forming the extender unit required for icumazole biosynthesis, albeit in the ACP-bound form instead of the common CoA substrate for CCRs. The ACP-tethered CCR substrate is rare, but alkylmalonyl-ACPs were previously proposed in pathways of FK506³⁷ and ansaseomycin.³⁸ According to the sequence alignment, ER₁₃, AT₁₄, and DH₁₄ were predicted to be inactive due to the loss or mutation of key residues (Data S1). However, myxobacterial natural products' PKS and NRPS biosyntheses often deviate from textbook rules,³⁹ and in *trans*-acting domains could supplement the missing function as proposed for other parts of the icumazole biosynthesis.

IcuL represents a unique group of CCRs

The only non-PKS/-NRPS gene in *icu* BGC is *icuC*, and its encoded protein shows high sequence identity to TgaD from the 5 pathway (48.6%) and Leu10 from the 4 pathway (52.6%). TgaD and Leu10 were proposed to be CCRs, albeit exhibiting extremely low similarity to the canonical CCRs.^{21,22} Similarly, *IcuL* only exhibits 18% sequence identity to CinF,¹⁸ a well-characterized actinobacterial CCR. Also, the ~380 aa average length of *IcuL*-type CCR is roughly 70 aa shorter than CinF-type CCR, which is ~450 aa. Most importantly, the actual function of *IcuL*, TgaD, and Leu10 lacks experimental evidence.^{13,21,22} We here propose that *IcuL* is involved in icumazole biosynthesis by providing MBM-ACP from MHE-ACP by reductive carboxylation (Figure 2B). To verify this hypothesis and to investigate the characteristics of this new type of CCRs, we overexpressed and purified *IcuL*, TgaD, and Leu10 as recombinant proteins. Six *N*-acetylcysteamine (SNAC) thioester-mimicking substrates with different chain lengths, namely 4-methyl-2-pentenoyl-SNAC (S1), 5-methyl-2-hexenoyl-SNAC (S2), (*S,E*)-4-methyl-2-hexenoyl-SNAC (S3), 2-pentenoyl-SNAC (S4), 2-hexenoyl-SNAC (S5), and 2-octenoyl-SNAC (S6), were chemically synthesized (Figure 3A; Data S2) to characterize these enzymes *in vitro*.

We employed *icuC* genes from different strains, but only *IcuL* from *So ce377* was solubly expressed and purified with an N-terminal MBP tag. After an overnight incubation of *IcuL* with reduced NADPH, NaHCO₃ and SNAC-ester S3, two new products with *m/z* values of 232.14 and 276.13 were formed (Figures 3B, 3Bi, and 3Bii), which are 2 and 46 daltons heavier than the *m/z* value of S3 (230.12), respectively, suggesting the formation of a reduced product 4-methylhexanoyl-SNAC (RP3) and a reductive carboxylated product MBM-SNAC (P3). The same types of products were also detected in assays of Leu10 (Figures 3B, 3Biii, and 3Biv) and TgaD (Figures 3B, 3Bv, and 3Bvi), corroborating their function as CCRs.



¹, testing with 50 mM NaHCO₃;
², testing without NaHCO₃.

Figure 3. Biochemical studies of IcuL-like CCRs

(A) Structure of the synthetic substrates.

(B) *In vitro* assays of IcuL (i and ii), Leu10 (iii and iv), and TgaD (v to viii). All assays were incubated at 30°C overnight. EICs of substrates and the corresponding products are shown in the left panel. Illustration of enzymatic reaction of each compound is shown in the right panel. Traces i, iii, v, and vii show reactions containing enzymes, cofactors, NaHCO₃, and corresponding substrates; traces ii, iv, vi, and viii are control reactions without enzymes. Traces i to vi use cofactor NADPH, while traces vii and viii use NADH. EICs of i to iv = 230.12, 232.14, and 276.13, respectively, and EICs of v to viii = 244.14, 246.15, and 290.14, respectively. P2, reductive carboxylated product of S2 (276.13); RP2, reduced product of S2 (232.14); P3, reductive carboxylated product of S3 (276.13); RP3, reduced product of S3 (232.14); P6, reductive carboxylated product of S6 (290.14); RP6, reduced product of S6 (246.15). See also Figure S1.

(C) Evaluation of substrate specificity of the three CCRs to all six SNAC substrates. For detailed data, see also Table S3. All assays were incubated at 30°C overnight in the presence of 50 mM NaHCO₃ in Tris-HCl (pH 7) buffer.

(D) Enzyme kinetic analysis of TgaD toward NADH/NADPH. See also Figures S2 and S3.

We then evaluated the potential substrate specificity of these three CCRs using all six alkenoyl-SNACs. IcuL catalyzed the reductive carboxylation of S2 and S5 at a moderate level compared with S3 with the authentic chain length, while almost no conversion was observed when using S1, S4, and S6 as substrates (Figures 3C and S1A). Except for the authentic SNAC S2, Leu10 is also actively employing S5 while it was found to be almost inactive on other SNAC esters (Figures 3C and S1B). We examined the structures of these substrates and found that S2, S3, and S5 share the same carbon chain length (Figure 3A), presumably suggesting a crucial role of chain length in the substrate specificity of this new type of CCR. In line with this assumption, none of the remaining five substrates, at least two carbons shorter than S6, can serve as substrate for TgaD (Figure S1C). Although SNAC esters might not perfectly mimic acyl-CoA substrates, these observations align with the fact that

the IcuL-like CCRs only introduce one kind of side chain in icumazoles, thuggacins, and leupyrrins, respectively. This property is divergent from CinF from the cinnabaramide BGC¹⁸ and PteB from the filipin BGC,⁴⁰ which are comparably much more promiscuous toward substrates with various carbon chain lengths.

TgaD utilizes both NADH and NADPH as cofactors

CCRs are recognized to be NADPH-dependent enzymes.¹³ Quade et al. determined that CinF strictly uses NADPH as its cofactor based on the *in vitro* biochemical characterization and cocrystal structures,¹⁸ presumably implying the general feature of tight cofactor specificity in canonical CCRs. We asked the question if members of the newly identified CCR family are capable of leveraging NADH as a cofactor. As expected, the catalytic activities of both IcuL and Leu10 were abolished when employing NADH as the cofactor (Figure S1D). However, octanoyl-SNAC (RP6) and hexylmalonyl-SNAC (P6) were successfully produced from S6 by TgaD when using NADH instead of NADPH (Figures 3B, 3Bvii, and 3Bviii).

To assess the specificity of TgaD toward NADH and NADPH, we decided to measure the kinetic parameters of TgaD. However, all six alkenoyl-SNACs degraded quickly (Figures S1A–S1C, S2A, and S2B), affecting an accurate kinetic study. In order to identify the reason for the degradation, we incubated S6 with each component from the reaction separately. The degradation only appeared when Tris-HCl (pH 8) and NaHCO₃ (Figure S2C) were used, suggesting that the relatively high pH should be the leading cause of the degradation. Therefore, a suitable condition for the CCR kinetic assay employing the SNAC substrates was required. Considering the pH of NaHCO₃ (8.3), we first tested the TgaD assay in various buffers without bicarbonate (Figure S2D), revealing that the degradation of S6 slowed down when using buffers with pH lower than 7. Further experiments showed that adding NaHCO₃ promotes degradation intensely even at a concentration of 25 mM when using a pH of 7. The degradation took place similarly when using NH₄HCO₃ at pH 7.8. Buffers with a pH lower than 7 are considered inappropriate for TgaD catalysis due to the low velocity. Notably, CCRs convert substrates via two independent steps, including forming a covalent NADPH-alkenoyl-CoA ene adduct intermediate in the first step, exhibiting maximum absorption at 260 and 370 nm.⁴¹ Therefore, NAD(P)H consumption can never be measured accurately if the intermediate is not immediately converted to the final product by the second reaction. Comparing k_{cat} and K_M constants at pH 7, we found that TgaD processes cofactor NADH with relative catalytic efficiency of 68% to NADPH in the presence of 50 mM NaHCO₃ (Figures 3D and S3A). TgaD has a higher K_M value when using NADH, suggesting a higher binding affinity for NADPH than NADH. Unexpectedly, the k_{cat} value of TgaD using NADH is 1.4-fold higher than for NADPH, indicating a higher catalytic velocity of NADH compared with NADPH. We also measured the enzyme kinetics in the absence of NaHCO₃, revealing different parameters, most probably due to the unstable ene adduct (Figures 3D and S3B). Nevertheless, TgaD generally showed a similar level of preference for both NADH and NADPH. To the best of our knowledge, TgaD is the first CCR reported to utilize both NADH and NADPH.

CCRs play a vital role in highly energy-consuming natural product biosynthesis by supplying special building blocks.¹³ The NADH-utilizing property of CCR provides an opportunity to engineer the metabolic flux for production improvement since the generation of NADP(H) is performed by NAD kinase from NAD(H) in bacteria,⁴² involving the consumption of ATP. A recent study reported establishing a CO₂ fixation synthetic cycle where two CCRs play a core role in the pathway.⁴³ Our finding thus enables future efforts on alteration of the cofactor preference of these CCRs,

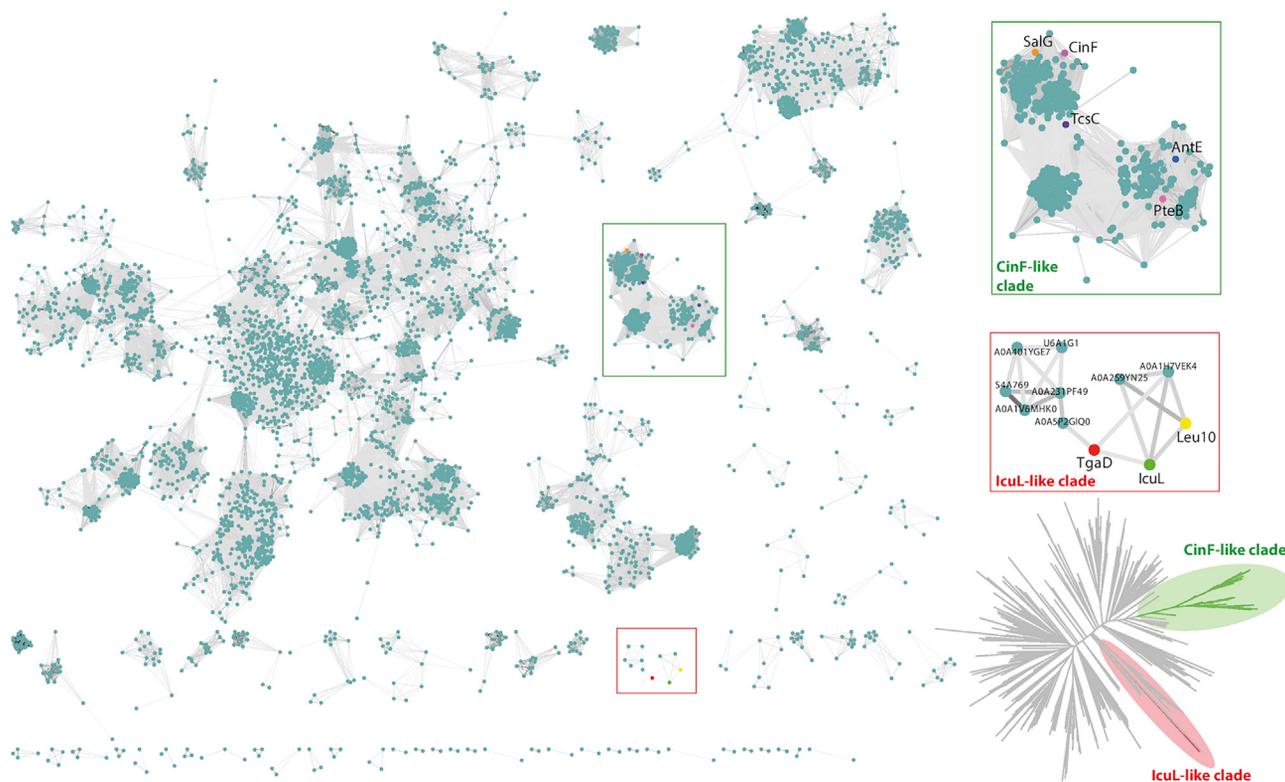


Figure 4. Global distribution of ADH family proteins located within BGCs

The canonical CinF-like CCR group (green frame) and the new IcuL-like CCR group (red frame) are highlighted. The edges are drawn between nodes, representing individual ADHs (darker edges indicate more significant sequence identity, edges drawn starting from >25% sequence identity values). On the top right, CCRs from well-characterized BGCs are colored and labeled. Phylogenetic tree with these groups highlighted correspondingly is shown at the bottom right.

which may result in more efficient *in vivo* pathways by saving two molecules of ATP for each catalytic cycle.

Distribution of IcuL-/TgaD-like CCRs in nature

To systematically study the new IcuL-like CCR family, we conducted a global sequence-based survey from the UniProt database⁴⁴ as described in the [supplemental experimental procedures](#) (Figure 4). As expected, IcuL, Leu10, and TgaD cluster in a unique group with eight other proteins (Table S4) apart from the canonical CCRs (a cluster of 586 proteins, including CinF,¹⁸ AntE,⁴⁵ SalG,⁴⁶ TcsC,⁴⁷ and PteB⁴⁰; see insets in Figure 4), suggesting their potential functional difference. Surprisingly, we also found that AsmB6 (UniProt: A0A5P2GIQ0), which was proposed as the CCR for ansaseomycin biosynthesis, was grouped into the new IcuL-like CCR family.³⁸ Moreover, the CCR from polyoxypeptin A³⁶ (identity/similarity of 41.6%/58.6% to IcuL) should also belong to this IcuL-like family but was absent because the corresponding genome data were missing in the public databases. Closer scrutiny of the structures of icumazoles, leupyrrins, thuggacins, ansaseomycins, and polyoxypeptins, which are all known compounds using the new family CCR in biosynthesis, revealed that only one kind of side chain was found,^{11,38,48–51} suggesting a precise substrate specificity of the IcuL-like CCRs controlling the chain length of building blocks for the polyketide chain extension.

We performed a phylogenetic analysis confirming that IcuL-like enzymes form a separate subgroup within the larger ADH family (Figure 4, bottom right). These

enzymes form a clade (in red) that branches very deeply in the phylogenetic tree of the ADH family and is separated from canonical CCRs (green clade). Amino acid conservation also supports a distinct mode of cofactor binding by the IcuL-like enzymes (Figure S4): the NADP-/NAD-binding pocket is significantly different in this group compared with canonical CCRs (CinF-like clade), as well as from other NADP- and NAD-binding ADHs. In particular, the amino acids that correspond to Ser253 and Lys257 in the pocket discriminating between NADP and NAD¹⁸ are mutated but not toward the consensus observed for NAD-binding ADHs. On the other hand, the CoA-binding pocket is largely conserved, but some consensus amino acids are different from the CinF-like clade (e.g., Trp80 is mutated to a His).

Comparing three-dimensional structures of CinF (experimentally resolved¹⁸) and TgaD (modeled with AlphaFold⁵²) confirms the overall similarity and important differences between the IcuL- and CinF-like enzymes (Figure S5). Considering that the TgaD model was produced in the absence of cofactors, and hence a certain degree of induced fit should be expected, one can observe an overall structural conservation of both NADP-/NAD- and CoA-binding pockets. However, the key cofactor-coordinating amino acids are frequently mutated, sometimes to amino acids with strikingly different chemical properties. For example, His218 (TgaD) might play a role in NADP/NAD selectivity; it is mutated to an Asn in Leu10 and an Arg in IcuL and the other IcuL-like proteins. Detailed mechanistic studies are thus required to fully elucidate the biochemistry of this new type of CCR.

Significance

The myxobacterial polyketides icumazoles exhibit potent antifungal bioactivity. However, the biosynthetic pathway of icumazoles remained unknown, leaving an especially intriguing open question of how the 2-methylbutyryl side chain of the isochromanone ring is introduced. Here, we describe the BGC of icumazoles, a hybrid pathway consisting of ten PKS genes, one NRPS gene, and one CCR gene (*icuL*). It is worth noting that the PKS/NRPS assembly line is divided into two parts by an unusual nonterminal TE domain. The former part comprises a loading module and 11 extension modules, which are proposed to build the backbone of icumazoles while requiring a highly unusual MBM building block. The latter PKS modules might be responsible for forming a rare precursor MHE-ACP, which deviates from the previously proposed MHE pathway. MHE-ACP is further converted to MBM-ACP by IcuL as a PKS building block for icumazole biosynthesis. We studied IcuL and its homologs TgaD and Leu10 *in vitro*, revealing a stricter substrate specificity of the IcuL-like CCRs compared with the canonical ones. Most intriguingly, we showed that TgaD unprecedentedly utilizes both NADPH and NADH as cofactors with similar efficiency, diverging from the NADPH-specific characteristic of canonical CCRs. Furthermore, a global sequence survey unambiguously shows that the IcuL-like CCR cluster is a new group separated from the canonical CinF-like CCRs. Further in-depth studies of the IcuL-like CCR, especially TgaD, are promising for synthetic biology applications because of their strict substrate specificity and the distinctive NADH-utilizing characteristic of TgaD, which will contribute to synthetic carbon fixation and production of high-value-added chemical products.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Prof. Rolf Müller (rolf.mueller@helmholtz-hips.de).

Materials availability

Plasmids generated in this study are available from the [lead contact](#) without restriction.

The bacterial strains, chemical reagents, oligonucleotides, and recombinant plasmids used in this work are listed in [Data S6](#).

Data and code availability

The nuclear magnetic resonance (NMR) spectra of synthetic substrates and their intermediates are listed in [Data S2](#). The DNA sequences of codon-optimized *icuL* and *tgaD* are listed in [Data S3](#). The genome sequence of strain So ce836 has been previously deposited in the NCBI database with an accession number of GenBank: CP012672.1. The genes in BGC *icu* include IDs from SOCE836_041090 to SOCE836_041200. UniProt identifiers of proteins in CinF-like and IcuL-like groups are listed in [Data S4](#). Amino acid sequence alignment of IcuL-like CCRs and typical NADH- and NADPH-dependent ADHs using for producing [Figure S4](#) is given in [Data S5](#).

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.xcrp.2023.101394>.

ACKNOWLEDGMENTS

This work was supported by grants to R.M., C.F., and A.K.H.H. from the following sources: Deutsche Forschungsgemeinschaft and the Helmholtz International Lab (InterLabs0007). O.V.K. acknowledges support from the Klaus Faber Foundation.

AUTHOR CONTRIBUTIONS

F.X., C.F., and R.M. performed the gene cluster analysis. F.X. and C.F. performed and analyzed biochemical experiments. A.F.K. and A.K.H.H. synthesized the SNAC substrates. O.V.K. performed phylogenetic and bioinformatics structural analysis. O.V.K. and C.F. performed and interpreted the sequence similarity network analysis and interpreted the bioinformatics analysis results. F.X., C.F., and R.M. wrote the manuscript with input from all authors. C.F. and R.M. conceived and designed the study.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 5, 2022

Revised: March 9, 2023

Accepted: April 11, 2023

Published: May 1, 2023

REFERENCES

1. Newman, D.J., and Cragg, G.M. (2020). Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* 83, 770–803.
2. Herrmann, J., Fayad, A.A., and Müller, R. (2017). Natural products from myxobacteria: novel metabolites and bioactivities. *Nat. Prod. Rep.* 34, 135–160.
3. Weissman, K.J., and Müller, R. (2010). Myxobacterial secondary metabolites: bioactivities and modes-of-action. *Nat. Prod. Rep.* 27, 1276–1295.
4. Bollag, D.M., McQueney, P.A., Zhu, J., Hensens, O., Koupal, L., Liesch, J., Goetz, M., Lazarides, E., and Woods, C.M. (1995). Epothilones, a new class of microtubule-stabilizing agents with a taxol-like mechanism of action. *Cancer Res.* 55, 2325–2333.
5. Gerth, K., Bedorf, N., Höfle, G., Irschik, H., and Reichenbach, H. (1996). Epothilons A and B: antifungal and cytotoxic compounds from *Sorangium cellulosum* (Myxobacteria). Production, physico-chemical and biological properties. *J. Antibiot.* 49, 560–563.
6. Baumann, S., Herrmann, J., Raju, R., Steinmetz, H., Mohr, K.I., Hüttel, S., Harmrolfs, K., Stadler, M., and Müller, R. (2014). Cystobactamids: myxobacterial topoisomerase inhibitors

- exhibiting potent antibacterial activity. *Angew. Chem. Int. Ed. Engl.* **53**, 14605–14609.
7. Hüttel, S., Testolin, G., Herrmann, J., Planke, T., Gille, F., Moreno, M., Stadler, M., Brönstrup, M., Kirschning, A., and Müller, R. (2017). Discovery and total synthesis of natural cystobactamid derivatives with superior activity against Gram-negative pathogens. *Angew. Chem. Int. Ed. Engl.* **56**, 12760–12764.
 8. Plaza, A., Garcia, R., Bifulco, G., Martinez, J.P., Hüttel, S., Sasse, F., Meyerhans, A., Stadler, M., and Müller, R. (2012). Aetheramides A and B, potent HIV-inhibitory depsipeptides from a myxobacterium of the new genus "*Aetherobacter*". *Org. Lett.* **14**, 2854–2857.
 9. Fu, C., Sikandar, A., Donner, J., Zaburannyi, N., Herrmann, J., Reck, M., Wagner-Döbler, I., Koehnke, J., and Müller, R. (2017). The natural product carolacton inhibits folate-dependent C1 metabolism by targeting Fold/MTHFD. *Nat. Commun.* **8**, 1529.
 10. Jansen, R., Irschik, H., Huch, V., Schummer, D., Steinmetz, H., Bock, M., Schmidt, T., Kirschning, A., and Müller, R. (2010). Carolacton - a macrolide ketocarboxylic acid that reduces biofilm formation by the Caries- and Endocarditis-associated bacterium *Streptococcus mutans*. *Eur. J. Org. Chem.* **2010**, 1284–1289.
 11. Barbier, J., Jansen, R., Irschik, H., Benson, S., Gerth, K., Böhlendorf, B., Höfle, G., Reichenbach, H., Wegner, J., Zeilinger, C., et al. (2012). Isolation and total synthesis of icumazoles and noricumazoles—antifungal antibiotics and cation-channel blockers from *Sorangium cellulosum*. *Angew. Chem. Int. Ed. Engl.* **51**, 1256–1260.
 12. Fischbach, M.A., and Walsh, C.T. (2006). Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev.* **106**, 3468–3496.
 13. Wilson, M.C., and Moore, B.S. (2012). Beyond ethylmalonyl-CoA: the functional role of crotonyl-CoA carboxylase/reductase homologs in expanding polyketide diversity. *Nat. Prod. Rep.* **29**, 72–86.
 14. Liu, H., and Reynolds, K.A. (1999). Role of crotonyl coenzyme A reductase in determining the ratio of polyketides monensin A and monensin B produced by *Streptomyces cinnamonensis*. *J. Bacteriol.* **181**, 6806–6813.
 15. Erb, T.J., Berg, I.A., Brecht, V., Müller, M., Fuchs, G., and Alber, B.E. (2007). Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. *Proc. Natl. Acad. Sci. USA* **104**, 10631–10636.
 16. Erb, T.J., Brecht, V., Fuchs, G., Müller, M., and Alber, B.E. (2009). Carboxylation mechanism and stereochemistry of crotonyl-CoA carboxylase/reductase, a carboxylating enoyl-thioester reductase. *Proc. Natl. Acad. Sci. USA* **106**, 8871–8876.
 17. Zhang, L., Mori, T., Zheng, Q., Awakawa, T., Yan, Y., Liu, W., and Abe, I. (2015). Rational control of polyketide extender units by structure-based engineering of a crotonyl-CoA carboxylase/reductase in antimycin biosynthesis. *Angew. Chem. Int. Ed. Engl.* **54**, 13462–13465.
 18. Quade, N., Huo, L., Rachid, S., Heinz, D.W., and Müller, R. (2011). Unusual carbon fixation gives rise to diverse polyketide extender units. *Nat. Chem. Biol.* **8**, 117–124.
 19. Ray, L., Valentic, T.R., Miyazawa, T., Withall, D.M., Song, L., Milligan, J.C., Osada, H., Takahashi, S., Tsai, S.-C., and Challis, G.L. (2016). A crotonyl-CoA reductase-carboxylase independent pathway for assembly of unusual alkylmalonyl-CoA polyketide synthase extender units. *Nat. Commun.* **7**, 13609.
 20. Pohl, N.L., Hans, M., Lee, H.Y., Kim, Y.S., Cane, D.E., and Khosla, C. (2001). Remarkably broad substrate tolerance of malonyl-CoA synthetase, an enzyme capable of intracellular synthesis of polyketide precursors. *J. Am. Chem. Soc.* **123**, 5822–5823.
 21. Kopp, M., Irschik, H., Gemperlein, K., Buntin, K., Meiser, P., Weissman, K.J., Bode, H.B., and Müller, R. (2011). Insights into the complex biosynthesis of the leupyrrins in *Sorangium cellulosum* So ce690. *Mol. Biosyst.* **7**, 1549–1563.
 22. Buntin, K., Irschik, H., Weissman, K.J., Luxenburger, E., Blöcker, H., and Müller, R. (2010). Biosynthesis of thuggacins in myxobacteria: comparative cluster analysis reveals basis for natural product structural diversity. *Chem. Biol.* **17**, 342–356.
 23. Buntin, K., Weissman, K.J., and Müller, R. (2010). An unusual thioesterase promotes isochromanone ring formation in ajudazol biosynthesis. *ChemBiochem* **11**, 1137–1146.
 24. Boldt, J., Lukoševičiūtė, L., Fu, C., Steglich, M., Bunk, B., Junker, V., Gollasch, A., Trunkwalter, B., Mohr, K.I., Beckstette, M., et al. (2023). Bursts in biosynthetic gene cluster transcription are accompanied by surges of natural compound production in the myxobacterium *Sorangium* sp. *Microb. Biotechnol.* **00**, 1–15.
 25. Chisuga, T., Nagai, A., Miyanaga, A., Goto, E., Kishikawa, K., Kudo, F., and Eguchi, T. (2022). Structural Insight into the reaction mechanism of ketosynthase-like decarboxylase in a loading module of modular polyketide synthases. *ACS Chem. Biol.* **17**, 198–206.
 26. Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blöcker, H., Höfle, G., and Müller, R. (2002). The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. *J. Biol. Chem.* **277**, 13082–13090.
 27. Liou, G.F., Lau, J., Cane, D.E., and Khosla, C. (2003). Quantitative analysis of loading and extender acyltransferases of modular polyketide synthases. *Biochemistry* **42**, 200–207.
 28. Long, P.F., Wilkinson, C.J., Bisang, C.P., Cortés, J., Dunster, N., Olynyk, M., McCormick, E., McArthur, H., Mendez, C., Salas, J.A., et al. (2002). Engineering specificity of starter unit selection by the erythromycin-producing polyketide synthase. *Mol. Microbiol.* **43**, 1215–1225.
 29. Rangan, V.S., and Smith, S. (1997). Alteration of the substrate specificity of the malonyl-CoA/acetyl-CoA:acyl carrier protein S-acyltransferase domain of the multifunctional fatty acid synthase by mutation of a single arginine residue. *J. Biol. Chem.* **272**, 11975–11978.
 30. Zhai, G., Wang, W., Xu, W., Sun, G., Hu, C., Wu, X., Cong, Z., Deng, L., Shi, Y., Leadlay, P.F., et al. (2020). Cross-module enoylreduction in the azalomycin F polyketide synthase. *Angew. Chem. Int. Ed. Engl.* **59**, 22738–22742.
 31. Dowling, D.P., Kung, Y., Croft, A.K., Taghizadeh, K., Kelly, W.L., Walsh, C.T., and Drennan, C.L. (2016). Structural elements of an NRPS cyclization domain and its intermodule docking domain. *Proc. Natl. Acad. Sci. USA* **113**, 12432–12437.
 32. Little, R.F., and Hertweck, C. (2022). Chain release mechanisms in polyketide and non-ribosomal peptide biosynthesis. *Nat. Prod. Rep.* **39**, 163–205.
 33. He, H.-Y., Tang, M.-C., Zhang, F., and Tang, G.-L. (2014). Cis-double bond formation by thioesterase and transfer by ketosynthase in FR901464 biosynthesis. *J. Am. Chem. Soc.* **136**, 4488–4491.
 34. Little, R., Trottmann, F., Preissler, M., and Hertweck, C. (2022). An intramodular thioesterase domain catalyses chain release in the biosynthesis of a cytotoxic virulence factor. *RSC Chem. Biol.* **3**, 1121–1128.
 35. Umezawa, K., Ikeda, Y., Kawase, O., Naganawa, H., and Kondo, S. (2001). Biosynthesis of polyoxypeptin A: novel amino acid 3-hydroxy-3-methylproline derived from isoleucine. *J. Chem. Soc. Perkin 1*, 1550–1553.
 36. Du, Y., Wang, Y., Huang, T., Tao, M., Deng, Z., and Lin, S. (2014). Identification and characterization of the biosynthetic gene cluster of polyoxypeptin A, a potent apoptosis inducer. *BMC Microbiol.* **14**, 30.
 37. Mo, S., Kim, D.H., Lee, J.H., Park, J.W., Basnet, D.B., Ban, Y.H., Yoo, Y.J., Chen, S.w., Park, S.R., Choi, E.A., et al. (2011). Biosynthesis of the allylmalonyl-CoA extender unit for the FK506 polyketide synthase proceeds through a dedicated polyketide synthase and facilitates the mutasynthesis of analogues. *J. Am. Chem. Soc.* **133**, 976–985.
 38. Liu, S.H., Wang, W., Wang, K.B., Zhang, B., Li, W., Shi, J., Jiao, R.H., Tan, R.X., and Ge, H.M. (2019). Heterologous expression of a cryptic giant type I PKS gene cluster leads to the production of ansaseomycin. *Org. Lett.* **21**, 3785–3788.
 39. Keatinge-Clay, A.T. (2017). The uncommon enzymology of cis-Acyltransferase assembly lines. *Chem. Rev.* **117**, 5334–5366.
 40. Yoo, H.-G., Kwon, S.-Y., Kim, S., Karki, S., Park, Z.-Y., and Kwon, H.-J. (2011). Characterization of 2-octenoyl-CoA carboxylase/reductase utilizing pteB from *Streptomyces avermitilis*. *Biosci. Biotechnol. Biochem.* **75**, 1191–1193.
 41. Rosenthal, R.G., Ebert, M.-O., Kiefer, P., Peter, D.M., Vorholt, J.A., and Erb, T.J. (2014). Direct evidence for a covalent ene adduct intermediate in NAD(P)H-dependent enzymes. *Nat. Chem. Biol.* **10**, 50–55.
 42. Spaans, S.K., Weusthuis, R.A., van der Oost, J., and Kengen, S.W.M. (2015). NADPH-generating

- systems in bacteria and archaea. *Front. Microbiol.* **6**, 742.
43. Schwander, T., Schada von Borzyskowski, L., Burgener, S., Cortina, N.S., and Erb, T.J. (2016). A synthetic pathway for the fixation of carbon dioxide *in vitro*. *Science* **354**, 900–904.
44. UniProt Consortium (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489.
45. Yan, Y., Zhang, L., Ito, T., Qu, X., Asakawa, Y., Awakawa, T., Abe, I., and Liu, W. (2012). Biosynthetic pathway for high structural diversity of a common dilactone core in antimycin production. *Org. Lett.* **14**, 4142–4145.
46. Eustáquio, A.S., McGlinchey, R.P., Liu, Y., Hazzard, C., Beer, L.L., Florova, G., Alhamadsheh, M.M., Lechner, A., Kale, A.J., Kobayashi, Y., et al. (2009). Biosynthesis of the salinosporamide A polyketide synthase substrate chloroethylmalonyl-coenzyme A from S-adenosyl-L-methionine. *Proc. Natl. Acad. Sci. USA* **106**, 12295–12300.
47. Blažič, M., Kosec, G., Baebler, Š., Gruden, K., and Petković, H. (2015). Roles of the crotonyl-CoA carboxylase/reductase homologues in acetate assimilation and biosynthesis of immunosuppressant FK506 in *Streptomyces tsukubaensis*. *Microb. Cell Fact.* **14**, 164.
48. Irschik, H., Reichenbach, H., Höfle, G., and Jansen, R. (2007). The thuggacins, novel antibacterial macrolides from *Sorangium cellulosum* acting against selected Gram-positive bacteria. *J. Antibiot.* **60**, 733–738.
49. Bode, H.B., Irschik, H., Wenzel, S.C., Reichenbach, H., Müller, R., and Höfle, G. (2003). The leupyrrins: a structurally unique family of secondary metabolites from the myxobacterium *Sorangium cellulosum*. *J. Nat. Prod.* **66**, 1203–1206.
50. Steinmetz, H., Irschik, H., Kunze, B., Reichenbach, H., Höfle, G., and Jansen, R. (2007). Thuggacins, macrolide antibiotics active against *Mycobacterium tuberculosis*: isolation from myxobacteria, structure elucidation, conformation analysis and biosynthesis. *Chemistry* **13**, 5822–5832.
51. Umezawa, K., Nakazawa, K., Uemura, T., Ikeda, Y., Kondo, S., Naganawa, H., Kinoshita, N., Hashizume, H., Hamada, M., Takeuchi, T., and Ohba, S. (1998). Polyoxypeptin isolated from *Streptomyces*: a bioactive cyclic depsipeptide containing the novel amino acid 3-hydroxy-3-methylproline. *Tetrahedron Lett.* **39**, 1389–1392.
52. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.