A role for antibiotic biosynthesis monooxygenase domain proteins in fidelity control during aromatic polyketide biosynthesis

Zhiwei Qin\textsuperscript{1,3}, Rebecca Devine\textsuperscript{2,3}, Matthew I. Hutchings\textsuperscript{2} & Barrie Wilkinson\textsuperscript{1}

The formicamycin biosynthetic gene cluster encodes two groups of type 2 polyketide antibiotics: the formicamycins and their biosynthetic precursors the fasamycins, both of which have activity against methicillin-resistant \textit{Staphylococcus aureus}. Here, we report the formicapyridines which are encoded by the same gene cluster and are structurally and biosynthetically related to the fasamycins and formicamycins but comprise a rare pyridine moiety. These compounds are trace-level metabolites formed by derailment of the major biosynthetic pathway. Inspired by evolutionary logic we show that rational mutation of a single gene in the biosynthetic gene cluster encoding an antibiotic biosynthesis monooxygenase (ABM) superfamily protein leads to a significant increase both in total formicapyridine production and their enrichment relative to the fasamycins/formicamycins. Our observations broaden the polyketide biosynthetic landscape and identify a non-catalytic role for ABM superfamily proteins in type II polyketide synthase assemblages for maintaining biosynthetic pathway fidelity.
A n idealised linear biosynthetic pathway to a complex natural product can be imagined proceeding through a series of intermediate structures which would exist for some finite time as the pathway product accumulates. These hypothetical intermediates would exist freely in solution or bound to enzymes in the case of assembly line processes, and eventually all flux through the pathway would end and only the final product would exist. In reality no pathways proceed in this manner as the situation is complicated by varying rates of reaction for the different steps, meaning that some intermediates accumulate, while the inherent reactivity of other intermediates, or their ability to act as substrates for housekeeping enzymes not dedicated to the pathway, means that shunt metabolites often arise. Matters are further complicated by the fact that pathways are often convergent, with multiple units made in parallel before assembly into the final product, for example in the biosynthetic pathways to macrolide or aminoglycoside antibiotics. In addition, some pathways are not linear, and the final product is accessed via several routes due to the inherent substrate plasticity of the biosynthetic enzymes; well-studied examples include the rapamycin and erythromycin pathways. Thus, in practice, any biosynthetic pathway will lead to the accumulation of a mixture comprising the final product plus varying concentrations of pathway intermediates and shunt metabolites. The composition of such a mixture will vary further when alternate growth conditions are used. Sometimes the final product of the pathway cannot be clearly discerned.

Such mixtures of compounds are said to comprise a series of biosynthetic congeners, and their identification can provide valuable information about a biosynthetic pathway and may sometimes lead to new biosynthetic understanding. We recently reported the formicamycins, aromatic polyketides produced by Streptomyces formicace KY5 isolated from the fungus-growing plant-ant Tetraponera penzigi (Fig. 1). In total, sixteen congeners were isolated including three more fasamycins, a group of compounds previously reported from the heterologous expression of a clone derived from environmental DNA. These congeners are the product of a type II polyketide synthase (PKS) operating in conjunction with post-PKS modifications that include O-methylation and halogenation, plus oxidative and reductive modifications. Intrigued by these compounds, which exhibit potent antibacterial activity, we employed targeted metabolomics to identify further congeners that may have been missed during manual analysis of culture extracts. This led us to identify the formicapyridines (1–9), pyridine containing polyketide alkaloids which represent additional products of the formicamycin (for) biosynthetic gene cluster (BGC). Products of type II PKS systems which contain a pyridine moiety are rare.

Challis and co-workers have shown that the majority of actinomycete derived polyketide alkaloids, including those containing a pyridine moiety, arise from reactive intermediates formed after transamination of aldehydes generated from reductive off-loading of the thioester bound polyketide chain from a type I modular PK. In contrast, the formicapyridines are minor shunt metabolites that arise due to derailment of the formicamycin biosynthetic pathway. We show here that reprogramming the for BGC, by deleting the forS gene encoding an ABM family protein, leads to significantly increased production of formicapyridines with concomitant reduction of the combined titre of fasamycins and formicamycins. These and other mutational data lead us to the hypothesis that ForS is not a cyclase but forms part of a multienzyme complex where it acts as a chaperone-like protein to aid in maintaining pathway fidelity and performance. The discovery and engineering of formicapyridine biosynthesis raises intriguing questions regarding the evolution of type II PKS biosynthetic pathways and the origins of natural product chemical diversity.

Results

Metabolomics led identification of the formicapyridines. The aromatic, polycyclic nature of these molecules limited fragmentation and the effectiveness of molecular networking via the Global Natural Product Social Molecular Networking web platform. Instead, as most fasamycin and formicamycin congeners are halogenated (Fig. 1), we established a bespoke dereplication method, making use of S. formicace KY5 mutants that we reported previously, including an entire BGC deletion strain (Δfor), and a strain in which the halogenase gene was deleted (ΔforV) (Fig. 2). Our earlier study showed that the Δfor mutant does not produce any fasamycins or formicamycins, and the ΔforV mutant produces only non-halogenated fasamycin congeners.

Replicate (n = 3) ethyl acetate extracts of the wild type (WT), Δfor and ΔforV strains, along with equivalent extracts from uninoculated mannitol soya flour (MS) agar plates, were analysed by liquid chromatography high-resolution mass spectrometry (LC-HRMS). Using the Profiling Solutions software (Shimadzu Corporation), the WT dataset (Supplementary Data 1 and 2) was filtered to remove ions also present in the other samples. Only two BGCs in the S. formicace KY5 genome contain putative halogenase encoding genes, so we hypothesised that any chlorine-containing ions present in the filtered dataset would likely derive from the formicamycin BGC. This process dramatically reduced the dataset complexity, leaving <200 unique ions from an original set of ~3000. Manual curation showed that most of the remaining ions corresponded to halogenated molecules (based on isotope patterns), leaving twelve previously identified fasamycin and formicamycin congeners, two additional fasamycin/formicamycin congeners that remain uncharacterized due to trace levels, and the known isoflavone 6-chlorogenistein (Supplementary Note 1 and Supplementary Figs. 11–14) along with a regioisomer. Analysis of C/H ratios and m/z data allowed us to identify a group of three additional metabolites (4–6) with mass spectra suggesting a close structural relationship to the fasamycins/formicamycins; these varied only in the number of methyl groups present. By searching for equivalent ions lacking chlorine atoms we identified three additional congeners (1–3). Compounds 1–6 were not initially observed in the UV chromatograms of the WT, and qualitative examination of the LCMS data suggested titres at least 100-fold lower than for the formicamycins. Furthermore, the m/z data implied the presence of a single nitrogen atom.

We repeated the experiment but included a set of WT strain fermentations to which sodium bromide (2 mM) was added, as our previous study showed this leads to the biosynthesis of brominated formicamycin congeners. This identified three conditional bromine-containing metabolites (7–9) with MS characteristics like those of 4–6 (Fig. 2 and Supplementary Fig. 1), supporting the hypothesis that these compounds represent biosynthetic congeners.

Isolation and structure elucidation and biological activity. HRMS data allowed us to predict molecular formulae for 1–9 that were used to search the online chemical database REAXYS. This suggested that they represented new structures. To isolate sufficient material for structure determination and antibacterial assays, the growth of S. formicace KY5 was scaled up (14 L; ~450 MS agar plates). After 9 days incubation at 30 °C the combined agar was chopped up, extracted with ethyl acetate, and the solvent removed under reduced pressure. The resulting extract was subjected to repeated rounds of reversed-phase HPLC followed by
Sephadex LH-20 size exclusion chromatography leading to small quantities of purified 1–6.

We first determined the structure of 2 which was isolated in the largest amount (~2 mg). HRMS and $^{13}$C NMR data indicated a molecular formula of $\text{C}_{28}\text{H}_{25}\text{NO}_6$ (calculated $m/z$ 472.1755 ([M + H]$^+$); observed $m/z$ 472.1753 ([M + H]$^+$); $\Delta = –0.4$ ppm) indicating 17 degrees of unsaturation. The UV spectrum showed absorption maxima at 229, 249, 272 and 392 nm indicating a relative activity with small $[\alpha]_20^D$ values between $+8^\circ$ and $+13^\circ$ indicating probable atropoisomerism about the axis of the C6–C7 bond. We attempted to assign the absolute configuration by combining computational approaches with electronic circular dichroism (ECD) spectra as done for the

Fig. 1 Chemical structures of metabolites isolated from Streptomyces formicae. Compounds 1–9 and 13 were discovered during this study whereas 10–12 and 14–26 were reported in an earlier study.\(^5\)
Fig. 2 Metabolomics pipeline. Dereplication, based on for biosynthetic mutants and exogenous bromide addition, led to the identification of the new congeners 4–9 with characteristic halogen-containing patterns (strain ΔforV lacks a halogenase gene; Δfor lacks the entire biosynthetic gene cluster)

Biosynthetic origins. Interrogation of LC-HRMS data from extracts of the Δfor and ΔforV mutants verified that the fasamycin/formicamycin BGC is required for formicapyridine production. The Δfor mutant does not produce 1–6 (Supplementary Fig. 3b vs. 3a), but production of all six compounds was restored upon ectopic expression of the P1-derived artificial chromosome (PAC) pESAC13–215-G containing the entire formicamycin BGC (Supplementary Fig. 3c). Similarly, the ΔforV mutant does not produce 4–6 (Supplementary Fig. 3d), but their production is re-established upon ectopic expression of forV under the control of the native promoter using an integrative plasmid (Supplementary Fig. 3e) (the construction of all strains, the requirement of the for BGC for production of the fasamycins and formicamycins, and the requirement of ForV for their halogenation were described previously5).

The very low level of formicapyridines made by the WT strain suggests they are shunt metabolites arising from aberrant derailment of fasamycin/formicamycin biosynthesis. On this basis, we suggest a biosynthetic pathway as described in Fig. 4. Assembly of the poly-β-ketone tridecaketide intermediate should proceed as previously proposed on route to the fasamycins5. This would be followed by a series of cyclization and aromatisation steps, presumably in a sequential manner, with the final cyclization event probably involving formation of the B-ring. However, premature hydrolysis of the acylcarrier protein from putative intermediate 27, prior to the action of a final cyclase, would liberate the enzyme-free, β-ketoacid species 28 that would be highly facile to spontaneous decarboxylation yielding the methylketone 29. The action of an endogenous aminotransferase could then generate either of the species 30 or 31, which could undergo cyclization, dehydration and oxidation to yield 1. Consistent with a shunt pathway, no aminotransferase encoding gene could be identified in the for BGC, and although numerous putative aminotransferase encoding genes were identified in the genome, none could be linked to a role in formicapyridine biosynthesis. Alternatively, direct addition of ammonia to the C23 ketone would generate an imine, tautomerize to the enamine, and the resulting amino group could then attack the C7 ketone. Subsequent dehydration would lead directly to the formicapyridine system with no requirement of an oxidation reaction.
This pathway requires that C25 of the formicapyridines originates from C2 of an acetate unit, with C1 lost via decarboxylation. To support our hypothesis, and the backbone structural assignments, *S. formicae* KY5 was cultivated on MS agar plates for two days and then overlaid with a solution of [1,2-13C2] sodium acetate (1 mL of 60 mM solution; final concentration 2 mM). This was repeated on the four following days, and after a total of 9 days incubation the agar was extracted.
with ethyl acetate and the most abundant congener was isolated (2, ~1 mg), and the $^{13}$C NMR spectrum was acquired. Due to the small amount of material, and overlapping signals, only eight of the intact acetate units could be unambiguously identified based on their coupling contests. In addition, a singlet was observed for C25 which showed weak enrichment (Fig. 5; Supplementary Fig. 4). This is consistent with our biosynthetic hypothesis, which requires that the carbon atom at C25 derives from C2 of a fragmented acetate unit. The data were in accordance with the proposed structure for 2.

**Targeted evolution of the for BGC.** Our biosynthetic proposal led to the following thought experiment. Suppose, in some environmental scenario, the presence of formicapyridines leads to a selection advantage. Is there then a single mutation in the BGC that could rapidly lead to enhanced levels of their production, and, in addition, could such changes lead to the reduction or even abolition of fasamycin/formicamycin biosynthesis?

The proposed biosynthetic pathway likely requires the action of multiple polyketide cyclase/ aromatase enzymes and this suggests that one cyclase could be dedicated to formation of ring-B in the final step of backbone biosynthesis (Fig. 4). On this basis we hypothesised that mutation of a gene encoding a putative ring-B cyclase might lead to the phenotype imagined in our thought experiment by eliminating the biosynthesis of fasamycins/formicamycins and shunting carbon flux into the proposed formicapyridine pathway. BLAST and conserved domain analysis, in conjunction with structural modelling using the Phyre2 web portal for protein modelling and analysis13, allowed us to identify five genes in the for BGC, which might encode potential cyclase/ aromatase enzymes (Table 1).

The gene product ForD shows significant sequence similarity to aromatase/cyclases such as the N-terminal domain (pfam 03364) of the archetypical tetracenomycin polyketide cyclase TcmN, which belongs to the Bet v1-like superfamily (cl10022)14. In vivo analysis15–17 and in vitro reconstruction18 showed that TcmN catalyses formation of the first two rings of tetracenomycin via sequential C9–C14 and C7–C16 cyclization/aromatization reactions. Thus, we predict ForD, which will play a key role in formation of the rings E and D during fasamycin biosynthesis. The ForL gene product belongs to the TcmI family of polyketide cyclases (cl24023; pfam 04673). The function of TcmI has been verified by in vivo mutational analysis19 and biochemical characterisation20,21. It catalyses cyclization of the final ring during formation of tetracenomycin F1 from tetracenomycin F2, and this reaction is remarkably like that proposed in our hypothesis for the formation of ring-B as the final step of fasamycin backbone assembly (Fig. 4). ForR is a homologue of the zinc containing polyketide cyclase RemF from the resistomycin BGC22,23. RemF is a single domain protein (pfam 07883) comprising the conserved barrel domain of the cupin superfamily (cl21464)24. Finally, the gene products ForS and ForU are both single domain proteins (pfam 03992) belonging to the ABM superfamily (cl1022). A notable member of this family is ActVA-orf6, which functions as a monoxygenase during biosynthesis of the polyketide antibiotic actinorhodin; the structure of this enzyme has been solved and it is topologically related to the PKS cyclase TcmI and homologues25. Notably, ABM family domains are found in several PKS cyclases including the C-terminal domain of BenH from the benastatin BGC26,27, which also comprises an N-terminal TcmN-like cyclase domain (pfam 03364); and WhiE Protein 1 and other members of the SchA/CurD-like family of PKS enzymes commonly associated with BGCs for the biosynthesis of spore pigments in *Streptomyces* spp28–30. SchA/CurD- and WhiE Protein 1-like enzymes are comprised of an N-terminal ABM domain and a C-terminal PKS cyclase domain (pfam 00486; superfamily cl24023).

To interrogate their roles, we used Cas9-mediated genome editing31 to make in-frame deletions in each of these five putative cyclase genes (forD, forL, forR, forS and forU). Three independent

---

**Fig. 5** The methyl group carbon C25 arises from C2 of acetate. Comparison of $^{13}$C NMR spectra for 2 isolated after growth in the presence of [1,2-$^{13}$C₂] sodium acetate. Integration of the $^{13}$C NMR spectra confirmed weak enrichment of the $^{13}$C isotope and the absence of coupling to any adjacent carbon atom. In contrast, the methyl group atom C24 shows enrichment and coupling to C1. The C2 atom of [1,2-$^{13}$C₂] sodium acetate is highlighted as a blue circle, the C1 atom as a black circle, and the coupled unit by a bold line. The methyl group carbon C25 arises from C2 of acetate. Comparison of $^{13}$C NMR spectra for 2 isolated after growth in the presence of [1,2-$^{13}$C₂] sodium acetate. Integration of the $^{13}$C NMR spectra confirmed weak enrichment of the $^{13}$C isotope and the absence of coupling to any adjacent carbon atom. In contrast, the methyl group atom C24 shows enrichment and coupling to C1. The C2 atom of [1,2-$^{13}$C₂] sodium acetate is highlighted as a blue circle, the C1 atom as a black circle, and the coupled unit by a bold line.
mutants generated from each gene deletion experiment, along with the WT strain, were grown on MS agar and cultured at 30 °C for 9 days. To assess differences in secondary metabolism production the ethyl acetate extracts from each culture were subjected to HPLC(UV) and LCMS analysis (Supplementary Data 3 and 4). The metabolic profiles showed that the ΔforD, ΔforL, ΔforR and ΔforU mutants lost the ability to produce formicapyridines, fasamycins and formicamycins (Supplementary Figs. 5–9), and no new shunt metabolites could be identified despite rigorous interrogation of the LCMS and LC(UV) data. These deletions could be rescued by complementation with the deleted gene under control of either the native promoter (ΔforD/forD and ΔforL/forL) or the constitutive ermE* promoter (ΔforR/forR), although the titres did not reach that of the WT strain in all cases. For the ΔforU mutant complementation with forU restored production of the non-halogenated congener 10 only, indicating a polar effect on the downstream halogenase forV gene. Subsequent complementation with a forUV cassette under the ermE* promoter in which the two genes were transcriptionally fused led to the production of halogenated fasamycin and formicamycin congeners.

In contrast, production of formicapyridines was increased ~25-fold in the ΔforS mutant when compared with the WT (Fig. 6). These effects were complemented by ectopic expression of forS under the control of the ermE* promoter. Moreover, the ΔforS mutant was significantly compromised in its ability to produce formicamycins, with their titre being reduced to approximately one third that of the WT strain. This is consistent with our hypothesis for formicapyride biosynthesis and suggests that ForS plays a role during B-ring closure, and that this constitutes the final step of fasamycin backbone biosynthesis. It also demonstrates that the final cyclization step can still occur without ForS. While carrying out analysis of the ΔforS mutant we identified a minor congener, which was not identified in any WT strain fermentations. Scale up growth (4 L) and solvent extraction, followed by isolation (3.4 mg) and structural elucidation identified this compound (13) as the C24-carboxyl analogue of fasamycin C (which we have named fasamycin F). This is the first congener identified with the C24-carboxyl group intact.

Table 1 Characteristics of the putative PKS cyclase genes in the for BGC

| Name   | Sequence identity | Top BLASTp hit                  | Annotation                                                                 | % Coverage | % Identity |
|--------|-------------------|---------------------------------|----------------------------------------------------------------------------|------------|------------|
| ForD   | WP_098245758.1    | WP_000362252.1                  | Polyketide cyclase (TcmN), Streptomyces clavuligerus ATCC 27064             | 95         | 73         |
| ForL   | WP_055544278.1    | WP_076977949.1                  | Tcm family type II polyketide cyclase, Streptomyces sparsogenes             | 97         | 52         |
| ForR   | WP_098245762.1    | AU14024.1                       | Polyketide WhiII Streptomyces sp., cupin domain                            | 96         | 75         |
| ForS   | WP_098245763.1    | SDF47214.1                      | Antibiotic biosynthesis monoxygenase, Lechevaleria fradiæ                  | 96         | 48         |
| ForU   | WP_098245765.1    | PZS28546.1                      | Antibiotic biosynthesis monoxygenase, Pseudonocardia sp.                   | 84         | 46         |

Our data suggest that formation of ring-B of the fasamycin scaffold is the final biosynthetic step, followed by thioster hydrolysis to liberate the ACP and subsequent decarboxylation to remove the carboxyl group attached to C24 (Fig. 5). Consistent with this, deletion of forS led to the isolation of shunt metabolite 13 with a C24-carboxyl group that was not observed from the WT. However, while ForS is implicated in ring-B cyclization it is not required for this role, nor, apparently, for the biosynthesis of any congeners. Rather, ForS seems to decrease the production of aberrant congeners from the pathway, e.g. formicapyridines, while increasing overall productivity. The most parsimonious interpretation of these data is that another of the for BGC gene products is the actual catalyst for ring-B formation, and that ForS acts as a chaperone, which modulates or stabilises the assembly, or arrangement, of a multienzyme complex to optimise production of the fasamycins, and therefore ultimately the formicamycins. This has the consequence of minimising the production of shunt metabolites (formicapyridines). Thus, while deletion of forS leads to the phenotype desired from our thought experiment, the mechanism by which this occurs is not what was anticipated. Interpretation of the for BGC bioinformatic analysis suggests the most likely candidate for a ring-B (final) cyclase is ForL due to its close relationship with TcmI, which catalyses a similar final cyclization step during tetracenomycin biosynthesis19–21. The apparent lack of any intermediates or shunt metabolites being accumulated by the remaining mutants is somewhat surprising and suggests an absolute requirement for the formation of a PKS-cyclase complex before biosynthesis can be initiated.

These observations are reminiscent of studies regarding the biosynthesis of pradimicin32. Pradimicins are pentangular polyketides similar to the benastatins26,27, and the BGC for their production contains three PKS cyclases equivalent to ForD (PdmD), ForL (PdmK) and ForR (PdmL)32,33. It also contains two ABM domain proteins (PdmH and PdmI; c.f. ForS and ForU) which were assigned monoxygenase roles. Through heterologous expression of PKS gene cassettes it was deduced that for the biosynthetic pathway to function correctly, and yield a pentangular backbone, all three cyclase genes (PdmDkl) plus the ABM domain monoxygenase PdmH must be co-expressed. In this pathway one oxidation reaction to form the quinone structure is required. These results led the authors to propose a model in which the two cyclases PdmKL and the monoxygenase PdmH form a multienzyme complex that engulfs the entire polyketide molecule during its assembly and works synergistically to ensure the correct reaction pathway occurs thereby minimising the production of shunt metabolites32. Similarly, formation of the unusual discoid metabolite resistomycin involves an extremely rare S-shaped folding pattern and requires the coordinated function, likely as a multienzyme complex, of core PKS proteins in addition to three distinct cyclase enzymes23. Moreover, heterologous reconstruction of the resistomycin pathway gave no products when the minimal PKS plus first cyclase were assembled23, a rare observation that is in keeping with our data showing the requirement of all the putative cyclases ForDLRU to produce any pathway derived metabolites, including shunts.
Biosynthesis of the for BGC polyketide backbones does not require a monooxygenase. Consistent with this, deletion of forS does not abolish polyketide production, but instead affects pathway productivity and fidelity. Thus, based on our mutational data and the observations discussed above, we hypothesise that ABM family enzymes can act as monooxygenases and/or as ancillary proteins to tune, in some way, the PKS enzyme complex function, and therefore the biosynthetic pathway, during aromatic polyketide biosynthesis. Genes encoding these proteins occur commonly in PKS BGCs, and multiple paralogues are often present even when monooxygenase reactions are not required. It is noteworthy that several PKS cyclases occur as fusion proteins commonly in PKS BGCs, and multiple paralogues are often isolated using a phenol:chloroform extraction method. Briefly, each strain was grown over the entire surface of the agar. Spores were harvested by applying 20% glycerol (1 mL). Glycerol stocks were stored at −80 °C. Streptomyces spores, material from a single colony was plated out using a sterile cotton bud and incubated at 30 °C for 7–10 days until a confluent lawn had grown over the entire surface of the agar. Spores were harvested by applying 20% glycerol (2 mL) to the surface of the agar plate culture and gently removing spores with a sterile cotton bud before storing them at −80 °C. E. coli strains were cultivated at 37 °C in LB Lennox Broth (LB), shaking at 220 rpm, or LB agar supplemented with antibiotics as appropriate. S. formicae was cultivated at 30 °C on MS agar or MYM agar with appropriate antibiotic selection.

Methods

Chemistry methods and materials. Unless stated otherwise all chemicals were supplied by Sigma-Aldrich or Fisher Scientific. [1,2-13C2] sodium acetate was purchased from CORTECNET. All solvents were of HPLC grade or equivalent. NMR spectra were recorded on a Bruker Avance III 400 MHz NMR spectrometer equipped with 5 mm BBFO Plus probe. 13C NMR spectra for 50,000 and isotopically enriched [1,2-13C2] sodium acetate were recorded on a Bruker Avance III 500 MHz NMR spectrometer equipped with a DUL cryoprobe at 30 °C. Mass spectrometry. The spray chamber conditions were: heat block, 300 °C; 25,000 °C; 50,000 °C; 25,000 °C. Unless otherwise stated samples were analysed by LCMS/MS on a Nexera/NX-1000 Prominence UHPLC system attached to a Shimadzu ion-trap time-of-flight mass spectrometer. The spray chamber conditions: heat block, 300 °C; 250 °C; 350 °C; 400 °C; 450 °C; 500 °C. Curved desorption line; interface (probe) voltage: 4.5 KV nebulizer gas flow rate 1.0 L min−1; drying gas on. The instrument was calibrated using sodium trifluoroacetate cluster ions according to the manufacturer’s instructions and run with positive–negative mode switching.

Standard microbiology and molecular biology methods. All strains used or made in this study are described in Supplementary Table 1. All plasmids and ePACS used are described in Supplementary Table 2. All PCR primers used are described in Supplementary Table 3. The compositions of media used are described in Supplementary Table 4. Standard DNA sequencing was carried out by Eurofins Genomics using the MiXisSeq kit (Ebersberg, Germany).

E. coli strains were cultivated at 37 °C in LB Lennox Broth (LB), shaking at 220 rpm, or LB agar supplemented with antibiotics as appropriate. S. formicae was cultivated at 30 °C on MS agar or MYM agar with appropriate antibiotic selection. To prepare Streptomyces spores, material from a single colony was plated out using a sterile cotton bud and incubated at 30 °C for 7–10 days until a confluent lawn had grown over the entire surface of the agar. Spores were harvested by applying 20% glycerol (2 mL) to the surface of the agar plate culture and gently removing spores with a sterile cotton bud before storing them at −80 °C. E. coli cells were made by pelleting the cells from an overnight E. coli culture (3–5 mL) in a bench top centrifuge (2733 x g, 5 min) and resuspending in fresh, sterile 1:1 2YT/40% glycerol (1 mL). E. coli cells were stored at −80 °C. S. formicae genomic DNA and cosmid ePAC DNA (from E. coli DH10B) was isolated using a phenolchloroform extraction method. Briefly, cells from an overnight culture (1 mL) were pelleted at 30,000 x g in a benchtop microcentrifuge.
and resuspended in solution 1 (100 µL) (50 mM Tris/HCl, pH 8; 10 mM EDTA). Alkaline lysis was performed by adding solution 2 (200 µL, prepared as follows: 200 mM NaOH; 1% SDS) and mixing by inverting. Solution 3 (150 µL) (3 M potassium acetate, pH 5.5) was added and samples mixed by inverting, before the soluable material was harvested by microcentrifugation at 30,000 × g for 5 min. The nucleic acid was extracted with 25:24:1 phenol/chloroform/isooamyl alcohol (400 µL) and DNA was precipitated in 600 µL ice-cold isopropanol. After centrifugation, the resulting DNA pellet was washed with 200 µL 70% ethanol and air dried before being resuspended in water for quantification on a Nanodrop 2000c UV–Vis spectrophotometer. Plasmid DNA was isolated from *E. coli* strains using a Qiagen miniprep kit according to the manufacturer’s instructions. PAC pESAC31–215-G DNA, containing the entire BGC, was used as the template for all PCR reactions. DNA amplification for cloning was conducted using Q5 Polymerase and diagnostic PCR was set up using PCRBI0 Taw Mix Red (PCR Biosystems), as per the manufacturer’s instructions. Amplified fragments and digested DNA products were purified on 1% agarose gels by electrophoresis and extraction using the Qiagen gel extraction kit as per the manufacturer’s instructions. For overlay bioassays soft LB agar (100% LB with 0.5% agar) was inoculated with *B. subtilis* 168 (10 mL; approximately OD600 = 0.6 mg). These pure compounds were subjected to LCMS and 1D and 2D NMR (100 MHz; 6500 scans; d3-methanol). Spectroscopic data is presented in Supplementary Figs. 13–43.

### Isolation and structure determination of fasamycin F (13)

*S. formicae* Aforos was cultivated on 4 L (~120 plates) of MS agar at 30 °C for nine days. The agar was extracted and purified using the methods described above to yield 13. Isolation and structure determination of fasamycin F (13) (2 mg), *pIJ10257*, respectively. Gibson assembly was used to fuse the gene(s) (and the *ermE* promoter and *pIJ10257*) into the plasmid containing the sgRNA at the *ermE* gene. The resulting transformants were selected on the appropriate selective medium (C18, 110 Å, 150 × 10 mm) using an Agilent 1100 series HPLC system and analysed by LCMS. Fractions 2–4 contained 1–6 and were further purified by chromatography over a Phenomenex Gemini-NX semi-prep reversed-phase column (C18, 110 Å, 150 × 10 mm) using an Agilent 1100 series HPLC system and eluting with the following gradient method: (mobile phase A: water +0.1% formic acid; mobile phase B: acetonitrile) 0–2 min 40%; 2–5 min 50%–100%; 20–21 min 100% B; and 40.1–45 min 40% B; flow rate 2 mL/min; injection volume 10 µL. Absorbance was monitored at 250 nm. Fractions (20 µL) were collected and analysed by LCMS. Fractions 2–4 contained 1–6 and were further purified by chromatography over a Phenomenex Gemini-NX semi-prep reversed-phase column (C18, 110 Å, 150 × 10 mm) using an Agilent 1100 series HPLC system and eluting with the following gradient method: (mobile phase A: water +0.1% formic acid; mobile phase B: acetonitrile) 0–2 min 40%; 2–20 min 40–100%; 20–21 min 100% B; and 40–41 min 100%–40% B; flow rate 2 mL/min; injection volume 100 µL. Absorbance was monitored at 390 nm. The samples were finally purified by Sephadex LH20 size exclusion chromatography with 100% methanol as the mobile phase. The isolated yields were: 1 (1 mg), 2 (2 mg), 3 (2 mg), 4 (0.7 mg), 5 (1 mg) and 6 (0.6 mg). These pure compounds were subjected to analysis by HRESIMS and 1D and 2D NMR and in the spectroscopic data (see Figs. 1 and 2). Spectroscopic and other data for each compound is presented in Supplementary Information (Supplementary Figs. 15–38 and Supplementary Note 1).

### Stable isotope feeding experiment

*S. formicae* was cultivated on MS agar (1 L, ~450 plates) at 30 °C for 9 days. The agar was sliced into small pieces and shaken twice with ethyl acetate (1 mL) using ultrasonication and successive extractions. These samples were subjected to LCMS and 1D and 2D NMR (100 MHz; 6500 scans; d3-methanol). The resultant spectra were matched to published data using Global Natural Products Social Molecular Networking. *S. formicae* was cultivated on 14 L (~120 plates) of MS agar at 30 °C for nine days. The agar was extracted and purified using the methods described above to yield 13. Isolation and structure determination of fasamycin F (13) (2 mg). These pure compounds were subjected to LCMS and 1D and 2D NMR (100 MHz; 6500 scans; d3-methanol). Spectroscopic data is presented in Supplementary Figs. 13–43.

### Coprogen content analysis of cyclostatine WT or mutant strains

(*n* = 3) were grown on MS agar at 30 °C for nine days. The agar was extracted and purified using the methods described above to yield 13. Isolation and structure determination of fasamycin F (13) (2 mg) was excised from each petri dish, sliced into small pieces and shaken with 100% methanol (1 mL) for 20 min. The ethyl acetate was transferred to a clean tube and the solvent removed under reduced pressure. The resultant extract was dissolved in methanol (200 µL) and analysed by LCMS but using the following modified UPLC method: Phenomenex Gemini C18 column (1000 x 2.1 mm, 100 Å); mobile phase A: water + 0.1% formic acid; mobile phase B: methanol. Elution gradient: 0–2 min, 50% B; 2–14 min, 50%–100% B; 14–18 min, 100% B; 18–18.1 min, 100–50% B; 18.1–20 min, 50% B; flow rate 1 mL/min−1; injection volume 10 µL. Calibration curves (Supplementary Fig. 10; Supplementary Data 3) were determined using standard solutions of fasamycin C (10, 20, 50 and 80 µM) for 20 min. The ethyl acetate was transferred to a clean tube and the solvent removed under reduced pressure. The resultant extract was dissolved in methanol (200 µL) and analysed by LCMS but using the following modified UPLC method: Phenomenex Gemini C18 column (100 x 2.1 mm, 100 Å); mobile phase A: water + 0.1% formic acid; mobile phase B: methanol. Elution gradient: 0–2 min, 50% B; 2–14 min, 50%–100% B; 14–18 min, 100% B; 18–18.1 min, 100–50% B; 18.1–20 min, 50% B; flow rate 1 mL/min−1; injection volume 10 µL. The content of 10 and 16 was determined by UV absorption at 285 nm. The content of 4 and 13 was determined by MS analysis of the base peak chromatogram (positive mode). Each standard solution was measured three times.

### 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 Fig. 6e–h are provided as a Source Data file.

Received: 11 March 2019 Accepted: 19 July 2019
Published online: 09 August 2019

### References

1. Flatt, P. M. & Mahmud, T. Biosynthesis of aminocyclitol–aminoglycoside antibiotics and related compounds. *Nat. Prod. Rep.* 24, 358–392 (2007).
2. Stauton, J. & Wilkinson, B. Biosynthesis of erythromycin and rapamycin. *Chem. Rev.* 97, 2611–2630 (1997).
3. Senges, C. H. R. et al. The secreted metabolome of *Streptomyces chartreusis* and implications for bacterial chemistry. *Proc. Natl Acad. Sci. USA* 115, 2490–2495 (2018).
4. Wilkinson, B. et al. Novel octaketide macrolides related to 6-deoxyerythronolide B provide evidence for iterative operation of the erythromycin polypeptide synthase. *Chem. Biol. 7*, 111–117 (2000).
5. Qin, Z. et al. Formycinamycins, antibacterial polypeptides produced by *Streptomyces formicae* isolated from African *Tetraporus* plant–ants. *Chem. Sci.* 8, 3218–3227 (2017).
6. Feng, Z., Kallida, K., Santoro, E., Brady, S. F. Functional analysis of environmental DNA-derived type II polypeptide synthases reveals structurally diverse secondary metabolic pathways. *Proc. Natl Acad. Sci. USA* 108, 12629–12634 (2011).
7. Panter, F., Krug, D., Baumann, S. & Muller, R. Self-resistance guided genome mining uncovers new topoisomerase inhibitors from myxobacteria. *Chem. Sci.* 9, 4988–4998 (2018).
8. Hedges, U. R., Ronan, J. L., Masschelein, J., de los Santos, E. L. C. & Challis, G. L. Thioester reduction and aldehyde transamination are universal steps in actinobacterial polypeptide alkyaloid biosynthesis. *Chem. Sci.* 8, 411–415 (2017).
9. Wang, M. et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* 34, 828–837 (2016).
10. Holmes, N. A. et al. Complete genome sequence of *Streptomyces formicae* KYS, the formycinamycin producer. *J. Biotechnol.* 265, 116–118 (2018).
11. Konig, W. A., Krauss, C. & Zahner, H. Metabolites from microorganisms, 6-chlorogenistein and 6,3'-dichlorogenistein. Helv. Chim. Acta 60, 2071–2078 (1977).
12. Zeigler, D. R. et al. The origins of 168, W23, and other Bacillus subtilis legacy strains. J. Bacteriol. 190, 6983–6995 (2008).
13. Kelley, L. A., Mezulis, S., Yates, C. M., Wasm, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modelling, prediction and analysis. Nat. Protoc. 10, 1864–1882 (2015).
14. Ames, D. et al. Crystal structure and functional analysis of tetracenomycin ARO/CYC: implications for cyclization of aromatic polyketides. Proc. Natl. Acad. Sci. USA 105, 5349–5354 (2008).
15. Summers, R. G., Wendt-Pienkowski, E., Motamed, H. & Hutchinson, C. R. Nucleotide sequence of the tcmF–tcmN region of the tetracenomycin C biosynthetic gene cluster of Streptomyces glaucescens and evidence that the tcmN gene encodes a multifunctional cyclase-dehydrogenase-O-methyl transferase. J. Bacteriol. 174, 1810–1820 (1992).
16. Shen, B., Summers, R. G., Wendt-Pienkowski, E. & Hutchinson, C. R. The Streptomyces glaucescens tcmKL polyketide synthase and tcmN polyketide cyclase genes govern the size and shape of aromatic polyketides. J. Am. Chem. Soc. 117, 6811–6821 (1995).
17. McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. Nature 373, 549–554 (1995).
18. Jawad, R. J. X. & Khosla, C. Heterologous expression, purification, reconstitution and kinetic analysis of an extended type II polyketide synthase. Chem. Biol. 6, 607–615 (1999).
19. Summers, R. G., Wendt-Pienkowski, E., Motamed, H. & Hutchinson, C. R. The tcmVI region of the tetracenomycin C biosynthetic gene cluster of Streptomyces glaucescens encodes the tetracenomycin F1 monooxygenase, tetracenomycin F2 cyclase, and, most likely, a second cyclase. J. Bacteriol. 175, 7571–7580 (1993).
20. Shen, B. & Hutchinson, C. R. Tetracenomycin F2 cyclase: intramolecular aldehyde condensation in the biosynthesis of tetracenomycin C in Streptomyces glaucescens. Biochemistry 32, 11149–11154 (1993).
21. Thompson, T. B., Katayama, K., Watanabe, K., Hutchinson, C. R. & Raymont, I. Structural and functional analysis of tetracenomycin F2 cyclase from Streptomyces glaucescens. A type II polyketide cyclase. J. Biol. Chem. 279, 37956–37963 (2004).
22. Jakobi, K. & Hertweck, C. A gene cluster encoding resistomycin biosynthesis in Streptomyces resistomycins. Structured and functional analysis of the resistomycin synthetase. J. Am. Chem. Soc. 126, 2298–2299 (2004).
23. Fritzsche, K., Ishida, K. & Hertweck, C. Exploration of polyketide cyclization in the resistomycin pathway. J. Am. Chem. Soc. 130, 8307–8316 (2008).
24. Silvennoinen, L., Sandalova, T. & Schneider, G. The polyketide cyclase RemF from Streptomyces resistomycins contains an unusual octahedral zinc binding site. FEBS Lett. 583, 2917–2921 (2009).
25. Sciara, G. et al. The structure of ActIVA-Orf6, a novel type of monooxygenase involved in actinorhodin biosynthesis. EMBO J. 22, 205–215 (2003).
26. Xu, Z., Schenk, A. & Hertweck, C. Molecular analysis of the benznatcin biosynthetic pathway and genetic engineering of altered fatty acid-polyketide hybrids. J. Am. Chem. Soc. 129, 6022–6030 (2007).
27. Lackner, G. et al. Biosynthesis of pentangular polyphenolic deductions from the benznatcin and grisoroehdin pathways. J. Am. Chem. Soc. 129, 9306–9312 (2007).
28. Davis, N. K. & Chater, K. F. Spore colour in Streptomyces coelicolor A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. Mol. Microbiol. 4, 1679–1691 (1999).
29. Blanco, G., Pereda, A., Méndez, C. & Salas, J. A. Cloning and disruption of a DNA fragment of Streptomyces halstedii involved in the biosynthesis of a spore pigment. Gene 112, 59–65 (2000).
30. Yu, T.-W. & Hopwood, D. A. Ectopic expression of the Streptomyces coelicolor whiE genes for the polyketide spore pigment synthesis and their interaction with the act genes for actinorhodin biosynthesis. Mol. Microbiol. 41, 2779–2792 (1999).
31. Cobb, R. E., Wang, Y. & Zhao, H. High-efficiency multiplex genome editing of Streptomyces species using and engineered CRISPR/Cas system. ACS Synth. Biol. 4, 723–728 (2015).
32. Zhan, J., Watanabe, K. & Tang, Y. Synergistic actions of a monooxygenase and cyclases in aromatic polyketide biosynthesis. ChemBioChem 9, 1710–1715 (2008).
33. Kim, R. C., Lee, J. M., Ahn, J. S. & Kim, B. S. Cloning, sequencing and characterization of the pradimicin biosynthetic gene cluster of Actinomadura hibusca P157–2. J. Microbiol. Biotechnol. 17, 830–839 (2007).
34. Miles, A. A., Misra, S. S. & Irwin, J. O. The estimation of the bactericidal power of the blood. J. Hyg. 38, 732–749 (1938).
35. Kieser, T., Bibb, N. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. Practical Streptomyces Genetics. (The John Innes Foundation Press, Norwich, 2000).

Acknowledgements
This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) via Institute Strategic Program Grant BBS/E/10009790 to the John Innes Centre, BBSRC Responsive Mode Grants BB/S00900/1 to B.W. and BB/S00811X/1 to M.I.H., and by a Norwich Research Park BBSRC Postdoctoral Training Programme Studentship BB/M011216/1 (R.D.). We thank Dr Lionel Hill and Dr Gerhard Saalbach (JIC) for excellent metabolomics support.

Author contributions
Z.Q., R.D., B.W. and M.I.H. designed the research. Z.Q., R.D., B.W. and M.I.H. wrote the paper and all authors commented. Z.Q. performed the chemical experiments, R.D. performed the molecular genetics experiments.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-11538-6.

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 Jurgen Rohr, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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