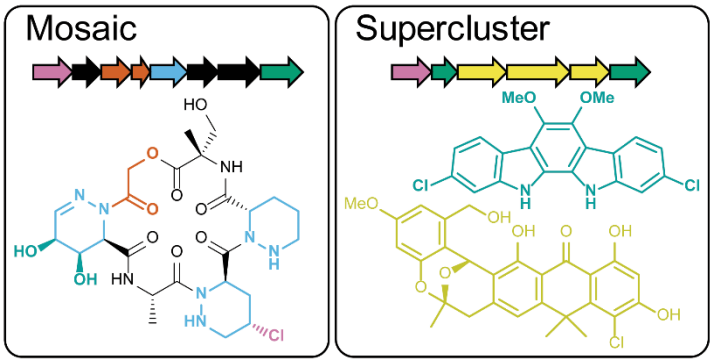


Chlorinated metabolites from *Streptomyces* sp. highlight the role of biosynthetic mosaics and superclusters in the evolution of chemical diversity

Mahmud T. Morshed,[a] Ernest Lacey,[a,b] Daniel Vuong,[b] Alastair E. Lacey,[b] Soo Sum Lean,[c] Stephen A. Moggach,[c] Peter Karuso,[a] Yit-Heng Chooi,[c] Thomas J. Booth*[c] and Andrew M. Piggott*[a]

LCMS-guided screening of a library of biosynthetically talented bacteria and fungi identified *Streptomyces* sp. MST-144321 as a prolific producer of chlorinated metabolites. We isolated and characterised six new and nine reported compounds from MST-144321, belonging to three discrete classes – the depsipeptide svetamycins, the indolocarbazole borregomycins and the aromatic polyketide anthrabenzoxocinones. Following genome sequencing of MST-144321, we describe, for the first time, the svetamycin biosynthetic gene cluster (*sve*), its mosaic structure and its relationship to several distantly related gene clusters. Our analysis of the *sve* cluster suggested that the reported stereostructures of the svetamycins may be incorrect. This was confirmed by single-crystal X-ray diffraction analysis, allowing us to formally revise the absolute configurations of svetamycins A–G. We also show that the borregomycins and anthrabenzoxocinones are encoded by a single supercluster (*bab*) implicating superclusters as potential nucleation points for the evolution of biosynthetic gene clusters. These clusters highlight how individual enzymes and functional subclusters can be co-opted during the formation of biosynthetic gene clusters, providing a rare insight into the poorly understood mechanisms underpinning the evolution of chemical diversity.

Introduction

To date, over 10,000 halogenated metabolites have been discovered from natural sources,[1-3] exhibiting a diverse range of biological activities.[4] Notable examples include the glycopeptide antibiotics vancomycin[5] and teicoplanin,[6, 7] the ribosome-targeting chloramphenicol[8] and chlortetracycline,[9] the first-in-class macrolide antibiotic fidaxomicin[10] and the marine natural product salinosporamide A,[11] a second-generation irreversible pan-proteasome inhibitor. The vast majority of halogenated natural products are chlorinated or brominated, with only limited examples of iodinated and fluorinated natural products having been reported to date.[12-14] Given the broad range of potential substrates, halogenation of secondary metabolites can be enzyme-catalysed by a variety of different oxidative and nucleophilic mechanisms.[15-17] Oxidative halogenations involve activation of the halogen *via* a one- or two-electron transfer, whereas nucleophilic halogenases utilise *S*-adenosyl methionine to catalyse nucleophilic substitutions.[18] The resulting halogenation is often central to the biological activity of the metabolite. For example, halogens can stabilise bioactive atropisomers through steric influence,[19, 20] facilitate the formation of covalent adducts within enzyme active sites by acting as leaving groups[21, 22] and facilitate transport across membranes and the blood-brain barrier.[23, 24]

Given the significant role that halogens play in shaping biological activity, the search for halogenated metabolites from natural sources has been an important focus of our research to date.[25-27] Halogenated compounds are readily identified in crude extracts using liquid chromatography–mass spectrometry (LCMS) due to their distinctive isotopic distributions. From our in-house library of extracts of >450,000 microorganisms, we have identified a subset of 10,000 talented strains based on their UV-vis spectra and biological activities. Systematic LCMS-guided screening of these 10,000 extracts identified *Streptomyces* sp. MST-144321 as a unique strain producing three discrete classes of chlorinated metabolites: the depsipeptide svetamycins, the indolocarbazole borregomycins and the aromatic polyketide anthrabenzoxocinones (ABXs). We isolated and characterised 6 new and 9 previously reported metabolites from MST-144321 and evaluated these compounds for activity against a panel of bacteria, fungi, protozoa and mammalian cells. We also sequenced and analysed the genome of MST-144321, leading to the identification of two biosynthetic gene clusters (BGCs) responsible for the biosynthesis of the three classes of metabolites: the *sve* BGC – a mosaic cluster responsible for the production of the svetamycins – and the *bab* BGC – a supercluster responsible for biosynthesis of the ABXs and borregomycins. These clusters provide unique insights into how individual enzymes and functional subclusters can be co-opted during evolution to generate new chemical diversity.

[a] *Department of Molecular Sciences, Macquarie University, NSW 2109, Australia.*
[b] *Microbial Screening Technologies, Smithfield, NSW 2164, Australia.*
[c] *School of Molecular Sciences, University of Western Australia, Perth, WA 6009, Australia.*

† Electronic Supplementary Information (ESI) available: General experimental details, cultivation and extraction of *Streptomyces* sp. MST-144321, isolation, purification, structure elucidation, characterisation, single-crystal X-ray diffraction analysis, biological screening of compounds, extraction and sequencing of genomic DNA and genomic features of MST-144321.

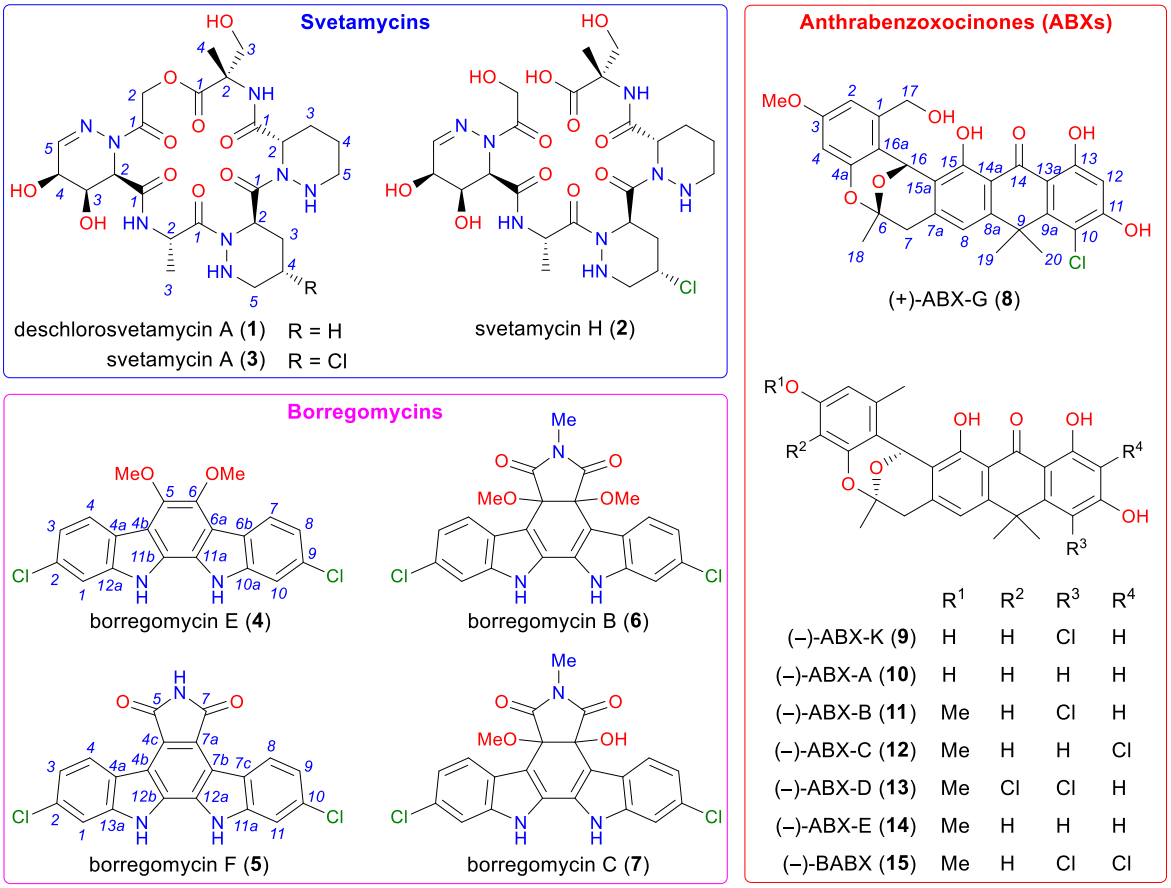


Fig. 1 Cyclic depsipeptides (**1**–**3**), indolocarbazoles (**4**–**7**) and anthrabenzoxocinones (**8**–**15**) isolated from *Streptomyces* sp. MST-144321.

Results and Discussion

Isolation, cultivation, extraction and purification of metabolites

Streptomyces sp. MST-144321 was isolated from pasture soil collected near Yass, New South Wales (NSW), Australia. Large-scale cultivation on pearl barley, followed by solvent partitioning, silica gel chromatography and preparative reversed-phase HPLC (Fig. S3†), facilitated the isolation and characterisation of fifteen major metabolites belonging to three discrete classes: svetamycins, borregomycins and ABXs (Fig. 1). Svetamycins are a family of piperazic acid-containing cyclic depsipeptides that were first isolated from *Streptomyces* sp. DSM 14386 in 2017.[28] Three svetamycins were isolated from MST-144321, including two new analogues, deschlorosvetamycin A (**1**) and svetamycin H (**2**), and the previously described analogue svetamycin A (**3**).[28] Borregomycins are a family of chlorinated indolocarbazoles that were identified in 2012 from a cosmid library of environmental DNA extracted from desert soil.[29] Four borregomycin analogues were isolated from MST-144321, including two new metabolites, borregomycins E (**4**) and F (**5**), and the previously described borregomycins B (**6**) and C (**7**).[29] ABXs are a family of chlorinated aromatic polyketides. The first member of the class (L-755,805) was isolated in 1995 from a dung-derived actinomycete culture,[30] with subsequent isolations from two *Streptomyces* spp.,[31, 32] and more recent biosynthetic pathway manipulations,[33] increasing the total number of reported ABXs to 20. Eight ABX analogues were isolated from MST-144321, including two new metabolites, (+)-ABX-G (**8**) and (–)-ABX-K (**9**), along with the previously reported (–)-ABX-A–E (**10**–**14**)[33] and (–)-bischloroanthrabenzoxocinone ((–)-BABX; **15**).[31] The isolated compounds were characterised by detailed spectroscopic analysis and comparison with literature data.

Structure elucidation of new metabolites

Deschlorosvetamycin A (**1**) was isolated as colourless amorphous powder. HR-ESI(+)-MS analysis of **1** revealed a protonated molecule ($[M + H]^+$ *m/z* 597.2623) indicative of a molecular formula $C_{24}H_{36}N_8O_{10}$ (Δmmu +0.4) containing one chlorine fewer and one hydrogen more than the previously reported compound svetamycin A (**3**).[28] The ^1H and ^{13}C NMR data for **1** (Table S1†; Figs. S4-S5†) were almost identical to those reported for **3**, with the only significant differences being the absence of the H-$4_{\gamma\text{-ClPip}}$ methine group (δ_H 4.35; δ_C 52.6) and the presence of an additional methylene group (δ_H 1.56, 1.50; δ_C 21.0). Diagnostic COSY correlations between H_2-3, H_2-5 and δ_H 1.56/1.50, and an HMBC correlation from H-2 to δ_C 21.0, positioned this methylene group at C-4 and confirmed **1** to be the deschloro analogue of **3**.

Svetamycin H (**2**) was isolated as colourless amorphous powder. HR-ESI(–)-MS analysis of **2** revealed an isotopic cluster ($[M - H]^-$

m/z 647.2196 / 649.2173 in a 3:1 ratio) indicative of a molecular formula $C_{24}H_{37}ClN_8O_{11}$ (Δmmu –0.1) containing two hydrogens and one oxygen (H_2O) more than **3**. The NMR data for **2** (Table S2†, Figs. S10-S11†) were also very similar to those reported for **3**,[28] with the only significant differences being the shielding of H_2-2_{Haa} (δ_H 4.34/4.29 in **2** vs. δ_H 5.32/4.61 in **3**), the slight deshielding of C-$1_{\alpha MeSer}$ (δ_C 173.6 in **2** vs. δ_C 171.8 in **3**) and the appearance of a broad exchangeable proton at δ_H 12.35. These observations, together with the absence of an HMBC correlation from H_2-2_{Haa} to C-$1_{\alpha MeSer}$, are indicative of **2** being the non-lactonised (free acid) analogue of **3**, as shown in Fig. 1.

Borregomycin E (**4**) was isolated as greenish-black amorphous powder. HR-ESI(–)-MS analysis of **4** revealed an isotopic cluster ($[M - H]^-$ *m/z* 383.3656 / 385.0327 / 387.0297 in a 10:6:1 ratio) indicative of a molecular formula $C_{20}H_{14}Cl_2N_2O_2$ (Δmmu –0.3). The ^{13}C NMR spectrum of **4** (Fig. S19†) contained only 10 resonances, suggesting the molecule must be symmetric. Examination of the ^1H NMR spectrum of **4** (Fig. S18†) revealed one pair of ortho-coupled aromatic protons (δ_H 8.15, d, *J* = 8.3 Hz; δ_H 7.21, dd, *J* = 8.3, 1.8 Hz), one meta-coupled aromatic proton (δ_H 7.76, d, *J* = 1.8 Hz), one methoxy group (δ_H 4.05, s) and one exchangeable proton (δ_H 11.27, s). Detailed analysis of the HMBC NMR spectrum of **4** (Fig. S21†) defined a 11,12-dihydroindolo[2,3-*a*]carbazole scaffold, with the two methoxy groups located on C-5 and C-6 of the central aromatic ring. The splitting patterns observed for the indole aromatic protons suggested chlorination at either C-2 and C-9 or at C-3 and C-8, with the former supported by ROESY correlations between 11/12-NH and H-10/1 and between H-4/7 and 5/6-OMe. Thus, the structure of **4** was elucidated as shown in Fig. 1.

Borregomycin F (**5**) was isolated as yellow amorphous powder. HR-ESI(–)-MS analysis of **5** revealed an revealed an isotopic cluster ($[M - H]^-$ *m/z* 391.9996 / 393.9964 / 395.9934 in a 10:6:1 ratio) indicative of a molecular formula $C_{20}H_9Cl_2N_3O_2$ (Δmmu –0.3) containing one nitrogen more and five hydrogens fewer than **4**. The NMR data for **5** (Table S5†, Figs. S24-S25†) were indeed very similar to those for **4**, with the main differences being the absence of signals associated with the two methoxy groups and the presence of signals for an additional carbonyl group (δ_C 171.0) and an additional exchangeable proton (δ_H 11.07, s) integrating for half the area of the other protons. HMBC correlations from this exchangeable proton to the carbonyl carbon, and also to the aromatic carbon at δ_C 120.1, suggested the presence of an imide ring fused to the central aromatic ring, as observed in the four previously reported borregomycin analogues.[29] The remaining 2D NMR correlations were consistent with the structure of **5** shown in Fig. 1.

(+)-ABX-G (**8**) was isolated as an amorphous yellow powder. HR-ESI(+)-MS analysis of **8** revealed an isotopic cluster ($[M + H]^+$ *m/z* 525.1303 / 527.2973 in a 3:1 ratio) indicative of a molecular formula $C_{28}H_{25}ClO_8$ (Δmmu –0.7) containing one oxygen more than the previously reported compound (–)-ABX-B (**11**).[33] The NMR data for **8** (Table S8†, Figs. S34-S35) were almost identical to those reported for **11**, with the only significant differences being the absence of signals for CH_3-17 and the presence of signals for an additional methylene group (δ_H 4.91, dd, *J* =14.4, 5.7 Hz; δ_H 4.54, dd, *J* = 14.4, 5.7 Hz) and an additional exchangeable proton (δ_H 5.09, t, *J* = 5.7 Hz). COSY correlations between the exchangeable proton and the methylene protons and HMBC correlations from the methylene protons to C-1, C-2 and C-16a, confirmed **8** to be the 17-hydroxy analogue of **11**. Interestingly, **8** exhibited a positive specific optical rotation (+19 in $CHCl_3$), while **11** exhibited a negative optical rotation (–269 in MeOH). Given the close structural similarity between **8** and **11**, this suggested that the compounds must have opposite configurations at C-6 and C-16. This hypothesis was supported by TDDFT calculations of the ECD spectra of (6*R*,16*R*)-**8** and (6*S*,16*S*)-**8**, with the latter closely matching the experimental ECD spectrum of **8** (Fig. S71†). Therefore, the structure of **8** was assigned as shown in Fig. 1.

(–)-ABX-K (**9**) was isolated as pale-yellow powder. HR-ESI(+)-MS analysis of **9** revealed an isotopic cluster ($[M + H]^+$ *m/z* 495.1198 / 497.1170 in a 3:1 ratio) indicative of a molecular formula $C_{27}H_{23}{}^{35}ClO_7$ (Δmmu +0.8) The ^1H and ^{13}C NMR spectra of **9** were identical to the those for the previously reported compound (+)-ABX-C.[33] However, **9** exhibited a negative specific optical rotation (–80 in $CHCl_3$), while (+)-ABX-C exhibited a positive specific optical rotation (+221 in MeOH), suggesting the compounds must be enantiomers. The ECD spectrum of **9** exhibited negative Cotton effects at 280 and 371 nm and a positive Cotton effect at 307 nm, in accord with all the other (6*R*,16*R*)-(–)-ABX analogues isolated (Fig. S70†). This was supported by TDDFT calculations of the ECD spectra of (6*R*,16*R*)-**9** and (6*S*,16*S*)-**9**, with the former closely matching the experimental ECD spectrum of **9** (Fig. S72†). Taken together, these observations allowed the structure of **9** to be assigned as shown in Fig. 1.

Bioassays

The fifteen metabolites isolated from MST-144321 were assessed for *in vitro* toxicity against the Gram-positive bacteria *Bacillus subtilis* (ATCC 6633) and *Staphylococcus aureus* (ATCC 25923), the Gram-negative bacterium *Escherichia coli* (ATCC 25922), the fungi *Candida albicans* (ATCC 10231) and *Saccharomyces cerevisiae* (ATCC 9763), the protozoan *Tritrichomonas foetus* (KV-1), and NS-1 (ATCC TIB-18) mouse myeloma cells (Table 1). None of the metabolites showed activity against *E. coli*, *C. albicans* or *S. cerevisiae* up to 100 μg/mL. The svetamycin analogues **1**–**3** showed no antibacterial, antifungal, antiprotozoal or antitumour activity, suggesting they must have a hitherto undiscovered biological role.[34] The borregomycin analogues **4**, **6** and **7** showed moderate activity against Gram-positive bacteria (MIC 3.1–6.3 μg/mL) and moderate tumour cell cytotoxicity (MIC 3.1–6.3 μg/mL), with **4** also showing weak antiprotozoal activity against *T. foetus* (MIC 25 μg/mL). However, the closely related analogue **5** showed no activity against any of the organisms tested, highlighting the importance of at least one methoxy group for bioactivity.

Table 1 *In vitro* biological activities of isolated compounds from *Streptomyces* sp. MST-144321

Compound	MIC (μg/mL)			
	B. subtilis (ATCC 6633)	*S. aureus* (ATCC 25923)	*T. foetus* (KV-1)	NS-1 (ATCC TIB-18)
deschlorosvetamycin A (**1**)	> 100	> 100	> 100	100
svetamycin H (**2**)	> 100	> 100	> 100	> 100
svetamycin A (**3**)	100	100	> 100	50
borregomycin E (**4**)	6.3	6.3	25	6.3
borregomycin F (**5**)	> 100	> 100	> 100	> 100
borregomycin B (**6**)	3.1	3.1	> 100	3.1
borregomycin C (**7**)	6.3	6.3	> 100	6.3
(+)-ABX-G (**8**)	3.1	6.3	100	100
(–)-ABX-K (**9**)	0.4	0.8	50	25
(–)-ABX-A (**10**)	0.8	0.8	50	25
(–)-ABX-B (**11**)	0.8	3.1	> 100	12.5
(–)-ABX-C (**12**)	0.8	1.6	50	6.3
(–)-ABX-D (**13**)	1.6	50	> 100	25
(–)-ABX-E (**14**)	3.1	100	100	12.5
(–)-BABX (**15**)	0.2	0.8	100	6.3
control	0.4[a]	0.8[a]	0.8[b]	0.04[c]

[a] tetracycline HCl; [b]metronidazole; [c]staurosporine

Borregomycins have previously been reported as having antibacterial and cytotoxicity activities,[29] with several modes of action having been ascribed to the indolocarbazoles and structurally related indolotryptolines.[35-38] Significantly, the indolocarbazoles are known to interact with the ATP binding sites of proteins, such as kinases and DNA topoisomerases.[35, 36, 39] It is interesting to note that **5**, lacking both methoxy groups, demonstrated no bioactivity, while **4**, lacking the entire pyrrolidinedione unit, retained antibacterial and cytotoxicity activities, and gained mild inhibitory activity against *T. foetus*.

The ABX analogues all showed potent antibacterial activities against *B. subtilis* (MIC 0.2–3.1 μg/mL) and *S. aureus* (MIC 0.8–6.3 μg/mL). Significantly, **9** exhibited activities against Gram-positive bacteria that were equivalent to the control compound tetracycline and two-fold more potent than the 6*S*,16*S* enantiomer (+)-ABX-C.[33] Several of the ABX analogues also showed weak antiprotozoal activity against *T. foetus* (MIC 50–100 μg/mL) and weak-to-moderate NS-1 tumour cell cytotoxicity (MIC 6.3–100 μg/mL). ABXs are known to inhibit bacterial fatty acid biosynthesis, reportedly by binding to the condensation protein FabF[31]. Given **8** and **9** retain the essential dimethylanthrone moiety,[33] it is likely that they achieve their antibacterial activity through the same mode of action.

Biosynthetic gene cluster analysis

The genome of *Streptomyces* sp. MST-144321 was sequenced *via* the Illumina sequencing platform (see ESI†). AntiSMASH[40] analysis of the sequence revealed the presence of 24 putative biosynthetic gene clusters (BGCs) and 6 partial BGCs within the genome of MST-144321 (Table S24†). To identify halogenases involved in natural product biosynthesis, protein coding sequences were assigned secondary metabolism Clusters of Orthologous Groups (smCOGs)[41] using HMMER.[42] An intact putative BGC for svetamycin biosynthesis (*sve*) and a borregomycin-anthrabenzoxocinone 'supercluster' (*bab*) were identified (Genbank MW626911 and MW810086, respectively).

Svetamycin is a biosynthetic mosaic

We identified the putative svetamycin BGC (*sve*) by searching the genome of MST-144321 for a six-modular NRPS in proximity to genes with homology to the BGCs of other piperazic acid-containing compounds.[43-46] The genomic assembly contained four predicted NRPS BGCs, only one of which contained the correct number of modules. The putative cluster contained 41 genes spanning 68.3 kb (Fig. 2 and Table S20†). To identify similar known BGCs, the predicted protein sequences of all 41 genes were searched against the MIBiG[47] database using cblaster.[48] Twenty-four BGCs were identified as containing more than three clustered homologues. The top three hits, each containing eight homologues, were cadasides,[49] kutzneride[44] and polyoxypeptin[46] BGCs (Table S21†). Interestingly, syntenic subclusters were also identified in the BGCs of naphithyridomycin,[50] sceliphrolactam[51] and quinocarcin.[52]

The *sve* BGC contains two genes, *sveA1* and *sveA2*, that together encode six NRPS modules (Fig. 2). Phylogenetic analysis of the six adenylation domains revealed that the first, third and fourth adenylation domains of *sveA2* claded together with known piperazic acid adenylating domains from previously described

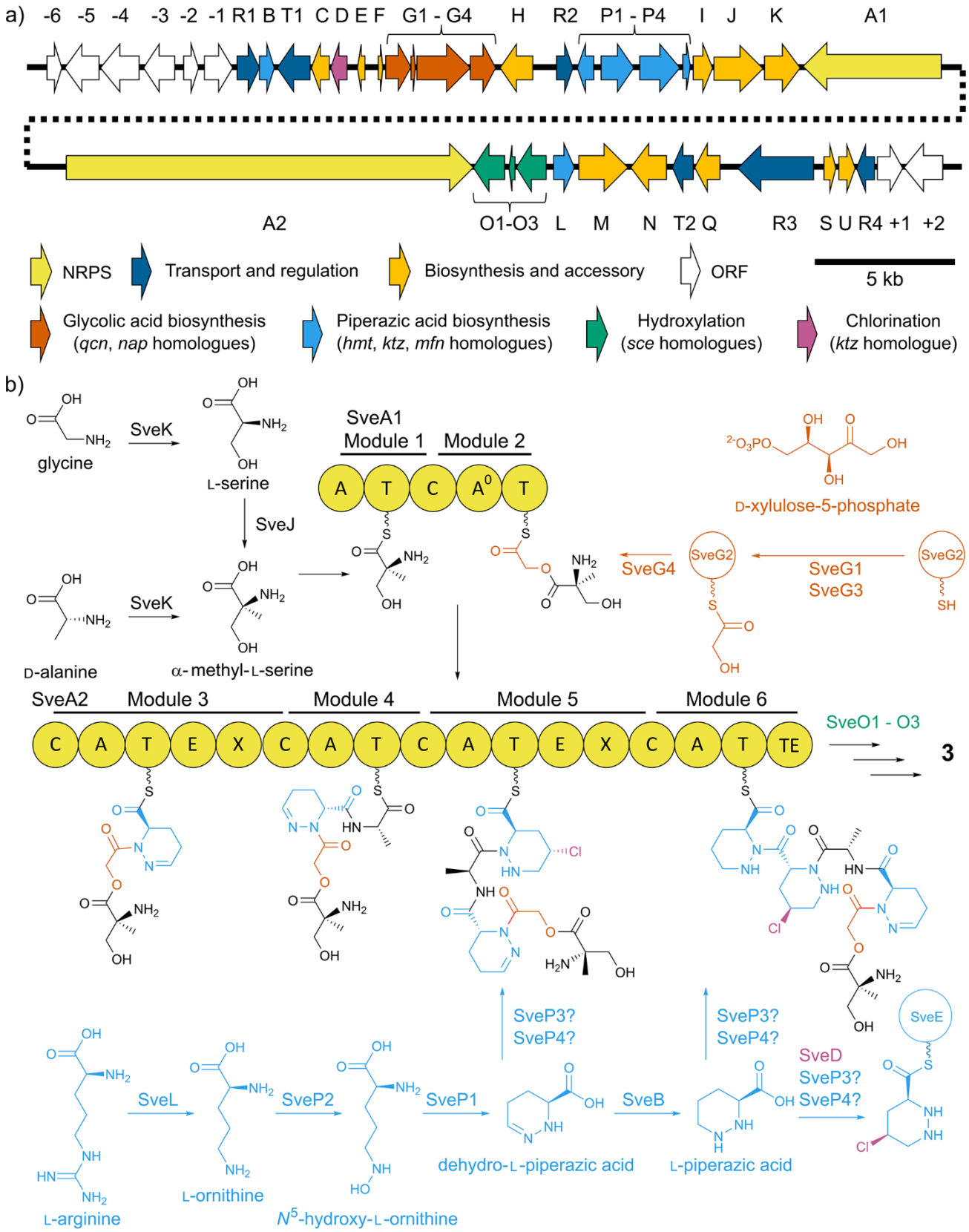


Fig. 2 a) The architecture of the svetamycin (*sve*) biosynthetic gene cluster; b) the biosynthesis of svetamycins. Biosynthetic steps are coloured by evolutionary relationship to the previously described biosynthesis of quinocarcins and naphthyridinomycin (orange), sceliphrolactam (green) and kutzneride and other non-ribosomal peptides containing piperazic acid (blue) and chloropiperazic acid (purple). The assembly-line is represented by the following domains: adenylation (A), thiolation (T), condensation (C), epimerisation (E), TIGR1720 (X) and thioesterase (TE).

NRPSs of kutzneride[44] (*ktz*) and polyoxypeptin[46] (*ply*). Additionally, *sveA2* encodes a terminal thioesterase, indicating it encodes the latter portion of the assembly line. Given this information, the order of incorporation must be as follows: α-methyl-L-serine, glycolic acid, β,γ-dihydroxydehydropiperazic acid, L-alanine, γ-chloropiperazic acid, terminating with a piperazic acid residue (Fig. 2).

The starter unit, α-methyl-L-serine, is likely biosynthesised by SveK, a serine hydroxymethyltransferase with homology to AmiS (59%) and FmoH (56%), which has been shown to produce α-methyl-L-serine through the hydroxymethylation of D-alanine (Fig. 2).[53, 54] As with the JBIR-34 (*fmo*) producer *Streptomyces* sp. Sp080513GE-23, we were able to find an exogenous alanine racemase in the genome of MST-144321 with 81% identity to the racemase SCO4745.[54] However, SveK also shows homology to AsmD[55] (51%), suggesting it may accept glycine as a substrate. In this case, α-methylation by the radical-SAM protein SveJ, encoded directly upstream of SveK, may provide an alternate route to α-methyl-L-serine. Homologues of SveJ are involved in branching methylations in a variety of natural product pathways, supporting this hypothesis.[56, 57]

The biosynthesis and subsequent incorporation of glycolic acid is homologous to the mechanisms involved in the biosynthesis of naphthyridinomycin (*nap* BGC) and quinocarcin (*qcn* BGC). There are four genes shared between the three BGCs, represented by *sveG1-G4*. Through this mechanism, glycolic acid is biosynthesised from D-xylulose-5-phosphate by SveG1 (NapB/Qcn7) and SveG3 (NapD/Qcn9).[58] SveG1G3 catalyses the transfer of glycolaldehyde to the prosthetic lipoyl group of SveG3 *via* a thiamine diphosphate intermediate. Then, the complex catalyses the transacylation of the resulting glycolyl group to the phosphopantetheine arm of the standalone ACP SveG2 (NapC/Qcn8).[58] Transacylation from SveG2 to the assembly line SveA1 is hypothesised to be catalysed by the 3-oxoacyl-ACP synthase SveG4 (NapD/Qcn10).[52] Analysis of the SveA1A2 A domains revealed that the phosphate-loop (Walker A motif) of the second A domain of SveA1 is completely mutated (Fig. S74†) rendering the domain inactive and therefore unable to load glycolic acid to the downstream T domain itself. This is analogous to the inactive A domains of the *nap* and *qcn* BGCs.[50, 52] It is noteworthy that this mechanism is not homologous to the biosynthesis of the structurally similar gerumycin family of peptides[28, 59, 60] (Table S28†). Following the incorporation of glycolic acids, the assembly line is able to proceed in a collinear fashion, incorporating β,γ-dihydroxydehydropiperazic acid, L-alanine, γ-chloropiperazic acid and L-piperazic acid residues.

SveP1-4 encode a subcluster comprising genes from other piperazic acid incorporating NRPSs. SveP2 and SveP1 encode an ornithine *N*-hydroxylase and a haem-dependent *N*-hydroxy ornithine cyclase homologous to KtzI and KtzT responsible for the production of piperazic acid during kutzneride biosynthesis.[44, 61, 62] Meanwhile, SveP3 and SveP4 encode a standalone adenylating protein and an acyl-carrier protein respectively. During marformycin biosynthesis, deletion of the *sveP3* homologue, *mnfK,* resulted in the accumulation of the peptide intermediate prior to incorporation of piperazic acid. Therefore, we hypothesised that *mfnK* and its neighbouring *mfnL* (encoding an acyl carrier protein) were required for the incorporation of piperazic acid into the nascent peptide.[45] Presumably, piperazic acid is incorporated into the svetamycin assembly-line in a similar fashion. With regards to 5-chloro-L-piperazic acid, SveD shares over 75% identity with KthP,[63] an exogenous chlorinase mononuclear iron halogenase family that has been shown to halogenate ACP bound substrates. Given that *sveD* is clustered together with *sveE*, encoding another standalone acyl-carrier protein, it is likely that these two genes are responsible for the conversion of L-piperazic acid to 5-chloro-L-piperazic acid prior to loading to the assembly line.

Interestingly, both modules 3 and 5 contain TIGR1720 domains, which are speculated to be involved in post-condensation modifications.[64] It is possible that these domains may recruit the required tailoring enzymes or PCPs to the assembly-line. Modules 3 and 5 also contain epimerase domains, which are predicted to convert the L-(3*S*)-piperazic acid derivatives to their D-(3*R*) isomers. Analysis of primary sequences of the E domains showed that all important residues[65] are conserved, suggesting that both domains should be catalytically active. However, svetamycins A–G were reported as containing β,γ-dihydroxydehydro-L-piperazic acid, rather than the predicted β,γ-dihydroxydehydro-D-piperazic acid, raising the possibility that this stereochemical assignment may be incorrect. This is certainly possible as the absolute configuration of the β,γ-dihydroxydehydropiperazic acid moiety was only tentatively assigned by comparison of experimental NMR data with DFT-calculated chemical shifts.[28]

Product stereochemistry is selected for by the downstream C domains in NRPS assembly lines and C domains clade into two evolutionary distinct families, LCL and DCL, depending on the stereochemistry of the substituent amino acids.[66-68] Phylogenetic analysis using NaPDoS[69] indicated that only SveA2 C1 and C3 were LCL family C domains (Fig. S75†). The remaining C domains, including SveA2 C2, clade within the DCL family (presumably a DCL domain is required in SveA1 to accommodate the α-methyl group of α-methyl-L-serine). While the isomerisation of piperazic acid residues is, for the most part, consistent with the presence of E domains (see polyoxypeptin and aurantimycin[46, 70]), a similar stereochemical contradiction is observed in the *ktz* BGC, with kutznerides 1–4 reported as containing L-piperazic acid, despite the presence of a downstream E domain in the first module of KtzH, again raising doubts about the original stereochemical assignments. The absolute configuration of the piperazic acid moiety in the kutznerides was also only tentatively assigned based on the GC retention time of the ornithine liberated following acid hydrolysis and catalytic hydrogenation of the natural products.[71] This revealed a mixture of L- and D-ornithine, suggesting partial racemisation of the sample had occurred during hydrolysis.

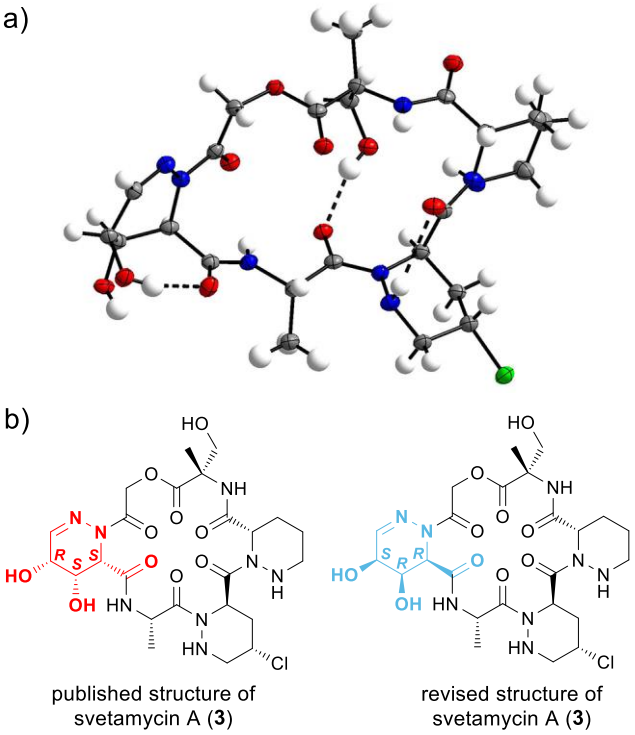


Fig. 3 a) Single-crystal X-ray structure of svetamycin A (**3**). b) Comparison of the previously published[28] and revised structures of svetamycin A (**3**).

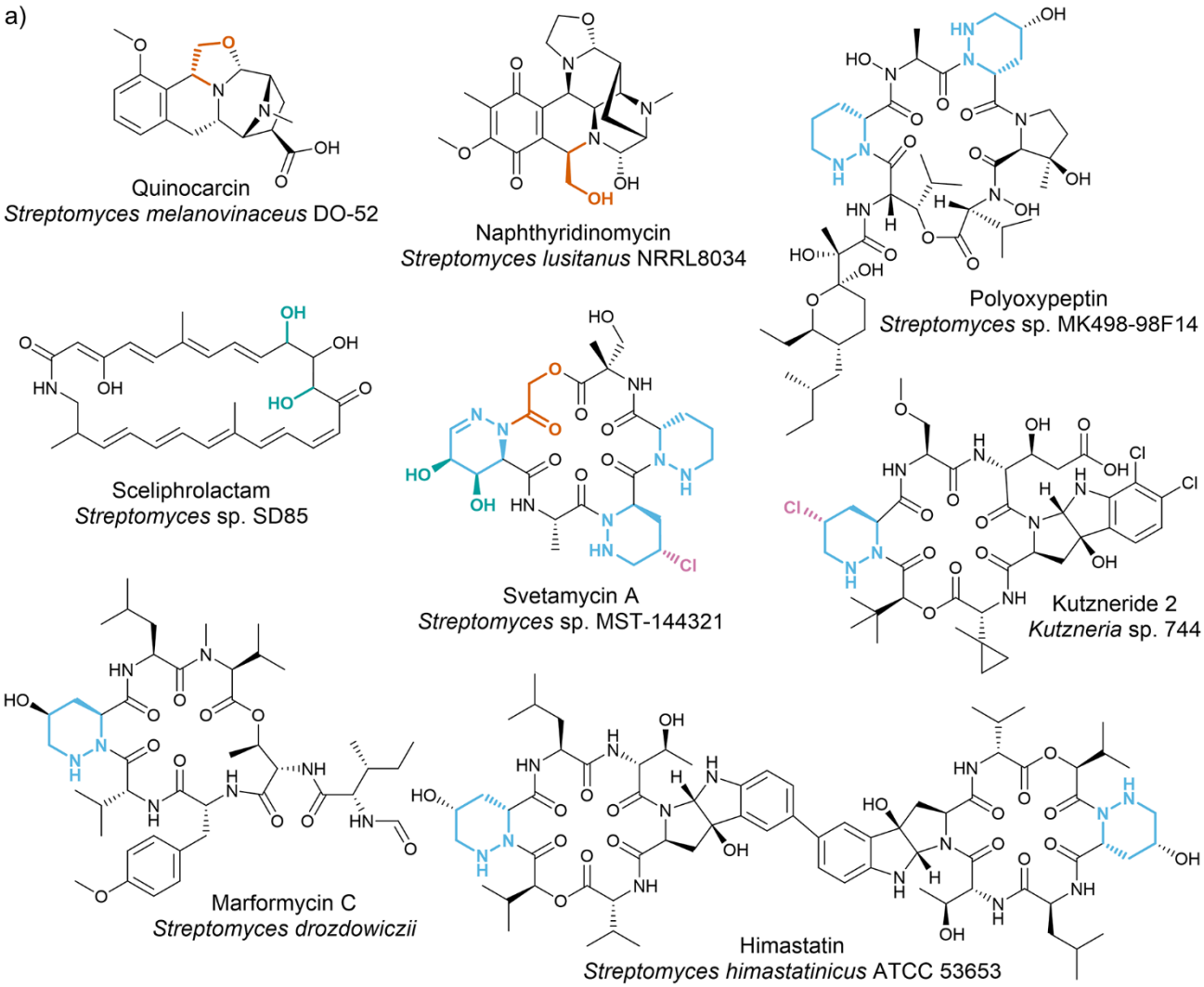
Given that the experimental evidence supporting the published stereochemical assignments of the svetamycins is tenuous and contradicts our biosynthetic hypothesis, we sought to solve the absolute configuration of **3** unequivocally. Suitable crystals were obtained by slow evaporation of a solution of **3** in 1:1 MeOH/MeCN and the absolute stereostructure was determined by single-crystal X-ray diffraction analysis (see ESI† and Table S29†). The crystal structure (Fig. 3a), which has been deposited with the Cambridge Crystallographic Data Centre (CCDC 2087152), confirmed that the three stereocentres present in the β,γ-dihydroxydehydropiperazic acid residue were indeed opposite in configuration to those reported in the published structure of **3**. This provides clear evidence that the associated epimerase domain is catalytically active, thus confirming our original biosynthetic hypothesis. As our NMR data for **3** are essientally identical to the published NMR data,[28] we propose the structure of **3**, and the structures of the biosynthetically related svetamycins B–G, should be revised to incorporate a (β*R*,γ*S*)-dihydroxydehydro-D-piperazic acid residue.

SveO1-3 are encoded by a three-gene cassette homologous to SceCDE from the sceliphrolactam BGC.[51] SceD and SceE encode two cytochrome P450s hypothesised to hydroxylate the polyene ring at the C-10 and C-12 positions. SceC encodes ferrodoxin, required for electron transfer to the cytochrome. Based on this relationship, we hypothesise that the P450s SveO1 and SveO3 and the ferredoxin SveO2 are required for the post assembly-line dihydroxylation of the dehydro-D-piperazic acid residue. Finally, the peptide chain is cyclised by the TE domain of module 6. Interestingly, *sve* also contains an analogue of DsaJ, a peptide cyclase required for the cyclisation of hexapeptides, such as desotamide[72] and ullungmycin,[73] and octapeptides, such as surugamide.[74-76] The presence of both a thioesterase domain and a peptide cyclase is apparently redundant and requires further investigation.

Svetamycin biosynthesis is a fascinating mosaic, sharing parts with piperazic and glycolic acid-incorporating NRPSs[46, 50, 52, 70] and more distantly related macrolactam[51] and cyclic peptide[73, 77, 78] BGCs (Fig. 4). It is notable that MST-144321 contains at least one macrolactam BGC in its genome, in addition to fragments of a naphthyridinomycin BGC (Table S24†). The mosaic structures of BGCs are a result of subclustering, whereby the enzymes responsible for the biosynthesis of specific functional moieties are conserved between otherwise distinct gene clusters. Elegant examples of subclustering have been described previously, such as the conservation of aminocoumarin, pyrrole and deoxysugar biosynthesis between clorobiocin, coumermycin and their relatives[79-81] and the intricate relationships between simocyclinone, rubradirin, everninomycin and polyketomycin.[82, 83] The development of new tools and databases have streamlined the process of identifying these relationships.[47, 48, 84] Nevertheless, *sve* is striking, both in the diversity of biosynthetic subclusters employed and its completeness. The *sve* BGC has the potential to become an important model for understanding how distinct biosynthetic parts are incorporated into an evolving BGC.

The borregomycin-anthrabenzoxocinone (*bab*) supercluster

The biosynthetic pathways of borregomycins and ABXs have been described previously.[29, 33, 85] As such, the genes responsible for their production were readily identified from the antiSMASH[40] analysis (Table S22†). Interestingly, the genes encoding the biosynthesis of the borregomycins and ABXs are collocated in the genome of MST-144321, encoded by a single supercluster, which we named the *bab* BGC (Fig. 5 and Table S22†). In the *bab* BGC, the genes responsible for the biosynthesis of ABXs are organised as previously described.[33] However, the genes encoding borregomycin biosynthesis are split across the flanks of the BGC, with the ABX biosynthetic genes apparently 'sandwiched' in between. The genes responsible for the biosynthesis of chromopyrrolic acid, the first intermediate in borregomycins biosynthesis, are situated downstream of the cluster. Meanwhile, genes encoding the FAD-binding protein responsible for oxidative decarboxylation of the pyrrole ring (*babO2*) are found upstream, along with genes encoding the tailoring methyltransferases (*babM1* and *babM2*) and the cytochrome P450s (*babO1* and *babO3*) required for the biosynthesis of mature borregomycins. With regards to the biosynthesis of borregomycin E, additional enzymes are required to dismantle the pyrrole ring (Fig. S78†). Chilczuk et al. recently published the BGC of the tjipanazoles (*tjp*), indolocarbazoles that also lack a pyrrole ring, from the cyanobacterium *Fischerella ambigua* (Näg.) Gomont 108b and predicted two putative biosynthetic pathways.[86] However, comparison of the *bab* and *tjp* BGCs revealed that only the

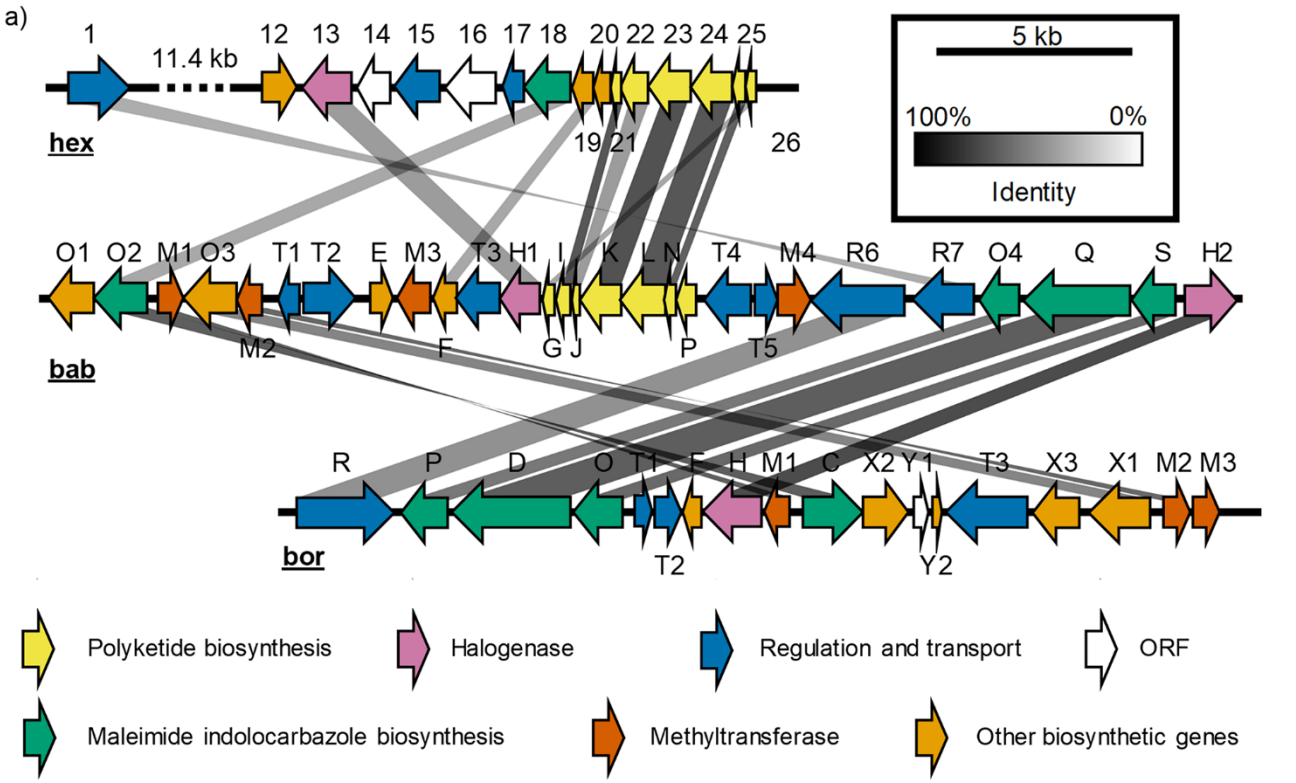


sve	D	E	F	G1	G2	G3	G4	H	R2	P1	P2	P3	P4	I	J	K	A1	A2	O1	O2	O3	L	M	N	T2	Q	R3
ktz																											
ply																											
hmt																											
mfn																											
qcn																											
nap																											
sce																											

Fig. 4 a) Structural features of common biosynthetic origin of svetamycin and related compounds; b) the conservation of biosynthetic subclusters responsible for the highlighted features of svetamycin (*sve*), kutzneride (*ktz*), polyoxypeptin (*ply*) himastatin (*hmt*), marformycin (*mfn*), quinocarcin (*qcn*), naphthyridinomycin (*nap*) and sceliphrolactam (*sce*).

genes required for chromopyrrolic acid were homologous (Fig. S79†), with the exception of the Tjp4 methyltransferase that shows homology to BabM3. Two proteins predicted to be involved in the cleavage of the pyrrole ring, Tjp3 and Tjp6, have no homologues within the *bab* BGC, or for that matter anywhere else in the genome of MST-144321 (Table S25†). Therefore, it seems likely that the enzymes required for the removal of the pyrrole ring have evolved convergently.

To our knowledge, this is the first supercluster to be described exhibiting this intertwined 'sandwich' architecture. Therefore, the *bab* supercluster has significant potential in helping to explain the mechanisms underpinning supercluster formation. In order to explore this idea further, we sought homologous genes involved in the biosynthesis of both aromatic polyketides and indolotryptolines/indolocarbazoles. Initial BLAST searches indicated that several genes (*babM1*, *babO3* and *babM2*) were present in both *bor* and *abx* BGCs, however experimental



b)

bab	O2	M1	O3	M2	R4	T2	E	M3	F	T3	H1	G	I	J	K	L	N	P	T4	R5	M4	R6	R7	O4	Q	S	H2	U
*abx**		■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■							
abx								■				■	■	■	■	■	■	■		■	■		■					
hex	■								■		■	■	■	■	■	■	■							■					
bor	■	■	■	■																		■		■	■	■	■	
reb	■																					■		■	■	■	■	■
sta	■																					■		■	■	■		

c)

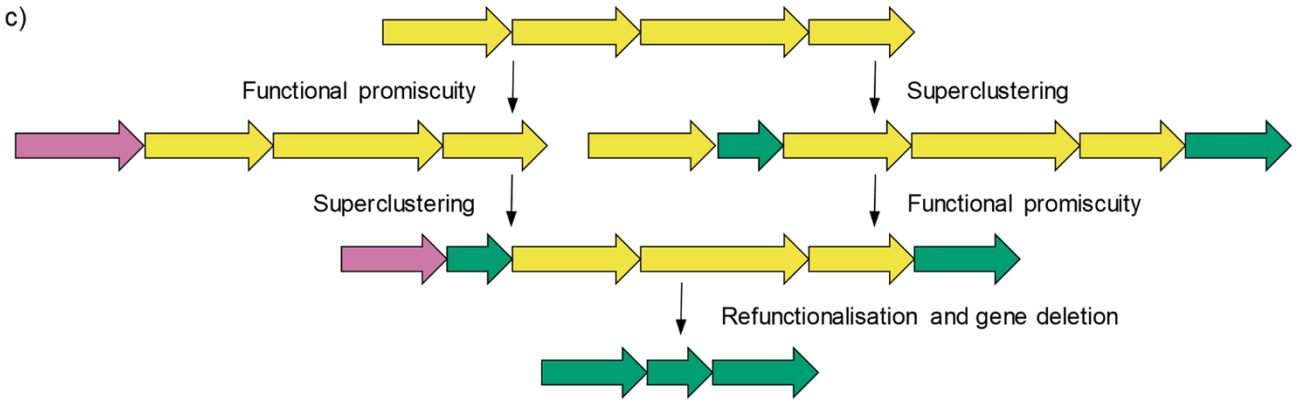


Fig. 5 a) Cluster comparison of the borregomycin-anthrabenzoxocinone (*bab*) super biosynthetic gene cluster (BGC) from *Streptomyces* sp. MST-144321 and related borregomycin (*bor*) and hexaricin (*hex*) BGCs; b) a simplified table of the cblaster[48] output demonstrating the 'sandwich' architecture of the supercluster, showing homology between the *bab* supercluster, the *hex* and *bor* BGCs and the BGCs for ABXs (from *Streptomyces* sp. MA6657 (*abx**) and *Streptomyces* sp. FJS31-2 (*abx*)), rebeccamycin (*reb*) and staurosporine (*sta*); c) potential models for the evolution of the indolotryptoline and indolocarbazole clusters from an ancestral supercluster.

evidence demonstrates that these homologues are not involved in the biosynthesis of ABXs. Knocking out the homologues *abxL, abxJ* and *abxQ* from a heterologously expressed cosmid had no effect on the production of ABXs.[33] It is likely given the related architectures of these BGCs that the remaining genes are present in the host organism, but were not present on the cosmid used in these experiments. Using cblaster,[48] we were able to identify 57 BGCs from the MIBiG[47] database (October 19, 2019) that contained at least 3 homologous genes with >30% identity and >50% coverage to *bab* proteins (Table S23†). Using this approach, we were able to find a single gene, *babO2,* which had homologues in both groups with a high level of consistency.

A phylogenetic tree of BabO2 homologues was constructed using RAxML[87] (100 replicates). BabO2 homologues form two major clades representing those associated with indolotryptoline/indolocarbazole BGCs (100% bootstrap support) and those associated with type II PKS BGCs (99% bootstrap support). Importantly, BabO2 clades with the indolotryptoline/indolocarbazole clade and was the first to diverge within this clade. This order of divergence indicates a specific series of events that lead to current diversity (Fig. S77†). We hypothesise that the ancestral BabO2 protein likely had a role in the biosynthesis of aromatic polyketides. Presumably, due to innate promiscuity, the ancestral protein acted in a bifunctional manner, allowing it to be neofunctionalised to its role in indolocarbazole biosynthesis. Loss of the T2PKS portion of the gene clusters would ultimately result in standalone indolocarbazole producing BGCs.

BabO2 is a FAD-binding oxygenase and its homologues are involved in a variety of oxidative reactions (Fig. S76†). In indolocarbazole biosynthesis, BabO2 homologues are involved in the oxidative carboxylation of the dicarboxypyrrole moiety of chromopyrrolic acid intermediates. Two well-characterised members of this group include RebC and StaC, from the biosynthetic pathways of rebeccamycin and staurosporine, respectively.[88-93] Both enzymes function in an FAD-dependent manner and bind the same substrate (7-carboxy-K252c), however small mutations in their primary sequence have altered the tautomeric specificity of the active site ultimately resulting in distinctive mono- or di-oxo-pyrroline products.[93] When associated with type II PKSs, BabO2 homologues take on a variety of roles. Homologues have been proposed as mono-, di- or tri-hydroxylases of a number of aromatic polyketides.[94-98] TcmG, for example, catalyses a triple hydroxylation of tetracenomycin A2 through a monooxygenase-dioxygenase mechanism.[99] An entirely different function is demonstrated by ForX, which catalyses a Baeyer–Villiger ring expansion that is required to convert the fasamycins to the formicamycin scaffold.[100] The diversity of functions catalysed by this family of FAD-binding proteins indicates a high degree of evolutionary plasticity that may have played a role in supercluster formation. Although BabO2 is required for the biosynthesis of the borregomycins, it is unclear what role, if any, BabO2 plays in the biosynthesis of the ABXs. We are currently investigating whether BabO2 maintains the functional promiscuity hypothesised for the ancestral protein.

The concept of biosynthetic superclusters is underexplored, but of increasing interest. The current model for the evolution of these superclusters relies on the two interdependent concepts of coregulation and synergistic activity. A well-characterised example of coregulation can be observed in strains of *Aspergillus fumigatus*, where genes for the biosynthesis of fumagillin and pseurotin are encoded by a single intertwined supercluster coordinated by the regulator FapR.[101] The coregulation of these two compounds hints towards their synergistic function, although the relationship has yet to be elucidated. A number of examples of synergistic activity have been described from *Streptomyces* spp. The archetypical example is the cephamycin-clavulanic acid supercluster, which is conserved across a number of *Streptomyces* spp.[102] Cephamycin is a beta-lactam antibiotic and is therefore susceptible to the activity of exogenous beta-lactamases. The cometabolite, clavulanic acid, is a beta-lactamase inhibitor, therefore enhancing the efficacy of cephamycin against organisms that produce beta-lactamases. Another classic example of synergistic antibiotics encoded by biosynthetic superclusters are the streptogramins, which synergistically bind to the 50S ribosomal subunit and inhibit protein synthesis.[103-106] Other examples involve the coordinated targeting of pathways,[107] as demonstrated by the Ras/TOR pathway targeting rapamycin-actinoplanic acid supercluster from *Streptomyces rapamycinicus* ATCC 29253,[108] or the cooperative binding of the 50S ribosome by lankamycin and lankacidin, encoded as a supercluster in *Streptomyces rochei* 7434AN4.[109, 110]

Our investigations into the regulation of the *bab* cluster are still ongoing and evidence of synergistic bioactivity between the borregomycins and ABXs has yet to be established. Our analysis, however, points towards a third aspect in the evolution of superclusters, namely collaborative biosynthesis. This has important implications for the evolution of natural product biosynthesis as in order for selection to occur, genes must not only be encoded, but expressed. Genes that are expressed in exclusive conditions are unable to interact and as such the neofunctionalisation of enzymes to new pathways cannot be selected for. Given that superclusters tend to be coregulated, genes in clustered pathways share a common temporality and therefore present an important opportunity for evolution of biosynthetic pathways to occur.

Conclusions

Detailed characterisation of genomes and metabolomes is crucial to developing a holistic understanding of small molecule biosynthesis in bacteria. Herein, we have uncovered three discrete families of chlorinated metabolites produced by the actinobacterium *Streptomyces* sp. MST-144321. By sequencing

the genome of the organism, we have not only linked each metabolite to a BGC, but have also provided insights into the evolution of BGCs that otherwise would have been invisible. The discovery of the *sve* BGC is particularly interesting as it appears to be a mosaic of several distantly related BGCs. While the precise nature of the relationship between *sve* and the related clusters (whether ancestral or descendent) could not be determined, the characterisation of related and intermediate pathways should allow these relationships to be elucidated in the future. This is a question of both theoretical and practical importance. Understanding how Nature combines biosynthetic machineries is an important facet of evolutionary biology. This understanding can, in turn, inform combinatorial biosynthesis, which is a major goal of synthetic biology.[111-114]

Cursory inspection of published genomes[115] reveals that the phenomenon of superclustering is widespread in *Streptomyces* spp. However, the number of superclusters where more than one metabolite has been characterised is far fewer. The *bab* supercluster represents an important addition to this list and the continued characterisation of superclusters will provide more robust evidence towards understanding the mechanisms and selective pressures behind the formation of superclusters. It has been shown that the products of superclusters may sometimes show complementary bioactivities.[102, 108] Given the broad range of bioactivities exhibited by the borregomycin analogues, it is difficult to comment on the hypothesis that the *bab* supercluster has evolved around complementary bioactivities. Rather, the presence of BabO2 homologues in both indolocarbazole and type II PKS BGCs points to cooperative biosynthesis as a potential driving force for superclustering, in this case (Fig. 5c). This is significant as it implicates superclusters as a potential nucleation point for the evolution of biosynthetic pathways, which may then diverge over time.

Conflicts of interest

There are no conflicts to declare

Acknowledgements

HRMS data were acquired by the Australian Proteome Analysis Facility, supported under the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS). This research was funded, in part, by the Australian Research Council (FT130100142, FT160100233, FT200100243), the Cooperative Research Centres Projects scheme (CRCPFIVE000119), the Australian Research Training Program (RTP), the University of Western Australia Scholarship for International Research Fees (SIRF) and Macquarie University (MQRTP scholarship to MTM).

Notes and references

1 G. W. Gribble, *Naturally occurring organohalogen compounds-a comprehensive update*, Springer, Wien/New York, 2010.
2 G. W. Gribble, *Mar. Drugs*, 2015, **13**, 4044-4136.
3 Dictionary of Natural Products 29.2, 2020 CRC Press, Taylor & Francis Group, Boca Raton, FL, USA.
4 A. R. Carroll, B. R. Copp, R. A. Davis, R. A. Keyzers and M. R. Prinsep, *Nat. Prod. Rep.*, 2020, **37**, 175-223.
5 M. McCormik, W. Stark, G. Pittenger, R. Pittenger and J. McGuire, *Antibiot. Ann.*, 1955, **1956**, 606-611.
6 F. Parenti, G. Beretta, M. Berti and V. Arioli, *J. Antibiot.*, 1978, **31**, 276-283.
7 J. C. Barna, D. H. Williams, D. J. Stone, T. C. Leung and D. M. Doddrell, *J. Am. Chem. Soc.*, 1984, **106**, 4895-4902.
8 M. C. Rebstock, H. M. Crooks, J. Controulis and Q. R. Bartz, *J. Am. Chem. Soc.*, 1949, **71**, 2458-2462.
9 B. M. Duggar, *Ann. N. Y. Acad. Sci.*, 1948, **51**, 177-181.
10 S. Omura, N. Imamura, R. Oiwa, H. Kuga, R. Iwata, R. Masuma and Y. Iwai, *J. Antibiot.*, 1986, **39**, 1407-1412.
11 W. Fenical, P. R. Jensen, M. A. Palladino, K. S. Lam, G. K. Lloyd and B. C. Potts, *Bioorg. Med. Chem.*, 2009, **17**, 2175-2180.
12 M. C. Walker and M. C. Chang, *Chem. Soc. Rev.*, 2014, **43**, 6527-6536.
13 M. F. Carvalho and R. S. Oliveira, *Crit. Rev. Biotechnol.*, 2017, **37**, 880-897.
14 L. Wang, X. Zhou, M. Fredimoses, S. Liao and Y. Liu, *RSC Adv.*, 2014, **4**, 57350-57376.
15 V. Agarwal, Z. D. Miles, J. M. Winter, A. S. Eustáquio, A. A. El Gamal and B. S. Moore, *Chem. Rev.*, 2017, **117**, 5619-5674.
16 C. Wagner, M. El Omari and G. M. König, *J. Nat. Prod.*, 2009, **72**, 540-553.
17 C. S. Neumann, D. G. Fujimori and C. T. Walsh, *Chem. Biol.*, 2008, **15**, 99-109.
18 B. R. Menon, D. Richmond and N. Menon, *Catal. Rev.*, 2020, doi.org/10.1080/01614940.01612020.01823788.
19 C. M. Harris, R. Kannan, H. Kopecka and T. M. Harris, *J. Am. Chem. Soc.*, 1985, **107**, 6652-6658.
20 J. E. Smyth, N. M. Butler and P. A. Keller, *Nat. Prod. Rep.*, 2015, **32**, 1562-1583.
21 M. Groll, R. Huber and B. C. Potts, *J. Am. Chem. Soc.*, 2006, **128**, 5136-5141.
22 A. J. Kale, R. P. McGlinchey, A. Lechner and B. S. Moore, *ACS Chem. Biol.*, 2011, **6**, 1257-1264.
23 G. Gerebtzoff, X. Li-Blatter, H. Fischer, A. Frentzel and A. Seelig, *ChemBioChem*, 2004, **5**, 676-684.
24 U. Bickmeyer, M. Heine, I. Podbielski, D. Münd, M. Köck and P. Karuso, *Biochem. Biophys. Res. Commun.*, 2010, **402**, 489-494.
25 K. Ragini, A. M. Piggott and P. Karuso, *Mar. Drugs*, 2019, **17**, 683.
26 M. T. Morshed, D. Vuong, A. Crombie, A. E. Lacey, P. Karuso, E. Lacey and A. M. Piggott, *Org. Biomol. Chem.*, 2018, **16**, 3038-3051.
27 K. Ragini, J. Fromont, A. M. Piggott and P. Karuso, *J. Nat. Prod.*, 2017, **80**, 215-219.
28 D. Dardić, G. Lauro, G. Bifulco, P. Laboudie, P. Sakhaii, A. Bauer, A. Vilcinskas, P. E. Hammann and A. Plaza, *J. Org. Chem*, 2017, **82**, 6032-6043.
29 F.-Y. Chang and S. F. Brady, *Proc. Natl. Acad. Sci. USA*, 2013, **110**, 2478-2483.
30 Y. T. Lam, O. Hensens, G. Helms, D. Williams Jr, M. Nallin, J. Smith, S. Gartner, L. H. Rodriguez and S. Stevens-Miles, *Tetrahedron Lett.*, 1995, **36**, 2013-2016.
31 K. B. Herath, H. Jayasuriya, Z. Guan, M. Schulman, C. Ruby, N. Sharma, K. MacNaul, J. G. Menke, S. Kodali, A. Galgoci, J. Wang and S. B. Singh, *J. Nat. Prod.*, 2005, **68**, 1437-1440.

32 Y. Lü, M. Shao, Y. Wang, S. Qian, M. Wang, Y. Wang, X. Li, Y. Bao, C. Deng and C. Yue, *Molecules*, 2017, **22**, 251.
33 X. Mei, X. Yan, H. Zhang, M. Yu, G. Shen, L. Zhou, Z. Deng, C. Lei and X. Qu, *ACS Chem. Biol.*, 2018, **13**, 200-206.
34 A. M. Piggott and P. Karuso, *Comb. Chem. High Throughput Screen.*, 2004, **7**, 607-630.
35 T. Tamaoki and H. Nakano, *Nat. Biotechnol.*, 1990, **8**, 732-735.
36 Y. Yamashita, N. Fujii, C. Murakata, T. Ashizawa, M. Okabe and H. Nakano, *Biochemistry*, 1992, **31**, 12069-12075.
37 T. Kimura, S. Kanagaki, Y. Matsui, M. Imoto, T. Watanabe and M. Shibasaki, *Org. Lett.*, 2012, **14**, 4418-4421.
38 A. Hörmann, B. Chaudhuri and H. Fretz, *Bioorg. Med. Chem.*, 2001, **9**, 917-921.
39 H. Nakano and S. Ōmura, *J. Antibiot.*, 2009, **62**, 17-26.
40 K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S. Y. Lee, M. H. Medema and T. Weber, *Nucleic Acids Res.*, 2019, **47**, W81-W87.
41 M. H. Medema, K. Blin, P. Cimermancic, V. de Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano and R. Breitling, *Nucleic Acids Res.*, 2011, **39**, W339–W346.
42 http://hmmer.org.
43 J. Ma, Z. Wang, H. Huang, M. Luo, D. Zuo, B. Wang, A. Sun, Y. Q. Cheng, C. Zhang and J. Ju, *Angew. Chem. Int. Ed.*, 2011, **50**, 7797-7802.
44 D. G. Fujimori, S. Hrvatin, C. S. Neumann, M. Strieker, M. A. Marahiel and C. T. Walsh, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 16498-16503.
45 J. Liu, B. Wang, H. Li, Y. Xie, Q. Li, X. Qin, X. Zhang and J. Ju, *Org. Lett.*, 2015, **17**, 1509-1512.
46 Y. Du, Y. Wang, T. Huang, M. Tao, Z. Deng and S. Lin, *BMC Microbiol.*, 2014, **14**, 30.
47 M. H. Medema, R. Kottmann, P. Yilmaz, M. Cummings, J. B. Biggins, K. Blin, I. De Bruijn, Y. H. Chooi, J. Claesen and R. C. Coates, *Nat. Chem. Biol.*, 2015, **11**, 625-631.
48 C. L. M. Gilchrist, T. J. Booth and Y.-H. H. Chooi, *bioRxiv*, 2020, doi.org/10.1101/2020.1111.1108.370601.
49 C. Wu, Z. Shang, C. Lemetre, M. A. Ternei and S. F. Brady, *J. Am. Chem. Soc.*, 2019, **141**, 3910-3919.
50 J. Y. Pu, C. Peng, M. C. Tang, Y. Zhang, J. P. Guo, L. Q. Song, Q. Hua and G. L. Tang, *Org. Lett.*, 2013, **15**, 3674-3677.
51 D. C. Oh, M. Poulsen, C. R. Currie and J. Clardy, *Org. Lett.*, 2011, **13**, 752-755.
52 T. Hiratsuka, K. Koketsu, A. Minami, S. Kaneko, C. Yamazaki, K. Watanabe, H. Oguri and H. Oikawa, *Chem. Biol.*, 2013, **20**, 1523-1535.
53 G. Zhang, H. Zhang, S. Li, J. Xiao, G. Zhang, Y. Zhu, S. Niu, J. Ju and C. Zhang, *Appl. Environ. Microbiol.*, 2012, **78**, 2393-2401.
54 A. Muliandi, Y. Katsuyama, K. Sone, M. Izumikawa, T. Moriya, J. Hashimoto, I. Kozone, M. Takagi, K. Shin-ya and Y. Ohnishi, *Chem. Biol.*, 2014, **21**, 923-934.
55 J. Shi, Y. J. Zeng, B. Zhang, F. L. Shao, Y. C. Chen, X. Xu, Y. Sun, Q. Xu, R. X. Tan and H. M. Ge, *Chem. Sci.*, 2019, **10**, 3042-3048.
56 D. R. Marous, E. P. Lloyd, A. R. Buller, K. A. Moshos, T. L. Grove, A. J. Blaszczyk, S. J. Booker and C. A. Townsend, *Proc. Natl. Acad. Sci. USA*, 2015, **112**, 10354-10358.
57 W. Lu, N. Roongsawang and T. Mahmud, *Chem. Biol.*, 2011, **18**, 425-431.
58 C. Peng, J. Y. Pu, L. Q. Song, X. H. Jian, M. C. Tang and G. L. Tang, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 8540-8545.
59 C. S. Sit, A. C. Ruzzini, E. B. Van Arnam, T. R. Ramadhar, C. R. Currie and J. Clardy, *Proc. Natl. Acad. Sci. USA,*, 2015, **112**, 13150-13154.
60 K. D. Morgan, R. J. Andersen and K. S. Ryan, *Nat. Prod. Rep*, 2019, **36**, 1628-1653.
61 C. S. Neumann, W. Jiang, J. R. Heemstra Jr, E. A. Gontang, R. Kolter and C. T. Walsh, *ChemBioChem*, 2012, **13**, 972-976.
62 Y. L. Du, H. Y. He, M. A. Higgins and K. S. Ryan, *Nat. Chem. Biol.*, 2017, **13**, 836-838.
63 W. Jiang, J. R. Heemstra Jr, R. R. Forseth, C. S. Neumann, S. Manaviazar, F. C. Schroeder, K. J. Hale and C. T. Walsh, *Biochemistry*, 2011, **50**, 6063-6072.
64 D. H. Haft, J. D. Selengut and O. White, *Nucleic Acids Res.*, 2003, **31**, 371-373.
65 T. Stachelhaus and C. T. Walsh, *Biochemistry*, 2000, **39**, 5775-5787.
66 S. L. Clugston, S. A. Sieber, M. A. Marahiel and C. T. Walsh, *Biochemistry*, 2003, **42**, 12095-12104.
67 C. Rausch, I. Hoof, T. Weber, W. Wohlleben and D. H. Huson, *BMC Evol. Biol.*, 2007, **7**, 78.
68 S. Meyer, J.-C. Kehr, A. Mainz, D. Dehm, D. Petras, R. D. Süssmuth and E. Dittmann, *Cell Chem. Biol.*, 2016, **23**, 462-471.
69 N. Ziemert, S. Podell, K. Penn, J. H. Badger, E. Allen and P. R. Jensen, *PLoS One*, 2012, **7**, e34064.
70 H. Zhao, L. Wang, D. Wan, J. Qi, R. Gong, Z. Deng and W. Chen, *Microb. Cell Fact.*, 2016, **15**, 160.
71 A. Broberg, A. Menkis and R. Vasiliauskas, *J. Nat. Prod.*, 2006, **69**, 97-102.
72 W. Ding, Y. Dong, J. Ju and Q. Li, *Appl. Microbiol. Biotechnol.*, 2020, **104**, 2603-2610.
73 S. Son, Y.-S. Hong, M. Jang, K. T. Heo, B. Lee, J.-P. Jang, J.-W. Kim, I.-J. Ryoo, W.-G. Kim and S.-K. Ko, *J. Nat. Prod.*, 2017, **80**, 3025-3031.
74 T. Kuranaga, K. Matsuda, A. Sano, M. Kobayashi, A. Ninomiya, K. Takada, S. Matsunaga and T. Wakimoto, *Angew. Chem. Int. Ed.*, 2018, **130**, 9591-9595.
75 K. Matsuda, M. Kobayashi, T. Kuranaga, K. Takada, H. Ikeda, S. Matsunaga and T. Wakimoto, *Org. Biomol. Chem.*, 2019, **17**, 1058-1061.
76 Y. Zhou, X. Lin, C. Xu, Y. Shen, S.-P. Wang, H. Liao, L. Li, H. Deng and H.-W. Lin, *Cell Chem. Biol.*, 2019, **26**, 737-744.
77 Y. Song, Q. Li, X. Liu, Y. Chen, Y. Zhang, A. Sun, W. Zhang, J. Zhang and J. Ju, *J. Nat. Prod.*, 2014, **77**, 1937-1941.
78 A. Ninomiya, Y. Katsuyama, T. Kuranaga, M. Miyazaki, Y. Nogi, S. Okada, T. Wakimoto, Y. Ohnishi, S. Matsunaga and K. Takada, *ChemBioChem*, 2016, **17**, 1709-1712.
79 A. S. Eustáquio, B. Gust, T. Luft, S.-M. Li, K. F. Chater and L. Heide, *Chem. Biol.*, 2003, **10**, 279-288.
80 Z. X. Wang, S. M. Li and L. Heide, *Antimicrob. Agents Chemother.*, 2000, **44**, 3040-3048.
81 F. Del Carratore, K. Zych, M. Cummings, E. Takano, M. H. Medema and R. Breitling, *Commun. Biol.*, 2019, **2**, 1-10.
82 A. Trefzer, S. Pelzer, J. Schimana, S. Stockert, C. Bihlmaier, H.-P. Fiedler, K. Welzel, A. Vente and A. Bechthold, *Antimicrob. Agents Chemother.*, 2002, **46**, 1174-1182.
83 M. H. Medema, P. Cimermancic, A. Sali, E. Takano and M. A. Fischbach, *PLoS Comput. Biol.*, 2014, **10**, e1004016.
84 J. C. Navarro-Muñoz, N. Selem-Mojica, M. W. Mullowney, S. A. Kautsar, J. H. Tryon, E. I. Parkinson, E. L. De Los Santos, M. Yeong, P. Cruz-Morales and S. Abubucker, *Nat. Chem. Biol.*, 2020, **16**, 60-68.
85 Y. Lü, C. Yue, M. Shao, S. Qian, N. Liu, Y. Bao, M. Wang, M. Liu, X. Li and Y. Wang, *Molecules*, 2016, **21**, 711.

86 T. Chilczuk, T. F. Schäberle, S. Vahdati, U. Mettal, M. El Omari, H. Enke, M. Wiese, G. M. König and T. H. J. Niedermeyer, *ChemBioChem*, 2020, **21**, 2170-2177.
87 A. Stamatakis, *Bioinformatics*, 2014, **30**, 1312-1313.
88 H. Onaka, S.-i. Taniguchi, Y. Igarashi and T. Furumai, *Biosci. Biotech. Bioch.*, 2003, **67**, 127-138.
89 A. R. Howard-Jones and C. T. Walsh, *J. Am. Chem. Soc.*, 2006, **128**, 12289-12298.
90 K. S. Ryan, A. R. Howard-Jones, M. J. Hamill, S. J. Elliott, C. T. Walsh and C. L. Drennan, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 15311-15316.
91 S. Asamizu, Y. Shiro, Y. Igarashi, S. Nagano and H. Onaka, *Biosci. Biotechnol. Biochem.*, 2011, **75**, 2184-2193.
92 K. Groom, A. Bhattacharya and D. L. Zechel, *ChemBioChem*, 2011, **12**, 396-400.
93 P. J. Goldman, K. S. Ryan, M. J. Hamill, A. R. Howard-Jones, C. T. Walsh, S. J. Elliott and C. L. Drennan, *Chem. Biol.*, 2012, **19**, 855-865.
94 E. R. Rafanan, C. R. Hutchinson and B. Shen, *Org. Lett.*, 2000, **2**, 3225-3227.
95 A. Li and J. Piel, *Chem. Biol.*, 2002, **9**, 1017-1026.
96 A. Ramos, F. Lombo, A. F. Brana, J. Rohr, C. Mendez and J. A. Salas, *Microbiology*, 2008, **154**, 781-788.
97 J. Tian, H. Chen, Z. Guo, N. Liu, J. Li, Y. Huang, W. Xiang and Y. Chen, *Appl. Microbiol. Biotechnol.*, 2016, **100**, 4189-4199.
98 T. e. Grocholski, K. Yamada, J. Sinkkonen, H. Tirkkonen, J. Niemi and M. Metsä-Ketelä, *ACS Chem. Biol.*, 2019, **14**, 850-856.
99 B. Shen and C. R. Hutchinson, *J. Biol. Chem.*, 1994, **269**, 30726-30733.
100 Z. Qin, R. Devine, T. J. Booth, E. H. Farrar, M. N. Grayson, M. I. Hutchings and B. Wilkinson, *Chem. Sci.*, 2020, **11**, 8125-8131.
101 P. Wiemann, C.-J. Guo, J. M. Palmer, R. Sekonyela, C. C. Wang and N. P. Keller, *Proc. Natl. Acad. Sci. USA*, 2013, **110**, 17065-17070.
102 J. M. Ward and J. E. Hodgson, *FEMS Microbiol. Lett.*, 1993, **110**, 239-242.
103 D. Vazquez, *Microbiology*, 1966, **42**, 93-106.
104 J. Noeske, J. Huang, N. B. Olivier, R. A. Giacobbe, M. Zambrowski and J. H. Cate, *Antimicrob. Agents Chemother.*, 2014, **58**, 5269-5279.
105 Y. Mast, T. Weber, M. Gölz, R. Ort-Winklbauer, A. Gondran, W. Wohlleben and E. Schinko, *Microb. Biotechnol.*, 2011, **4**, 192-206.
106 Y. Xie, B. Wang, J. Liu, J. Zhou, J. Ma, H. Huang and J. Ju, *ChemBioChem*, 2012, **13**, 2745-2757.
107 M. Alanjary and M. H. Medema, *J. Biol. Chem.*, 2018, **293**, 19996-19997.
108 P. Mrak, P. Krastel, P. P. Lukančič, J. Tao, D. Pistorius and C. M. Moore, *J. Biol. Chem.*, 2018, **293**, 19982-19995.
109 S. Mochizuki, K. Hiratsu, M. Suwa, T. Ishii, F. Sugino, K. Yamada and H. Kinashi, *Mol. Microbiol.*, 2003, **48**, 1501-1510.
110 M. J. Belousoff, T. Shapira, A. Bashan, E. Zimmerman, H. Rozenberg, K. Arakawa, H. Kinashi and A. Yonath, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 2717-2722.
111 D. Sardar and E. W. Schmidt, *Curr. Opin. Chem. Biol.*, 2016, **31**, 15-21.
112 R. H. Baltz, *ACS Synth. Biol.*, 2014, **3**, 748-758.
113 S. Chen, W. A. Kinney and S. Van Lanen, *World J. Microbiol. Biotechnol.*, 2017, **33**, 66.
114 A. A. Malico, M. A. Calzini, A. K. Gayen and G. J. Williams, *J. Ind. Microbiol. Biot.*, 2020, **47**, 675-702.
115 K. Blin, S. Shaw, S. A. Kautsar, M. H. Medema and T. Weber, *Nucleic Acids Res.*, 2021, **49**, D639–D643.