Accepted manuscript

Molecular BioSystems

New insights into polyene macrolide biosynthesis in *Couchioplanes caeruleus*

J. Sheehan, C. D. Murphy and P. Caffrey†

School of Biomolecular and Biomedical Science, University College Dublin, Belfield, Dublin 4, Ireland.

† Correspondence: patrick.caffrey@ucd.ie.

Published as: *Molecular BioSystems*, 2017, **13**, 866 – 873

DOI: 10.1039/c7mb00112f

Abstract

Couchioplanes caeruleus DSM43634 synthesises 67-121C, an aromatic heptaene macrolide that contains a mannosyl-mycosaminyl disaccharide. An improved draft genome sequence was used to obtain the biosynthetic gene cluster for this antifungal. Bioinformatic analysis of the polyketide synthase indicated that extension modules 7 and 8 contain A-type ketoreductase and dehydratase domains. These modules are therefore predicted to form *cis* double bonds. The deduced stereostructure of the 67-121C macrolactone is identical to that experimentally determined for the partricin subgroup of aromatic heptaenes. Some of these polyenes are *N*-methylated on the aminoacetophenone moiety. The *C. caeruleus* AceS protein was shown to methylate 4-aminoacetophenone and esters of 4-aminobenzoate, but not 4-aminobenzoate. This suggests that the substrate specificity of AceS prevents it from interfering with folate biosynthesis. The methyltransferase should be valuable for chemoenzymatic alkylation of compounds that contain aminobenzoyl moieties.

1. Introduction

Polyene macrolides were first identified as antifungal agents in the 1950s. Since that time, several of these compounds have been used as antibiotics, antiparasitic drugs and preservatives, and other possible applications have been proposed.[1, 2] Polyenes interact with ergosterol to disrupt membrane functions in sensitive cells. In fungal pathogens, this vulnerability is not easily overcome by resistance mechanisms.[3] However, polyene antifungals have serious side effects in humans and considerable effort has gone into reducing this toxicity.[4] In the period from 2000 to 2005, biosynthetic gene clusters for nystatin, pimaricin, amphotericin, candicidin and rimocidin were cloned and sequenced. Methods for genetic manipulation of the producer micro-organisms were

developed. These studies generated potentially valuable polyene analogues, and provided enzymes and information that will be useful for synthetic biology.[5] In addition, analysis of polyene biosynthetic enzymes helped to reveal protein sequence motifs that allow prediction of alcohol stereochemistry in complex polyketides.[6]

Aromatic heptaenes contain an aminoacetophenone moiety.[7] These polyenes fall into two groups that differ in the locations of two consecutive *cis* double bonds in the heptaene unit, and the presence or absence of a methyl branch at C40. Candicidin D (**1**) and partricin B (**2**) are representative members of each group (Fig 1; Fig S1). Some members of the partricin group, such as partricin A (**3**), are *N*-methylated on the 4-aminobenzoyl moiety.[8-13] All of the aromatic heptaenes that have been characterised are synthesised as complexes of several analogues that differ slightly in the polyol chain, as a result of variable processing of β-ketones during polyketide chain assembly.[14]

Candicidins, levorins and ascomycins are identical complexes synthesised by different producer micro-organisms.[15] In the other group, partricin B and vacidin A are identical, as are partricin A **3** and gedamycin, while perimycin has a similar macrolactone that lacks an exocyclic carboxyl group and is glycosylated with perosamine rather than mycosamine (Fig. S1).[16, 17]

Aromatic heptaenes are more potent than amphotericin B, the gold standard among antifungal antibiotics. Candicidin has been used in medicine,[15] and semisynthetic non-toxic derivatives of partricin A have been taken into clinical trials.[18] The 67-121 complex is of interest as it includes the only naturally occurring aromatic heptaene that is known to be modified with a disaccharide, 4-O-β-D-mannopyranosyl-D-mycosamine.[19]

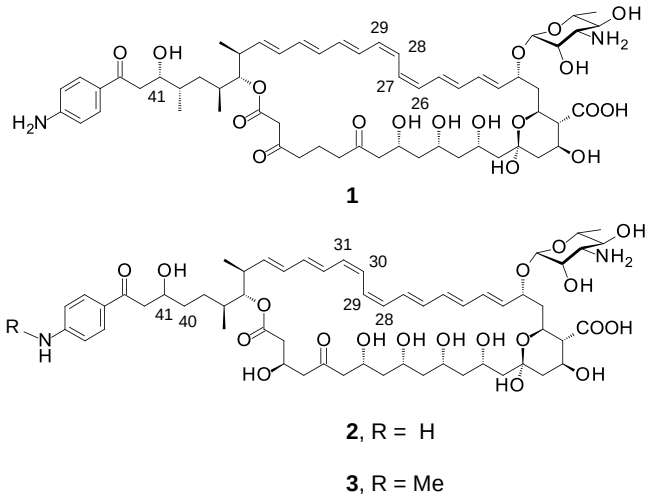


Fig. 1 Structures of candicidin D (**1,** synonyms, levorin A2, ascocin A2), partricin B (**2,** synonym vacidin A), and partricin A (**3**, synonym gedamycin).

67-121B (**4**), 67-121A (**5**) and 67-121C (**6**) (Fig. 2) are produced by *Couchioplanes caeruleus* subsp. *caeruleus* DSM 43634, formerly *Actinoplanes caeruleus*.[20] Only a gross planar structure has been experimentally determined for the 67-121 macrolactone. Analysis of separated components revealed that the extra mannosyl sugar made little apparent difference to antifungal activity. However, the methylated forms **5** and **6** were more protective than the unmethylated **4** in mouse models of systemic mycoses.[21] The 67-121 complex also had greater oral bioavailability than candicidin. Here we investigate 67-121 biosynthesis further.

The candicidin biosynthetic genes from *Streptomyces* sp FR008 have been extensively studied,[22] and homologous clusters have been identified in the genomes of many other micro-organisms.[23] However, to our knowledge no cluster has been completely sequenced for any member of the partricin group that has been chemically characterised. Here we used an improved draft genome sequence of *C. caeruleus* to complete the sequence of the 67-121 biosynthetic gene cluster. Prediction methods indicate that the macrolactone core has the same stereostructure as that of partricins. The *N*-methyltransferase that acts on the 4-aminoacetophenone moiety was characterised *in vivo* and *in vitro*. This work also reveals that there is no conflict between biosyntheses of folate and aromatic heptaenes, pathways of primary and secondary metabolism that process 4-aminobenzoyl precursors in different ways.

2. Materials and Methods

2.1 Bacterial strains

C. caeruleus DSM43634 was obtained from Leibniz Institute DSMZ. *Streptomyces cinnamoneus* DSM 40114 was used for isolation of trichomycins.[24] The *aceS* methyltransferase gene was expressed in the candicidin producer *Streptomyces albidoflavus* DSM40624 (formerly *Streptomyces griseus* IMRU3570). *Escherichia coli* TG1 was used as a host for the construction and propagation of recombinant plasmids. *E. coli* BL21(DE3) was used as host for pET28 expression plasmids.

2.2 DNA methods

Chromosomal DNA was isolated from *C. caeruleus* DSM43634 as described previously.[25] The genome was sequenced by Source Bioscience in Nottingham using a MiSeq instrument. CLC Bio software was used for sequence assembly. The draft genome sequence has been submitted to GenBank and has the accession number MEIA00000000.

PCR was carried out using Phusion® GC high fidelity polymerase in a TECHNE TC-3000 thermocycler. Plasmid minipreps or PCR products were purified using Qiagen kits. Templates were sent to Source Bioscience Ireland for dideoxy sequencing.

2.3 *Streptomyces* methods

For polyene production, *C. caeruleus* was grown initially in DSMZ GYM medium 65, and *S. albidoflavus* and *S. cinnamoneus* were grown on tryptic soy broth. These starter cultures were incubated for five days at 30 °C with agitation at 200 rpm. From these flasks, 5 ml volumes were used to inoculate 250 ml flasks containing 100 ml each polyene production medium (20 g fructose, 60 g dextrin, 30 g soya flour, 10 g $CaCO_3$, 50 g Amberlite XAD16 beads and 10 g glycerol per litre H_2O). Flasks also contained compression springs to disperse mycelia and increase aeration. Production cultures were incubated at 30 °C with shaking at 200 rpm for seven days. Polyenes were then extracted as described previously.[26]

Transformation of *S. albidoflavus* was carried out using methods adapted from Kieser and co-workers.[27]

2.4 Construction of expression plasmids for AceS

The 1 kb *aceS* gene was amplified with MetF3 and MetR primers for cloning into pIAGO, and with MetF4 and MetR for construction of pIAGO-*pegA* + *aceS* (Table 1). To construct pET28-MetN, the *aceS* gene was amplified with primers MetF2 and MetR, digested with Nde I and Hin dIII, and cloned into pET28 to give construct pET28-MetN. This encodes a methyltransferase with an N-terminal hexahistidine tag. The pET28-MetN plasmid was shown to have the correct insert by subsequent restriction mapping and sequencing. The construct was used to transform *E. coli* BL21 DE3.

2.5 Overproduction and purification of the AceS methyltransferase

To overproduce the AceS protein, *E. coli* BL21 DE3 pET28-MetN cells were grown at 30 °C in 2TY medium containing 50 μg kanamycin ml^{-1}. Flasks were shaken at 180 rpm. Once the OD_{600} had reached 0.6, expression was induced with 0.4 mM isopropylthiogalactoside (IPTG) and incubation was continued for a further 6 h. The cells from a total culture volume of 1 litre were sedimented at 4,500 x g for 15 min at 4 °C, and then re-suspended in 50 ml 10 mM NPI lysis buffer (300 mM NaCl, 50 mM sodium phosphate and 10 mM imidazole at pH 8.0) with DNase (1 mg/ml), 100 mM phenylmethanesulphonyl fluoride (PMSF). Cells were disrupted in a pre-chilled French press at 20,000 psi. The resulting lysate was centrifuged at 15,000 x g for 20 min. The soluble fraction was retained as a source of AceS protein. This was purified using a 1 ml Ni-NTA affinity column that had been equilibrated with 20 mM NPI equilibration buffer (300 mM NaCl, 50 mM sodium phosphate and 20 mM imidazole pH 8.0). The column was washed with 10 ml of the same buffer then the tagged AceS was eluted with elution buffer (300 mM NaCl, 50 mM sodium phosphate and 250 mM imidazole pH 8.0). Fractions containing the methyltransferase were identified by SDS-PAGE and pooled. Dialysis was used to exchange the imidazole buffer for 20 mM sodium phosphate, 150 mM NaCl (pH 7.0) containing 5 % (v/v) glycerol. Purified proteins were concentrated using Amicon® ultra 0.5 ml centrifugal filters.

SDS-polyacrylamide gel electrophoresis was carried out by the method of Laemmli,[28] with 12% separating gels and 4% stacking gels.

Table 1 PCR primers used for cloning *aceS*. Sequences in bold indicate restriction sites added to each end of the coding sequence.

Primer	Sequence 5′ to 3′	Construct
MetF2	gcaa**catatg**accatcggcacactcatcgacgc	pET28-MetN
MetF3	ctga**ggatcc**cgagggagcaagccatgaccatc	pIAGO-*aceS*
MetF4	ctga**aagctt**cgagggagcaagccatgaccatc	pIAGO-*pegA*+*aceS*
MetR	gatc**aagctt**gatggacgacgcgtcctgatcagac	All of the above

2. 6 *In vitro* methyltransferase assay

For assessment of enzymatic activity, the standard 1 ml reaction mixture contained 1 mM SAM, 1 mM aminobenzoyl substrate, 50 μg purified AceS methyltransferase, 20 mM sodium phosphate buffer (pH 7.5). The reaction mixture was incubated at 30 °C with agitation at 950 rpm. After 120 min, reaction mixtures were quenched by extraction twice with equal volumes of ethyl acetate. Extracts were dried using nitrogen, re-suspended in 60 μl ethyl acetate and analyzed by GC-MS. All reactions were performed in triplicate.

2. 7 Analytical methods

HPLC was carried out on a reverse phase C18 column (SB-C18, 4.6 x 150mm, 5μm) on a system with a ProStar 335 photodiode array detector. Solvent A was 0.05% (v/v) formic acid in water, solvent B was 0.05% (v/v) formic acid in methanol. Gradients of 50 – 100% B were run over 47 minutes at a flow rate of 1 ml min^{-1}.

LC-MS analysis was carried out in the Department of Chemistry at the University of Leicester. The column used was an Acquity UPLC BEH C18 (2.1 mm x 50 mm, 1.7 μm) coupled to a XEVO quadrupole time-of-flight mass spectrometer. For the LC solvent A was 0.1% (v/v) formic acid in water and solvent B was 0.1% (v/v) formic acid in acetonitrile.

GC-MS analysis was carried out using a 6890N network GC system, 7683B series injector and a 5973 inert mass selective detector, all by Agilent technologies. 1 μL volumes of samples were injected onto a HP5 MS column. The oven temperature was held at 70 °C for 3 min before being raised to 250 °C over 8 min with a run time of 19 min. The operating mass range was from 100-600 Da.

3. Results

3.1 Sequencing the 67-121 cluster

In our previous work, sequencing the *C. caeruleus* genome was carried out using an Illumina instrument. This gave 2,194 contigs with an average size of 3,696 bp.[25] The total length of DNA sequenced was 8,109,345 bp. This identified the gene for the PegA extending glycosyltransferase (GT) that catalyses attachment of the mannosyl sugar to the mycosaminyl residue of 67-121A.[25, 29] The *pegA* gene is flanked by insertion sequences and is not located within the main 67-121 cluster. Some of the late genes were also identified and the *aceDII* mycosamine synthase and *aceN* cytochrome P450 genes were shown to complement *amphDII* and *amphN* mutations in *Streptomyces nodosus* strains. In this project, sequencing was repeated using the MiSeq system.[30] The second sequence gave 8,360,800 bp as 665 contigs with an average length of 12, 787 bp. This assisted assembly of the complete 67-121 cluster, although many gaps remained. These were closed by amplifying missing regions by PCR and carrying out dideoxy sequencing. Manual annotation and AntiSmash analysis gave 18 clusters for natural products.[31]

3. 2 Predicting the stereostructure of 67-121C

The 67-121 cluster is 130,524 bp in length. The order of genes is similar to that in the candicidin cluster (Table S1). The polyketide synthase (PKS) modules are housed within six multienzyme polypeptides (Fig. S2). The last two of these, AceP5 and AceP6, contain two and three modules, respectively, whereas the corresponding modules in the candicidin PKS are organized as the tetramodular FscE and single-module FscF proteins.[22] Most of the KR domains are B-type and form 3D-3-hydroxyacyl-ACP thioester intermediates (Fig. S3). B-type KRs and DH domains together form *trans* double bonds. A-type KRs form a 3L-3-hydroxyacyl ACP thioester. The combination of an A-type KR with a DH occurs less frequently, but results in formation of a *cis* double bond (Fig. S4).[32, 33] Interestingly, extension modules 7 and 8 contain A-type KRs paired with DH domains. This indicates that *cis* double bonds are formed in cycles 7 and 8 of 67-121 polyketide biosynthesis (Fig. S2). Where a methylmalonyl extender unit is used, KR domains also determine the chirality at C-2 of the 2-methyl-3-ketoacyl-ACP thioester

(Fig. S5). The original KR stereospecificity motifs have been extended to allow prediction of methyl as well as hydroxyl stereochemistry (Table S2).[34, 35] The configurations of the remaining two methyl groups were predicted using these extended KR motifs. Where a 2-methyl-3-ketoacyl chain is fully reduced, the enoylreductase (ER) domain decides the final stereochemistry of the C-2 methyl branch (Fig. S6). The ER3 domain of the Ace PKS contains the diagnostic tyrosine characteristic of ERs that form (2*S*)-2-methyl branched intermediates.[36] On this basis the configuration of the methyl group imposed by this domain (at C-38) was assigned as shown in Fig. 2. The stereospecificity motifs in the ketoreductase (KR) and enoylreductase (ER) domains of the 67-121 PKS are shown in Fig. S7. The deduced structure of the macrolactone agrees with the gross planar structure determined in 1977,[19] with one exception. KR20 lacks an essential active site tyrosine and must be inactive. This would result in a ketone at C-5 rather than the previously proposed hydroxyl group. Re-analysis of heptaenes from *C. caeruleus* by LC-MS supported the predicted structure (see below).

The stereostructure predicted for the 67-121 macrolactone (Fig. 2) matches that experimentally determined for partricin, although the chirality of the hydroxyl group at C-41 in that class of polyene has yet to be assigned.[10-12] In the candicidin PKS, modules 8 and 9 contain A-type KRs and DH domains.[22] This agrees with the locations of the two *cis* double bonds in the stereostructure of candicidin D (Fig. 1).[8]

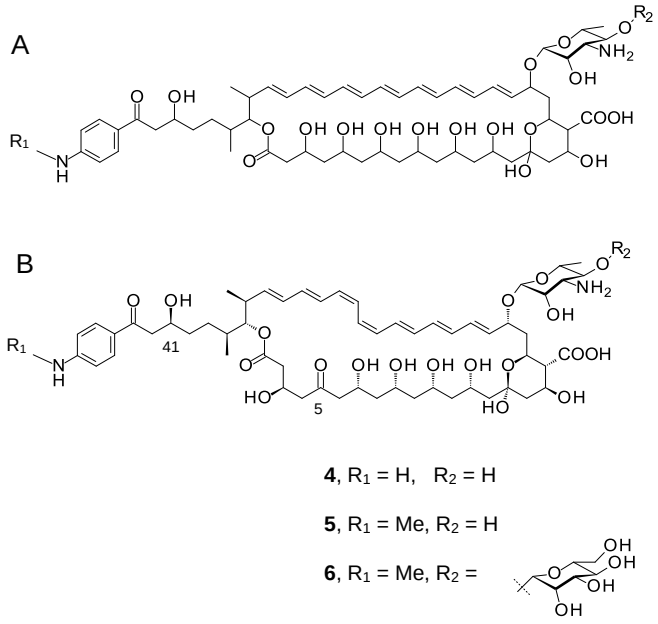


Fig. 2 Structures of 67-121 polyenes. A. Planar structure determined by Wright and co-workers in 1977.[19] B. Stereostructure deduced here from stereospecificity motifs in the polyketide synthase.

Candicidin clusters occur in many genome sequences and all appearances of the PKS appear to have the same stereospecificity motifs as the original FR008/candicidin PKS. One intriguing feature of the candicidin PKS is that the KR of module 2 is B1 type and is predicted to give a (2*R*, 3*R*)-2-methyl-3-hydroxyacyl-ACP.[1, 22] However, NMR analysis of candicidin D indicates that the alcohol at this position (C-41) has the opposite stereochemistry.[8] This anomaly might be explained by further work on the second modules of aromatic polyene PKSs.

3.3 Analysis of polyenes from *C. caeruleus*

C. caeruleus was found to produce heptaenes in yields of about 10 mg/litre. These extracts contained fewer polyene analogues than the candicidin complex, which contains about five components.[14] LC-MS analysis revealed that the main *C. caeruleus* heptaene had a mass appropriate for 67-121C ($[M + 2H]^{2+}/2$ = 645.3177; $[M + H]^{+}$ = 1289.6353)(Fig. S8). The abundance of the doubly protonated form indicates that the methylated amino group has a stable positive charge. Methylation could contribute to antifungal activity since positively charged biocides are more active. The second most abundant heptaene had a mass expected for 67-121A ($[M + H]^{+}$ = 1127.5916). These masses are consistent with forms of 67-121 polyenes with ketone rather than hydroxyl groups at C5.

The AceN cytochrome P450 oxidises a methyl branch at C-18 to a carboxyl group.[25] Inactivation of AceN homologues has led to production of polyenes lacking exocyclic carboxyl groups, which have reduced haemolytic activity. At present it is not possible to manipulate *C. caeruleus* genetically. Lambalot and Cane found that the

cytochrome P450 inhibitor 8-methoxypsoralen caused *Streptomyces venezuelae* to accumulate 10-deoxymethynolide rather than the hydroxylated products methymycin and neomethymycin.[37] *C. caeruleus* cultures were grown in the presence of 8-methoxypsoralen to investigate whether this would cause biosynthesis of 67-121 analogues in which methyl groups replace the free carboxyl groups. However, even low concentrations of 8-methoxypsoralen abolished polyene production.

3. 4 Expression of the *aceS* methyltransferase gene in *S. albidoflavus*

The AceS protein shows 50% sequence identity with known methyltransferases, and is the only plausible candidate for the enzyme that methylates the 4-aminobenzoyl moiety of 67-121 polyenes. We expressed the gene in *S. albidoflavus*, to see whether AceS acts on candicidin, which has an un-methylated aminoacetophenone group. The same heterologous expression approach previously revealed that the *C. caeruleus* PegA extending GT works weakly on candicidin.[29] We also investigated whether a combination of AceS and PegA might give N-methylated candicidins containing mannosyl-mycosaminyl disaccharides. The *aceS* gene was cloned into a pIAGO vector downstream from a strong *ermE* promoter. The pIAGO-*aceS* construct was used to transform *S. albidoflavus*. The methyltransferase gene was also cloned into the Hind III site of pIAGO-*pegAI* plasmid,[25] downstream from the *pegA* gene. Clones with the glycosyltransferase and methyltransferase genes in the same orientation were identified by PCR. The resulting pIAGO-*pegA*+*aceS* plasmid was also introduced into *S. albidoflavus*. Polyenes were extracted from *S. albidoflavus* strains transformed with pIAGO, pIAGO-*aceS* or pIAGO-*pegA*+*aceS*. LC-MS revealed at least five heptaenes in the extract from the control strain containing the empty vector (Fig. S9 and Fig. S10). These were identified as components of the candicidin complex from the masses of the major ions present (Table S3). The extract from the transformant containing the AceS methyltransferase gave a chromatogram that was almost identical to the control (Fig. S10B). The major ions had masses of unmethylated candicidins. A species was detected with a mass appropriate for methylated candicidin III, but this did not correlate with a major heptaene peak that was unique to cells containing AceS. The transformant with both PegA extending glycosyltransferase and AceS methyltransferase produced two new polyenes (Fig. S10). These had masses appropriate for mannosylated candicidins I and either II or III. No ions were detected that might represent forms that were both methylated and mannosylated.

These results indicate that expression of the *aceS* methyltransferase gene in *S. albidoflavus* does not result in efficient methylation of candicidins. It is uncertain whether this is because candicidins are poor substrates, or because polyene export out-competes methylation. These results confirm the previous findings that PegA acts on candicidins *in vivo*, to some extent.[29] It is now clear that 67-121A and candicidin differ in the locations of their two consecutive double bonds in the heptaene unit. The PegA GT shows some tolerance towards candicidin as an alternative acceptor.

3.5 *In vitro* studies on the methyltransferase

A form of AceS was overproduced in *E. coli* as a soluble protein with an N-terminal hexahistidine tag sequence (Fig. S11). The methyltransferase was purified and tested for activity in an assay system containing *S*-adenosylmethionine (SAM) and polyene acceptors. The acceptors used were candicidins and trichomycins (Fig. S1), which are unmethylated aromatic heptaenes that lack a methyl branch at C-40. Analysis of reaction mixtures by HPLC indicated no *in vitro* methylation of either substrate. Attempts to demonstrate the reverse reaction, *S*-adenosylhomocysteine (SAHC)-dependent de-methylation of 67-121C were also unsuccessful. These failures were thought to be due to low solubility of aromatic heptaenes in aqueous reaction mixtures. The enzyme was then tested on a range of low molecular weight compounds containing 4-aminobenzoyl groups (Fig. 3). Nine of these were assessed as potential surrogate substrates, and methylated products were detected quantitatively by GC-MS.

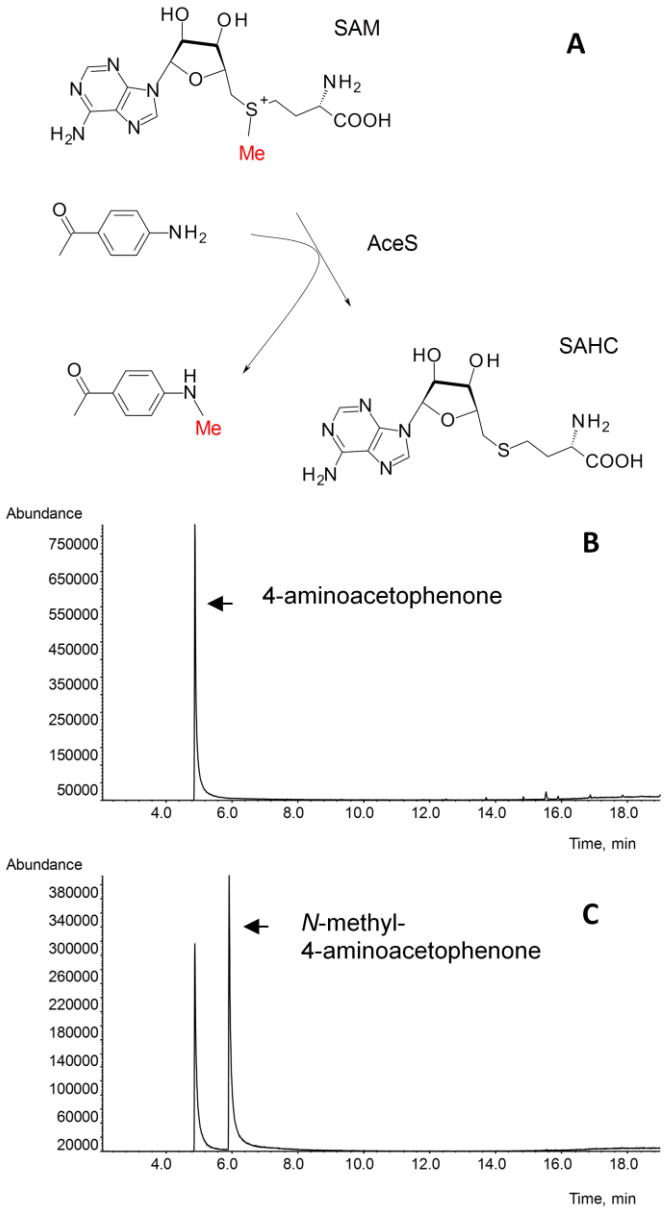


Fig. 3 *In vitro* assay for methyltransferase activity. A. 4-Aminoacetophenone was incubated with SAM and AceS then reaction mixtures were analysed by GC-MS. B. Chromatogram from analysis of 4-aminoacetophenone. C. Chromatogram from analysis of reaction products. N-methyl-4-aminoacetophenone was identified by comparison with a standard and by mass spectrometry.

Fig. 4 shows the compounds recognized by the methyltransferase and those that did not act as substrates. Table 2 gives the extent of methylation of each substrate, as determined by peak areas in total ion chromatograms. The enzyme methylated 4-aminoacetophenone (**7**), converting over 50% of the starting material, but not 4-aminobenzoic acid (**8**) or the related 4-aminosalicyclic acid (**9**), which has been used as an anti-tuberculosis drug. The methyltransferase showed modest activity (6% yield) towards 3-aminoacetophenone (**10**) but none towards 2-aminoacetophenone (**11**). This indicates a high degree of regioselectivity for the para position.

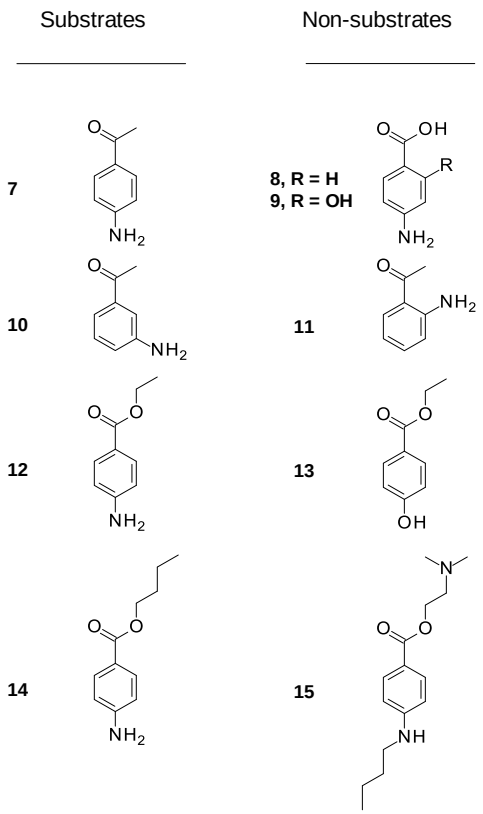


Fig. 4 Acceptors used in methyltransferase assays. **7**, 4-aminoacetophenone; **8**, 4-aminobenzoic acid; **9**, 4-aminosalicylic acid; **10**, 3-aminoacetophenone; **11**, 2-aminoacetophenone; **12**, 4-aminobenzoyl ethyl ester; **13**, 4-hydroxybenzoyl ethyl ester; **14**, 4-aminobenzoyl butyl ester; **15**, tetracaine.

Significant activity (57% yield) was observed with 4-aminobenzoate-ethyl ester (**12**). This compound, also known as benzocaine, is used as a component of a topical anaesthetic. While the AceS methyltransferase acted on **12**, it showed no activity towards 4-hydroxybenzoate-ethyl ester (**13**). The enzyme is therefore an *N*-methyltransferase only, not an *O*-methyltransferase. Almost 100% methylation occurred with 4-aminobenzoate-butyl-ester (**14**). The relative transformation of substrate to product increased with the length of the chain esterified to the carboxyl group. These results suggest that, at the earliest, AceS only acts after the 4-aminobenzoyl unit has been loaded onto the PKS and extended by one or more extension cycles.

Table 2. GC-MS results for methyltransferase assays with surrogate substrates.

Compound	RT (min)	*m/z*	RT product (min)	*m/z* product	% Yield
7	4.87	135	5.93	149	53
10	4.12	135	4.87	149	6
12	8.2	165	6.74	179	57
14	8.49	193	9.27	207	97

A number of N-dimethylases function in biosynthesis of dTDP-deoxyaminosugars. These include TylM1, DesVI, MegD3, and SpnS that function in biosynthesis of mycaminose, desosamine, megosamine and forosamine.[38-40] The glycopeptide N-methyltransferase MtfA also dimethylates peptide aglycones *in vitro*.[41] No AceS-catalysed dimethylation of any substrate was observed. The topical anaesthetic tetracaine (**15**) contains a 4-amino benzoyl moiety which is N-alkylated with a butyl group. The AceS methyltransferase did not modify the secondary amine of this compound.

4. Discussion

The sequence of the 67-121 PKS allows prediction of a stereostructure for the polyene macrolactone product. This indicates that 67-121 polyenes belong to the partricin group of aromatic heptaenes. In general, there is complete agreement between experimentally determined and predicted alcohol stereochemistry in complex polyketides.[42] The only exception known to us is the C-41 hydroxyl in candicidin D.

Some *cis* double bonds in polyketides result from *trans-cis* isomerisation as a late step.[43] This work indicates that with aromatic heptaenes, *cis* double bonds are introduced at an earlier stage, as a result of dehydration of a 3L-3-hydroxyacyl-ACP intermediates by the PKS modules. The aromatic heptaene PKSs provide more examples where pairing of A-type KRs with DH domains correlates with initial formation of *cis* double bonds. This strengthens the view that initial double bond geometry can be predicted from sequence motifs in KR domains.

This work is the first characterisation of a methyltransferase that functions in polyene biosynthesis. Previous studies isolated 67-121B (**4**), 67-121A (**5**) and 67-121C (**6**) in yields of 12%, 12% and 76% respectively.[21] This suggests that the last stages in the pathway are N-methylation of **4** to form **5**, then finally mannosylation of **5** to give **6**. Heterologous expression of *aceS* and *pegA* genes in *S. albidoflavus* did not provide convincing evidence for methylation of candicidins, but it did confirm that partial mannosylation of these unmethylated substrates can occur. This is of interest because studies on biosynthesis of two other disaccharide-modified polyenes revealed that extending GTs can show strict substrate specificity. In *Pseudonocardia autotrophica*, the NppY extending GT recognises 10-deoxynystatin but not nystatin.[44] After disaccharide formation, final C-10 hydroxylation occurs prior to export. The *Pseudonocardia* sp. P1 NypY GT is also discriminating, showing a preference for 8-deoxyamphotericins over amphotericins.[45] With aromatic heptaenes there is no cytochrome P450-mediated hydroxylation of the polyol chain. In 67-121C biosynthesis, the specificity of the ABC transporters may ensure that all late modifications are complete before export occurs. In heterologous hosts, rapid export of the normal product may preclude non-native late modifications. This would account for the lack of AceS-catalysed methylation of candicidins in *S. albidoflavus*.

The AceS protein was shown to modify surrogate substrates *in vitro*, thus confirming its function as an *N*-methyltransferase. The enzyme acted on 4-aminoacetophenone and esters of 4-aminobenzoic acid, with butyl esters being 100% methylated. The appearance of unmethylated 67-121B **4** and methylated 67-121A **5** as fermentation products suggests that methylation occurs as a late modification of the completed macrolactone. Our results indicate that methylation of nascent polyketide chains is also possible, if these intermediates are available to the methyltransferase while still attached to the PKS.

4-Aminobenzoic acid serves as a primer for aromatic polyene polyketides and also features in biosynthesis of other natural products. These include albicidin, a polyketide-peptide antibacterial that targets DNA gyrase,[46] and xanthomonic acid, a meroterpenoid with anticancer activity.[47] 4-Aminobenzoate provides a primer for the polyenyl pyrone polyketide aureothin, but the amine is first oxidised to a nitro group.[48] A similar oxidation reaction occurs with 4-aminophenylpyruvate in chloramphenicol biosynthesis.[49] It will be of interest to investigate whether AceS can be used to methylate intermediates or products in other pathways. Polyene *N*-methyltransferases may have the potential to interfere with pactamycin biosynthesis, in which the amino group of 3-aminoacetophenone is attached to an aminocyclitol core.[50] While many actinomycetes contain multiple biosynthetic gene clusters that might compete for precursors, pathways for pactamycin and a methylated aromatic heptaene might also actively antagonise each other.

In primary metabolism, 4-aminobenzoate is a precursor in folate biosynthesis, a target for sulphonamide antibiotics.[51] Methylation of 4-aminobenzoate might remove one of the normal substrates for the dihydropteroate synthase reaction. This would interfere with biosynthesis of tetrahydrofolate, which is essential for viability of many bacteria. This project showed that AceS methylates 4-aminoacetophenone but not 4-aminobenzoic acid, which indicates that methyltransferase substrate specificity prevents a potential clash between secondary and primary metabolism. The extra 4-aminobenzoate synthase gene within the cluster ensures that the requirements for aromatic polyene primers are met.

Methylation is a common late reaction in secondary metabolism. The method described here could be used as an alternative assay for SAM-dependent *N*-methylation of aromatic natural products, one which does not rely on radioactive labels or fluorescence. The short incubation time and rapid identification offers a new approach towards generating alkylated aromatic substrates with potentially altered physicochemical properties.

In cells, the methyl group donor SAM is synthesised from Met and ATP by SAM synthetase. Analogues of SAM with alternative alkyl groups have been synthesised by enzymatic and chemical methods. These include *S*-

adenosyl-L-ethionine and *S*-adenosyl-allyl-L-homocysteine. *O*-Methyltransferases capable of using these alternative alkyl donors have been used to produce analogues of rebeccamycin and rapamycin.[52-54] Some of these alkyl groups serve as reactive handles that allow further chemical modification. The AceS methyltransferase could potentially be used to generate analogues of aromatic heptaenes with unnatural alkyl groups attached to the amino group of the aminoacetophenone moiety. AceS may also be useful in alkylating 4-aminobenzoyl esters in efforts to synthesise antifolate antibiotics, or compounds derived from other natural products such as albicidin and xanthomonic acid.

5. Conclusions

The sequence of the 67-121 PKS reveals that pairing of A-type KR domains with DH domains results in *cis* double bond geometry in aromatic heptaene macrolides. This will assist bioinformatic prediction of polyketide structures from sequences of biosynthetic gene clusters. The AceS *N*-methyltransferase will be useful in chemoenzymatic alkylation of other bioactive compounds that contain aminoacetophenone groups.

Abbreviations

ACP	Acyl carrier protein
ER	Enoylreductase
GC-MS	Gas chromatography-mass spectrometry
GT	Glycosyltransferase
HPLC	High performance liquid chromatography
IPTG	Isopropylthiogalactoside
KR	Ketoreductase
LC-MS	Liquid chromatography-mass spectrometry
Ni-NTA	Nickel nitrilotriacetic acid
PKS	Polyketide synthase
PMSF	Phenylmethanesulphonylfluoride
SAHC	*S*-adenosylhomocysteine
SAM	*S*-adenosylmethionine

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

We thank Dr. Bernard Rawlings, Simon Walmsley and Mike Lee, Department of Chemisty, University of Leicester, for LC-MS analysis. James Sheehan was supported by a UCD research demonstratorship. Some of this work was supported by Science Foundation Ireland, grant number 09/RFP/GEN2132.

Notes and references

1. P. Caffrey, E. De Poire, J. Sheehan, P. Sweeney, *Appl Microbiol Biotechnol.,* 2016, **100**, 3893.

2. L. Soler, P. Caffrey, H. McMahon, *Biochim. Biophys. Acta,* 2008, **1780**, 1162.

3. A. Lemke, A. F. Kiderlin, O. Kayser, *Appl. Microbiol. Biotechnol.* 2005, **68**,151.

4. D. Cereghetti, E. Carreira, *Synthesis*, 2006, **6**, 914.

5. P. Caffrey, J. F. Aparicio, F. Malpartida, S. B. Zotchev, 2008, *Curr. Top. Med. Chem.*, **8**, 639.

6. P. Caffrey, *ChemBiochem*, 2003, **4**, 654.

7. S. Omura, H. Tanaka, 1984, Production, structure and antifungal activity of polyene macrolides. pp351-404 In *Macrolide Antibiotics: Chemistry, Biology and Practice*, Omura S. ed. Academic Press Inc. Harcourt Brace Jovanovich Publishers, New York.

8. K. Szwarc, P. Szczeblewski, P. Sowinski, E. Borowski, J. Pawlak, 2015, *J. Antibiot*. **68**, 504.

9. K. Szwarc, P. Szczeblewski, P. Sowiński, E. Borowski, J. Pawlak, *Magn. Reson. Chem*. 2015, **53**, 479.

10. P. Sowinski, P. Gariboldi, A. Czerwinski, E. Borowski, *J. Antibiotics* 1989, **42**, 1631.

11. P. Sowinski, P. Gariboldi, J. K. Pawlak, E. Borowski, *J. Antibiotics,* 1989, **42**, 1639.

12. P. Sowinski, J. Pawlak, E. Borowski, P. Gariboldi, *Polish. J. Chem.* 1995, **69**, 213.

13. J. Golik, J. Zielinski, E. Borowski, *J. Antibiotics,* 1980, **33**, 904.

14. Y. Zhou, J. Li, J. Zhu, S. Chen, L. Bai, X. Zhou, H. Wu, Z. Deng, *Chem. Biol.* 2008, **15**, 629.

15. P. Szczeblewski, T. Laskowski, B. Kubacki, M. Dziergowska, M. Liczmańska, J. Grynda, P. Kubica, A. Kot-Wasik, E. Borowski, *Scientific Reports* 2017, **7**, 40158.

16. J. K. Pawlak, P. Sowinski, E. Borowski, P. Gariboldi, *J. Antibiotics* 1995, **48**, 1034.

17. E. Hutchinson, B. Murphy, T. Dunne, C. Breen, B. Rawlings, P. Caffrey, *Chem. Biol.,* 2010, **17**,174.

18. M. S. Butler, *Nat. Prod. Rep*. 2008, **25**, 475.

19. J. Wright, D. Greeves, A. K. Mallams, D. H. Picker, J. C. S. Chem. Commun., 1977, **1977**: 710.

20. T. Tamura, Y. Nakagaito, T. Nishii, T. Hasegawa, E. Stackebrandt, A. Yokota, *Int J Syst Bacteriol*. 1994, **44**, 193.

21. M. J. Weinstein, G. Wagman, J. A. Marquez, K. G. Patel, 1977, United States Patent 4,027,015

22. S. Chen, X. Huang, X. Zhou, L. Bai, J. He, K. J. Jeong, S. Y. Lee, Z. Deng Z. *Chem. Biol*., 2003, **10**,1065.

23. H. Jørgensen, E. Fjærvik, S. Hakvåg, P. Bruheim, H. Bredholt, G. Klinkenberg, T. E. Ellingsen, S. B. Zotchev SB, *Appl Environ Microbiol.,* 2009, **75**, 3296.

24. T. Komori, *J. Antibiotics* 1990, **43,** 778.

25. N. Stephens, B. Rawlings, P. Caffrey, *Biosci. Biotechnol. Biochem*. 2013, **77**, 880.

26. B. Murphy, K. Anderson, C. Borissow, P. Caffrey, G. Griffith, J. Hearn, O. Ibrahim, N. Khan, N. Lamburn, M. Lee, K. Pugh, B. Rawlings, *Org. Biomol. Chem*. 2010, **8**, 3758.

27. T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood, 2000, *Practical* Streptomyces *Genetics*. John Innes Foundation, Norwich United Kingdom.

28. U. K. Laemmli, *Nature,* 1970,**227**, 680.

29. E. De Poire, N. Stephens, B. Rawlings, P. Caffrey, *Appl. Environ. Microbiol*., 2013, **79**, 6156.

30. L. Liu, Y. Li, S. Li, N. Hu, Y. He, R. Pong, D. Lin, L. Lu, M. Law, *J. Biomedicine and Biotechnol.* 2012, Volume **2012**, Article ID 251364. doi:10.1155/2012/251364

31. J. Sheehan, 2017, Ph. D. thesis, University College Dublin.

32. M. M. Alhamadsheh, N. Palaniappan, S. DasChouduri, K. A. Reynolds, *J. Am. Chem. Soc.*, 2007, **129**, 1910.

33. J. W. Labonte, C. A. Townsend, *Chem. Rev*., 2013, **113**, 2182.

34. A. T. Keatinge-Clay, *Chem. Biol.,* 2007, **14**, 898.

35. A. T. Keatinge-Clay, *Nat. Prod. Rep.,* 2016, **33**, 141.

36. D. H. Kwan, Y. Sun, F. Schulz, H. Hong, B. Popovic, J. C. Sim-Stark, S. F. Haydock, P. F. Leadlay, *Chem. Biol*., 2008, **15**,1231.

37. R. H. Lambalot, D. E. Cane, *J. Antibiotics,* 1992, **45**, 1981.

38. C. J. Thibodeaux, C. E. Melançon III, H-W. Liu, *Angew. Chem. Int. Ed.,* 2008, **47**, 9814.

39. M. Useglio, S. Peirú, E. Rodríguez, G. R. Labadie, J. R. Carney, H. Gramajo, *Appl Environ Microbiol.,* 2010, **76**, 3869.

40. L. Hong, Z. Zhao, C. E. Melançon 3rd, H. Zhang, H.-W. Liu, *J Am Chem Soc*., 2008, **130**, 4954.

41. D. P. O'Brien, P. N. Kirkpatrick, S. W. O'Brien, T. Staroske, T. I. Richardson, D. A. Evans, A. Hopkinson, J. B. Spencer, D. H. Williams, *Chem. Commun*., 2000, **2000**, 103.

42. H. Gao, S. Gruschow, J. Barke, R. F. Seipke, L. M. Hill, J. Orivel, D. W. Yu, M. Hutchings, R. J. M. Goss, *RSC Adv*., 2014, **4**, 57267.

43. N. Kandziora, J. N. Andexer, S. J. Moss, B. Wilkinson, P. F. Leadlay, F. Hahn, *Chem. Sci*., 2014, **5**, 3563.

44. H. J. Kim, M. K. Kim, M. J. Lee, H. J. Won, S. S. Choi, E. S. Kim, *PLoS One* 2015, **10**(4), e0123270.

45. S. Walmsley, E. De Poire, B. J. Rawlings, P. Caffrey, *Appl. Microbiol. Biotechnol.*, 2017, **101**, 1899.

46. S. Cociancich, A. Pesic, D. Petras , D. Uhlmann,J. Kretz, V. Schubert, L. Vieweg, S. Duplan, M. Marguerettaz, J. Noëll, I Pieretti, M. Hügelland, S. Kemper, A. Mainz, P. Rott, M. Royer, R. D. Süssmuth, *Nature Chem. Biol*., 2015, **11**, 195.

47. H. Saleh, D. Petras, A. Mainz, D. Kerwat, A. Nalbantsoy, Y. Erzurumlu, R. D. Süssmuth, *J. Nat. Prod*., 2016, **79**, 1532.

48. J. He, C. Hertweck, *Chem. Biol.,* 2003, **10**, 1225.

49. L. T. Fernández-Martínez, C. Borsetto, J. P. Gomez-Escribano, M. J. Bibb, M. M. Al-Bassam, G. Chandra, M. J. Bibb, *Antimicrob Agents Chemother*., 2014, **58**, 7441.

50. F. Kudo, Y. Kasama, T. Hirayama, T. Eguchi, 2007, *J. Antibiotics* **60**, 492.

51. A. Bermingham, J. P. Derrick, *BioEssays* 2002, **24**, 637.

52. C. Zhang, R. L. Weller, J. S. Thorson, S. R. Rajski, *J. Am. Chem. Soc.,* 2006, **128**, 2760.

53. B. J. C. Law, A.-W. Struck, M. R. Bennett, B. Wilkinson, J. Micklefield, *Chem. Sci.,* 2015, **6**, 2885

54. S. Singh, J. Zhang, T. D. Huber, M. Sunkara, K. Hurley, R. D. Goff, G. Wang, W. Zhang, C. Liu, J. Rohr, S. G. Van Lanen, A. J. Morris, J. S. Thorson, *Angew. Chem. Int. Ed. Engl*., 2014,**53**, 3965.

Electronic Supplementary Information

Molecular BioSystems

New insights into polyene macrolide biosynthesis in *Couchioplanes caeruleus*

J. Sheehan, C. D. Murphy and P. Caffrey†

School of Biomolecular and Biomedical Science, University College Dublin, Belfield, Dublin 4, Ireland.

† Correspondence: patrick.caffrey@ucd.ie.

List of figures and tables

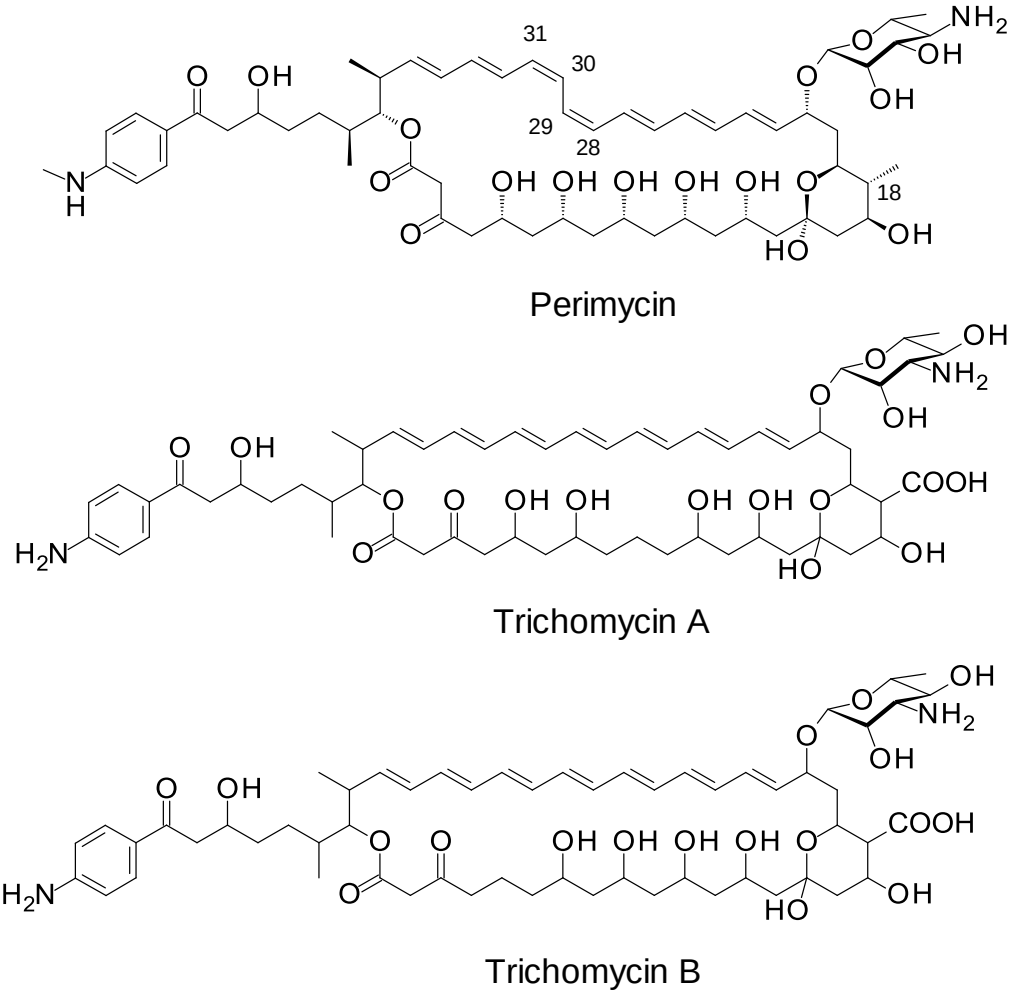


Fig. S1 Additional aromatic heptaenes mentioned in this work. Perimycin A differs from other members of the partricin group in that the methyl branch at C18 is not oxidised to a carboxyl group, and the sugar is D-perosamine rather than D-mycosamine.

Table S1. Polyene 67-121 biosynthetic genes

Gene product	Size, AA	Location in contig 1, accession number MEIA00000000
Hypothetical protein	521	335-1900
PabC 4-aminobenzoate synthase	251	2115-2870c
Propionyl CoA carboxylase	469	2998-4407
AceR3 Transcriptional regulator	923	4557-7328
AceR2 Transcriptional regulator	893	7331-10012
AceR1 Transcriptional regulator	191	10685-11260
AceD3 GDP-α-D-mannose dehydratase	342	11338-12366
Hypothetical protein (chorismate mutase)	102	12437-12745c
AceP4 PKS, modules 11 – 16	9002	12847-39852
AceP5 PKS, modules 17 – 20	3185	39852-49409
AceP6 PKS, module 21	5128	49443-64829
AceP2 PKS, modules 2 – 4	5049	64811-79960
AceP3 PKS, modules 5 - 10	9692	79957-109035
AceS Methylase	264	109056-109910
AceP1 PKS, module 1	1615	109911-114758c
pABA synthase	698	114774-116869c
AceTE	259	116907-117719c
AceM Ferredoxin	63	117747-117938c
AceN Cytochrome P450	405	117968-119185c
AceD2 Mycosamine synthase	323	119182-120240c
AceD1 Mycosaminyltransferase	457	120237-121610c
AceR5 Transcriptional regulator	548	121873-123520
Phosphatase	241	123488-124249c
Ace T1 ABC transporter	576	124326-126056
AceT2 ABC transporter	627	126053-127937
Amidohydrolase	408	128093-129319c
Hypothetical protein	104	129297-129611c

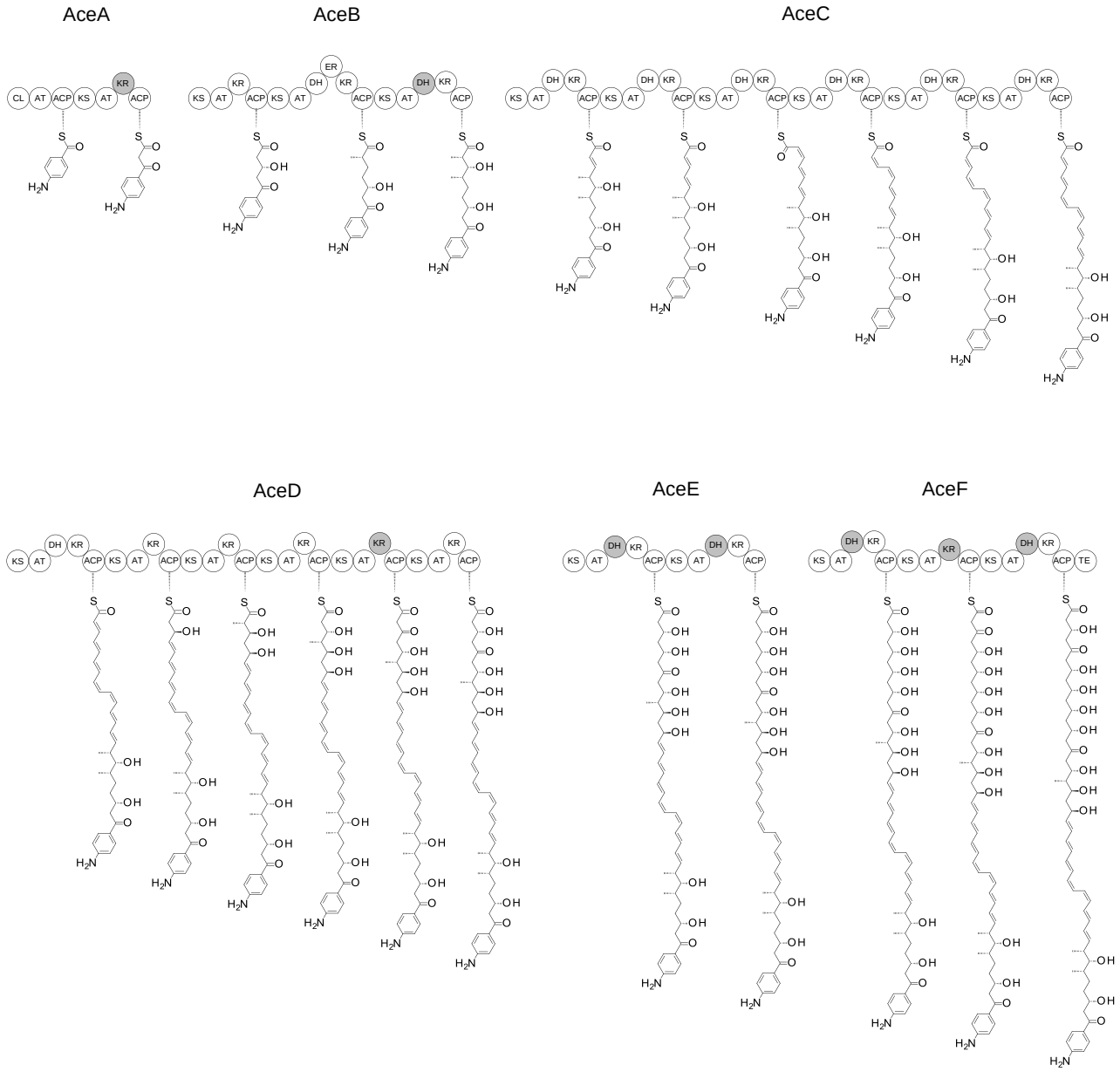


Fig. S2 The *C. caeruleus* (Ace) polyketide synthase. Inactive domains are shaded in grey.

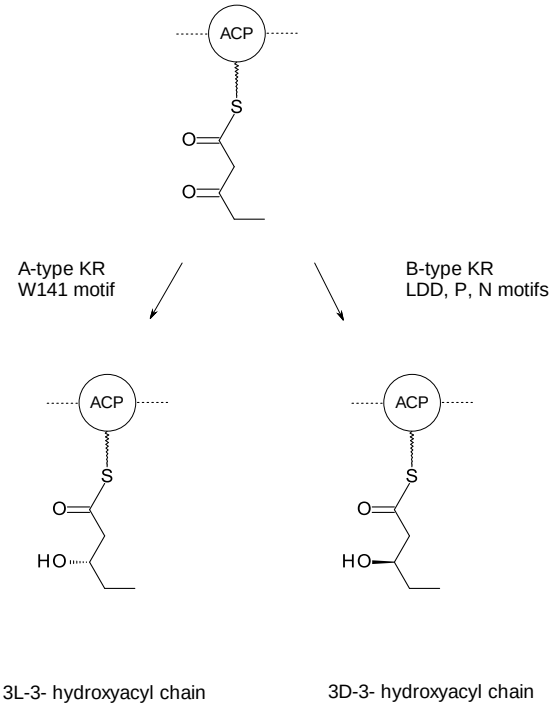


Fig. S3 A and B-type KR domains specify 3L ("3*S*") and 3D ("3*R*") alcohol stereochemistry. A 3-ketopentanoyl-ACP thioester substrate is shown as an example. A type KRs give 3L-3-hydroxyacyl chains, B-type KRs give 3D-3-hydroxyacyl chains. A-type KRs have a conserved tryptophan W-141 whereas B-type KRs have the LDD motif.

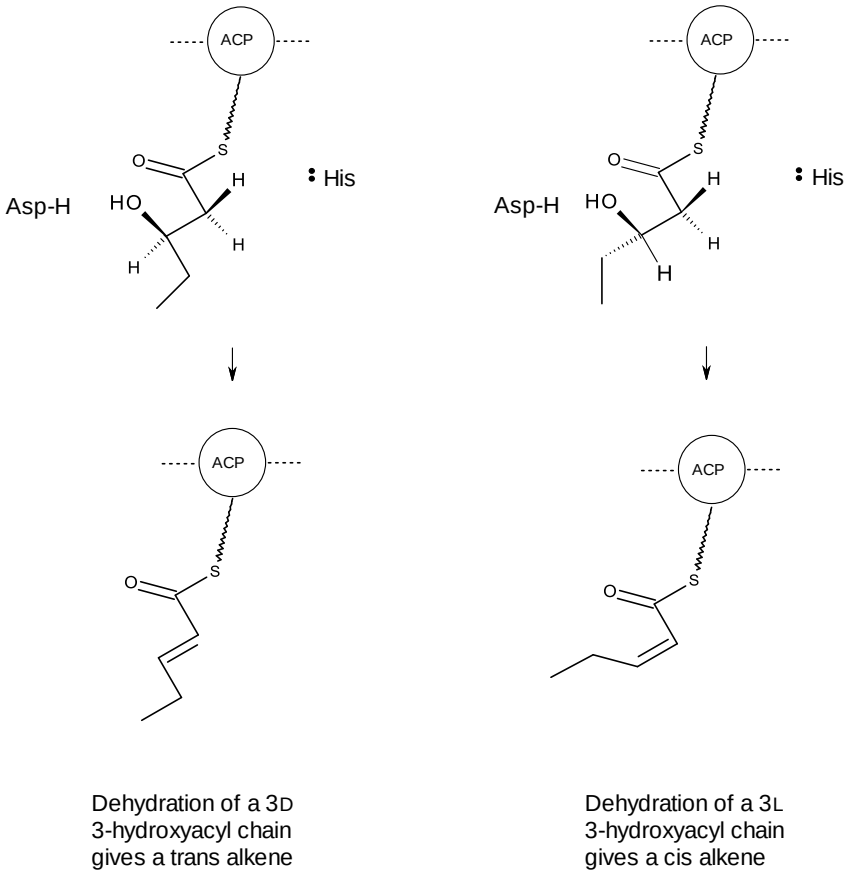


Fig. S4 DH domains give *trans* double bonds when paired with B-type KRs and *cis* double bonds when paired with A-type KRs. His and Asp residues in the DH active site act as base and acid catalysts during the syn elimination reaction.[33, 35]

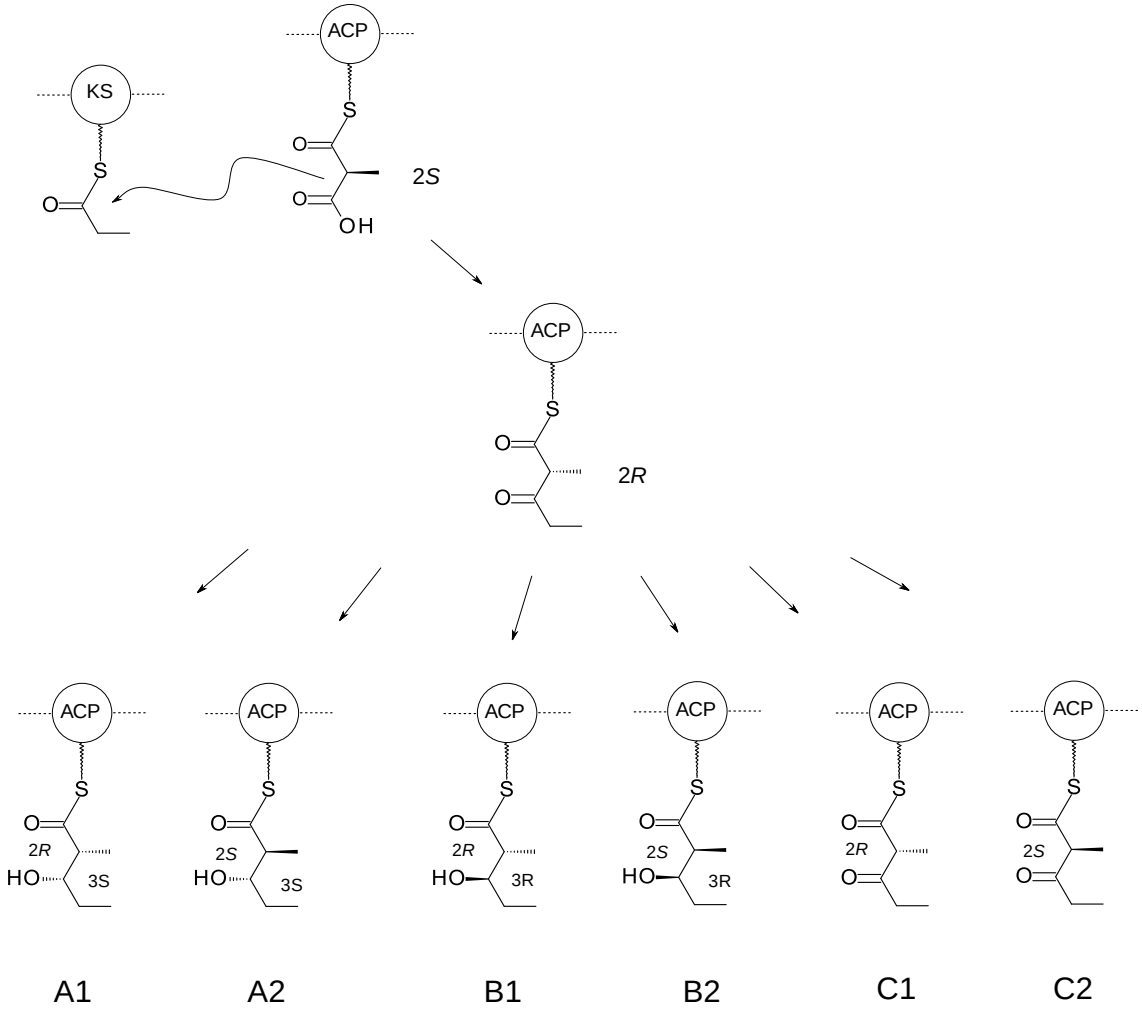


Fig. S5 Role of KR domains in specifying methyl stereochemistry. PKS KR domains process their 2*R* 2-methyl-3-ketoacyl-ACP substrates in different ways (a 2*R* 2-methyl-3-ketopentanoyl-ACP thioester is shown here as an example). A1 and B1 KRs reduce this substrate as shown. A2 and B2 KRs epimerise C-2 prior to ketoreduction. C-type KRs are inactive as reductases but the C2-subtype retain epimerase activity. Keatinge-Clay (2007) has identified additional key residues in these KRs that allow prediction of the chiral configuration of the methyl-branched centre as well as that of the secondary alcohol. [34]

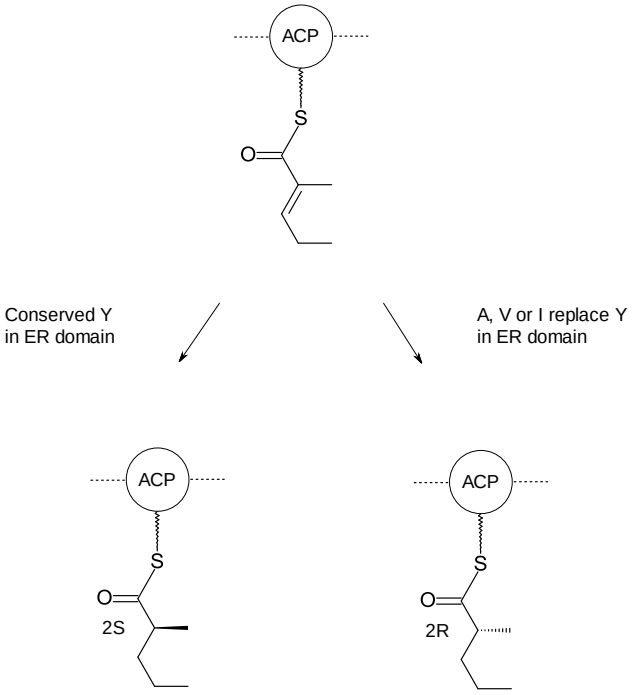


Fig. S6 The ER determines methyl stereochemistry when the β-ketone is fully processed. A conserved tyrosine correlates with stereochemical outcome.

Table S2 Amino acids revealing KR type in stereospecificity motifs.

KR type	1	2	3	4	5	6
A	Not LDD	W	-	Y	-	
A1	Not LDD	W	Not H	Y	-	
A2	Not LDD	W	H	Y	-	
B	LDD	-	-	Y	-	
B1	LDD	-	-	Y	Not P	
B2	LDD	-	-	Y	P	
C1				Not Y		
C2						Not N

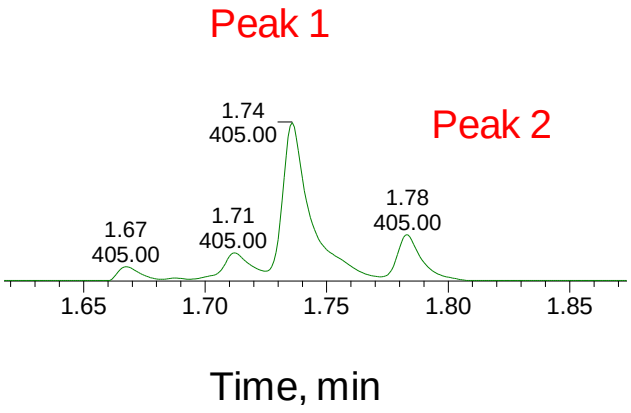
```
                1                                                2    3  4 5 6
AceA KR1  Deleted
AceB KR2  HAAGVLDDGLLTTLTPAKLDAVLRAKAQAAANLDDLT--GDLDMFVLFSSIAGSVGNHGQANYAAAN  B-type
AceB KR3  HAAGQLDDGTVASLTPDRIRAVMRPKADAARHLDELTRGHDLAEMVYFSSAAGVFGSPGQGNYAAAN  B-type
AceB KR4  HAAGVLDDGLIESLTPQRLDAVLRPKADAAIHLDELTRDRDLRQFVLFSSFAGVAGGMAQANYAAAN  B1-type
AceC KR5  HSAGVLDDGVIGSLTPERLATVLRPKVDAAWNLHTATLVRDLDAFVLFSSVSGLFGGPGQGSYSAAN  B-type
AceC KR6  HCAGVLDDGVIGSLTRERLATVLAPKVDAAWNLHTATLGRDLDAFILFSSVAGVFGAAGQGNYAAGN  B-type
AceC KR7  HAAGVGDNGLITALTPERLDAVLAPKADAAWYLHELTADMDLTAFVLISSVGGLVLTAGQGNYAAAN  A-type
AceC KR8  HAAGVGDNGLITALTPERLDAVLAPKADAAWYLHELTADMDLTAFVMFSSAGGTVLTGGQGNYAAAN  A-type
AceC KR9  HAAGVLDDGVIESLTPERADRVLQPKITAAWNLHAATRDRDLSAFVLFSSVAGLLGNPGQASYAAAN  B-type
AceC KR10 HAAGVLDDGVIGSLTPDRLDAVLRPKVDAAWNLHKATKDLDV--FVLFSSMAGLLGNPGQASYAAGN  B-type
AceD KR11 HAAGVLDDGVVESLTPQRLSTVLRPKADAVWNLHRA--AGDVAGFVVFSSFSGTAGAAGQANYAAAN  B-type
AceD KR12 HAAGVGQAGPLTAATLDEVAATVSAKMTGAAHLDSLLEGHDLDFLVLVSSIAGVWGSAGQSAYGAAN  A-type
AceD KR13 HTAAVIELASIEATSLDAFDRVMHAKVTGARLLDELLGDDLDDF-VLYSSTAGMWGSGQHAAYVAAN  A2-type
AceD KR14 HTAGIVDDGVIDALTPQRFAAVQRAKMDATRSLHELT-PDARAF-VLFSSTAGVLGAAGQGNYAAAN  B-type
AceD KR15 Deleted
AceD KR16 HAAGVLDDGILDGLTAAQFATVFRAKVTSALLLDELTAGRDLTVFALFSSASAAVGNPGQANYAAAN  B-type
AceE KR17 HTAGVLDDGVITALNPDRLATVLRPKVDAAWNLHAATKDLDA--FVLFSSISGIMGSAGQANYAAGN  B-type
AceE KR18 HTAGVLDDGVITALNPDRLATVLRPKVDAAWNLHAATKDLDA--FVLFSSISGIMGSAGQANYAAGN  B-type
AceF KR19 HAAGILDDGILTSLTPQRLSAVLEPKVDGAWNLHLATASRHLDAFVLFSSISGVTGTAGQANYAAGN  B-type
AceF KR20 HAASAVDHGVVADLTADRLRLVVDAKVRPAILLDELTAGLDLDAFVLFSSVSGSVGSPGRAAIAAVG  C-type
AceF KR21 HIAGVLDDAVLTSLTPDRMERVLRPKVDVAWNLHELTCDMGLAAFVSFSSGAGIMGNPGQGNYAAAN  B-type


ER domain

AceB ER3 (2S) AGLNFRDVLNVLGMYPGGARYLGSEAAGVVVEVADDVTTLAPGDRVTGMVAGGFGTHAIA
```

Fig. S7 Stereospecificity motifs in *C. caeruleus* (Ace) polyene polyketide synthase domains. Most KRs have the LDD motif characteristic of B-type KRs that form 3D-3-hydroxyacyl-ACP intermediates. KR12 and KR13 contain the conserved W typical of A-type KRs, which give 3L stereochemistry. KR19 lacks the active site Y (green type) and must be inactive. KR13 is predicted to form a (2*S*, 3*S*)-2methyl-3-hydroxyacyl intermediate (fingerprint H residue is magenta). KR4 is predicted to generate a (2*R*, 3*R*)-2-methyl-3-hydroxyacyl intermediate (fingerprint residue in magenta is A not P). The ER3 domain is predicted to give a (2*S*)-2-methyl-branched intermediate (fingerprint Y residue is magenta). KR7 and KR8 are A-type KRs paired with dehydratase domains. Modules 7 and 8 are predicted to generate *cis* double bonds.

A



B

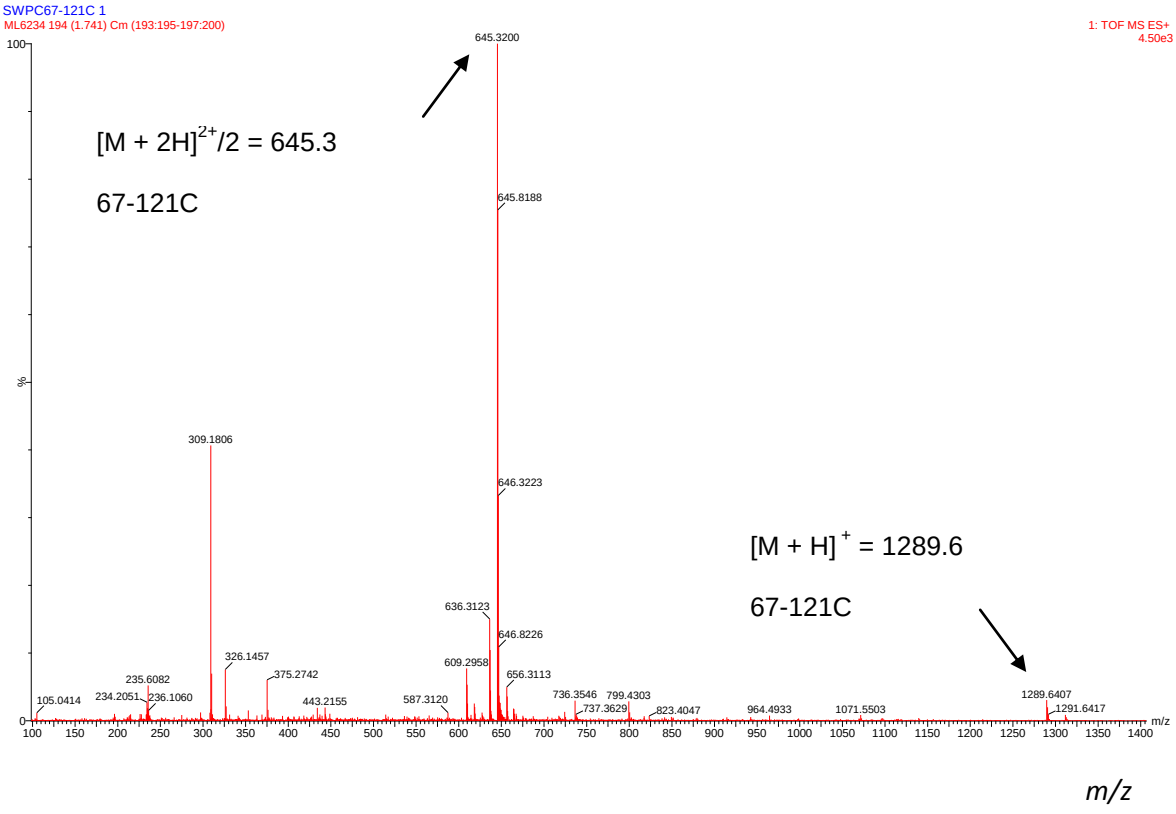
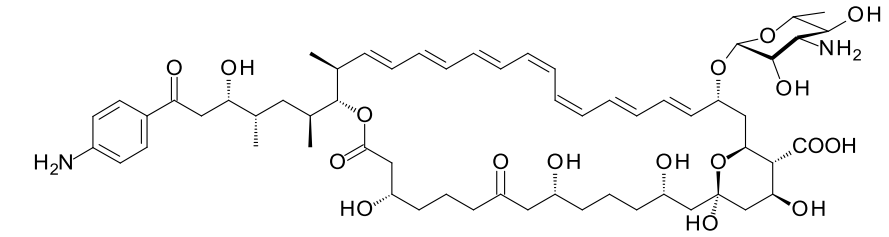
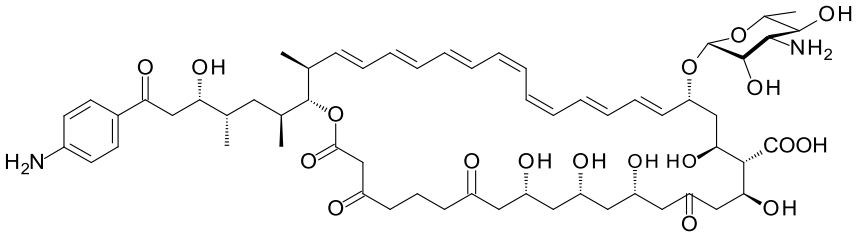


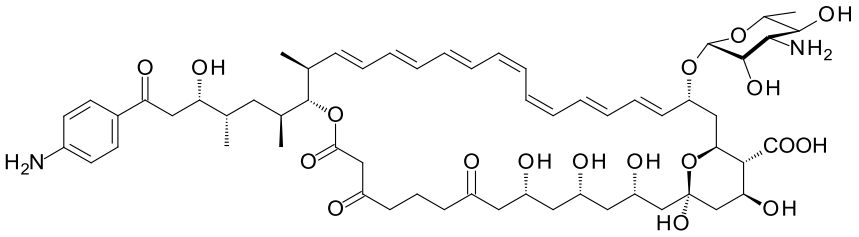
Fig. S8 LCMS analysis of polyenes from *C. caeruleus*. A. Section of HPLC chromatogram showing heptaene peaks. B. Mass spectrum of major peak 1 showing doubly protonated and singly protonated 67-121C ions.



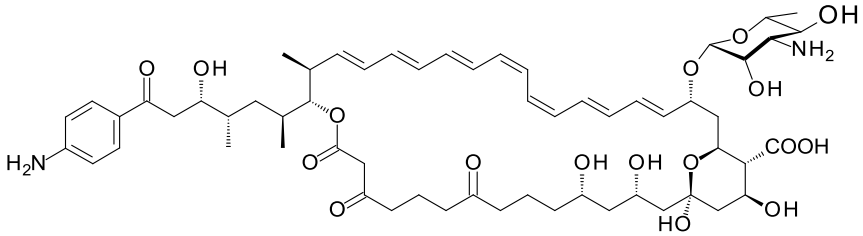
Candicidin I (candicidin A3)



Candicidin II



Candicidin III (candicidin D)



Candicidin IV (candicidin A1)

Fig. S9 Main components of the candicidin complex.

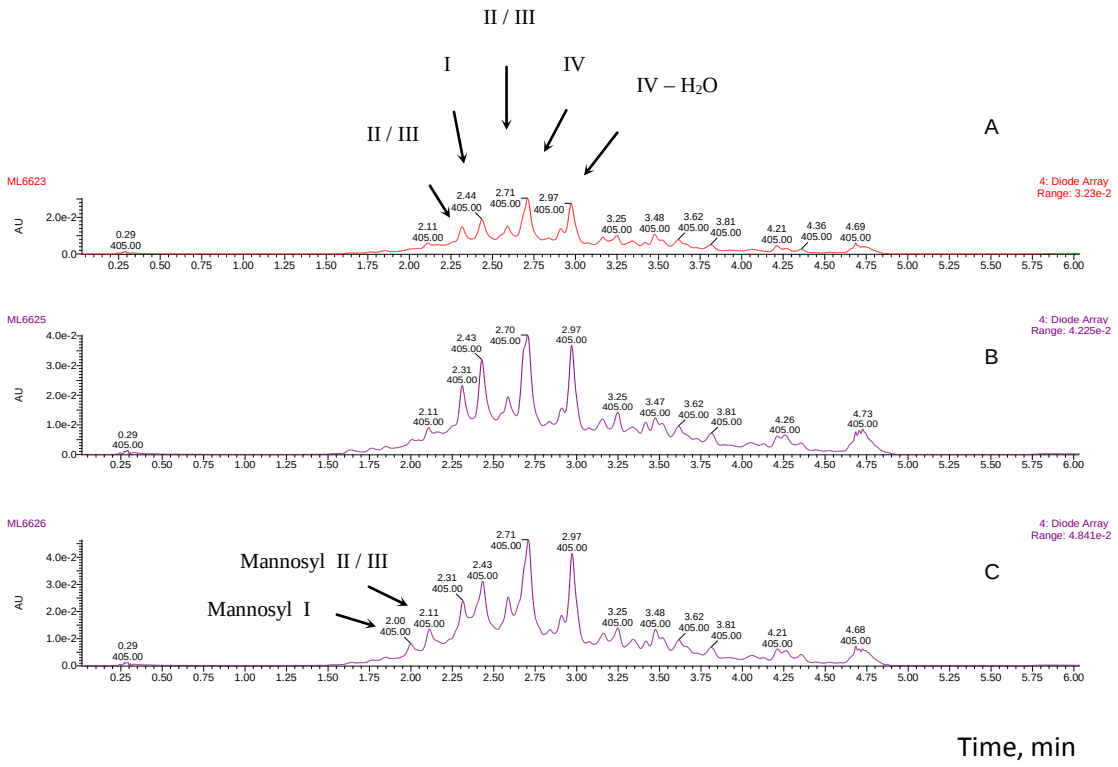


Fig. S10 LC-MS analysis of candicidins from (A) *S. albidoflavus* pIAGO, (B) *S. albidoflavus* pIAGO-*aceS*, (C) *S. albidoflavus* pIAGO-*pegA*+*aceS*. The compounds were identified from mass spectra, candicidins I, II, III and IV, and mannosylated forms in (C). Candicidins II and III have the same mass and cannot be distinguished from this analysis.

Table S3 Masses of candicidins detected in Fig S10.

Candicidin	**Mass**	**$[M + H]^+$ observed**
Candicidin I	1110.6	1111.6
Candicidin II	1108.6	1109.6
Candicidin III	1108.6	1109.6
Candicidin IV	1092.6	1093.6
Mannosyl-candicidin I	1272.6	1273.6
Mannosyl-candicidin II	1270.6	1271.6
Mannosyl-candicidin III	1270.6	1271.6
Mannosyl-candicidin IV	1254.6	Not detected

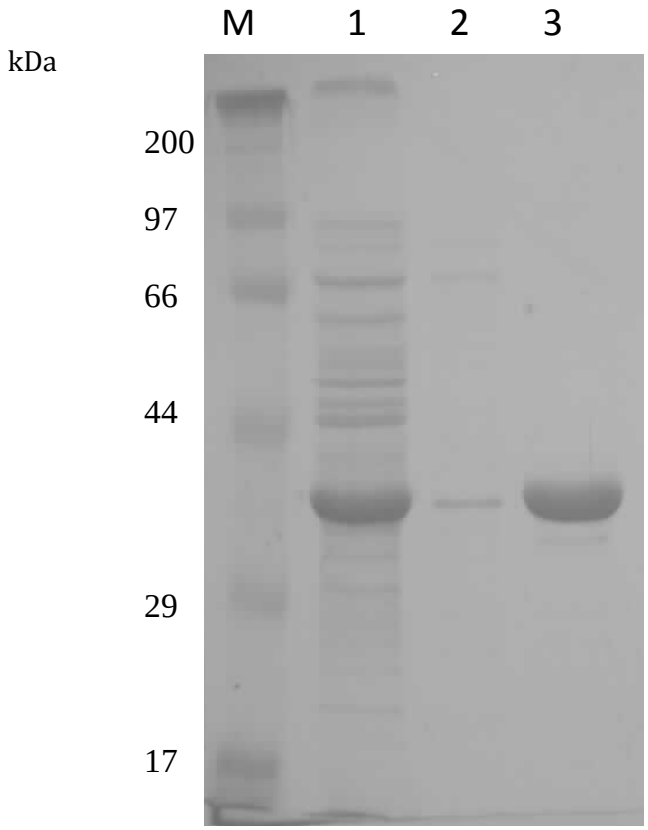


Fig. S11 SDS-PAGE analysis of purified AceS methyltransferase. Lanes: M = Protein molecular weight markers; 1 = soluble fraction from *E. coli* BL32 DE3 pET28-MetN, 2 = AceS after purification on a Ni-NTA column; 3 = concentrated purified AceS protein.

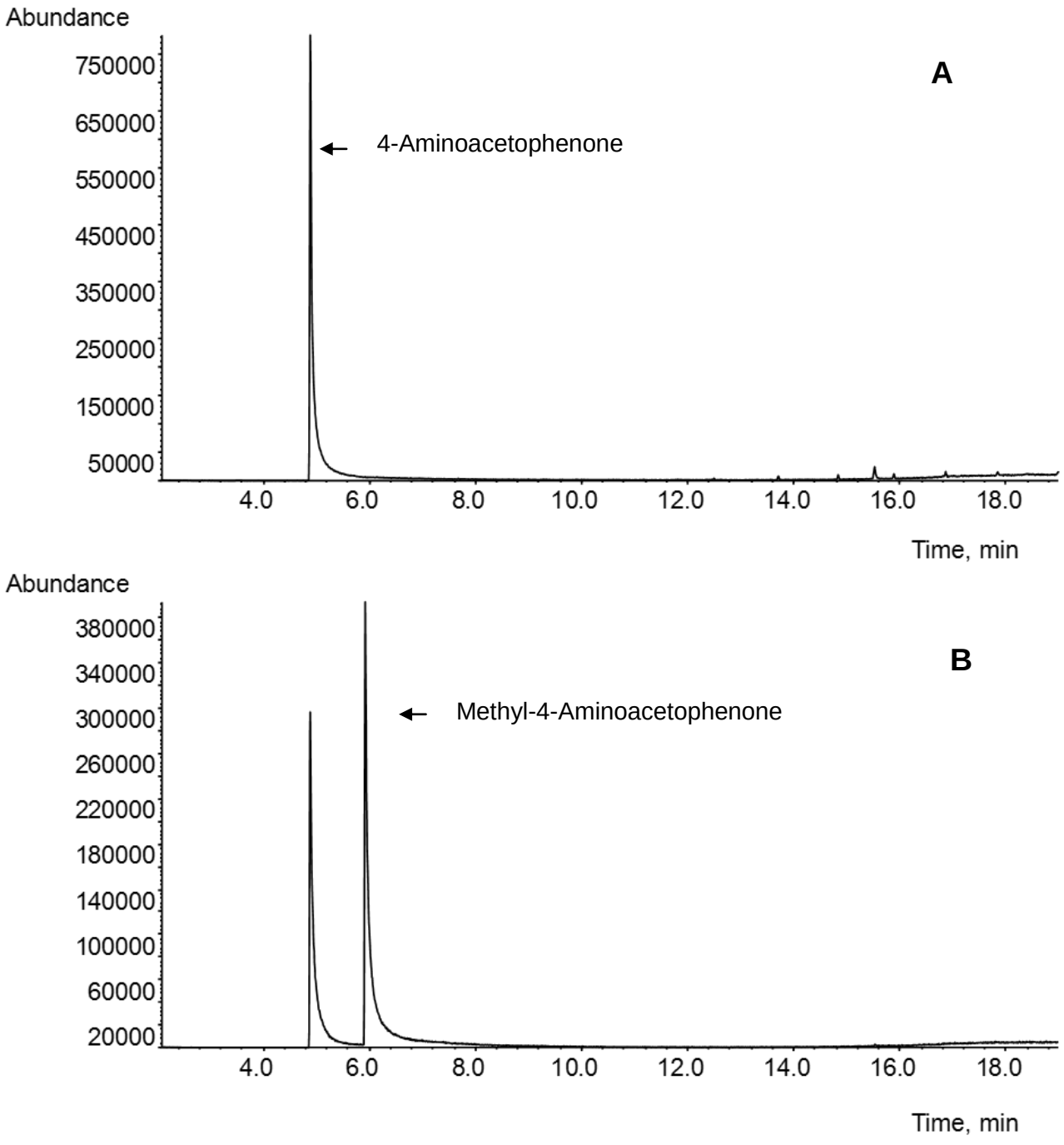


Fig. S12 GC-MS of AceS-catalysed methylation of 4-aminoacetophenone. A. Chromatogram from analysis of 4-aminoacetophenone control. B. Chromatogram from analysis of reaction mixture after 120 minute incubation.

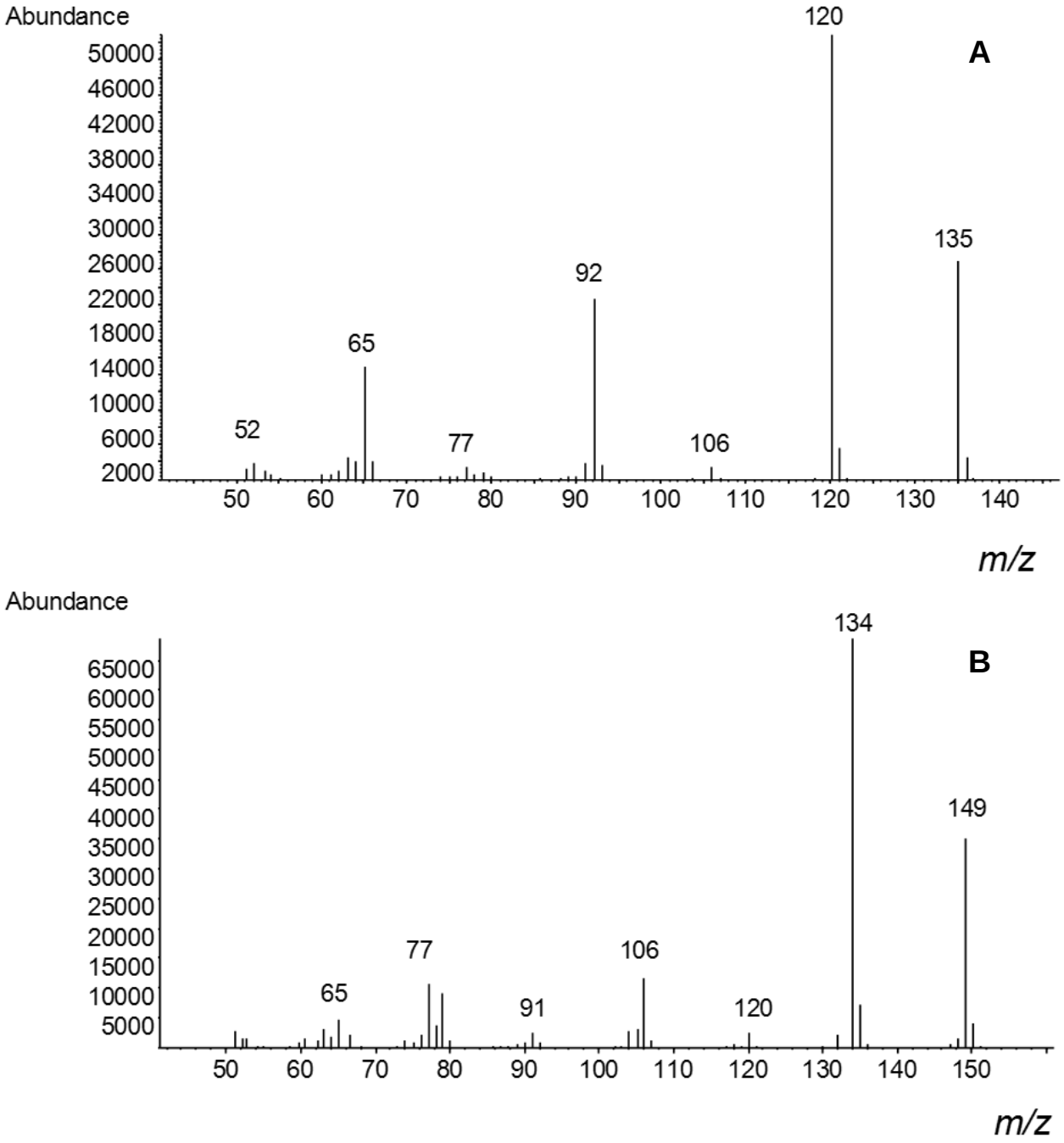


Fig. S13 GC-MS of AceS-catalysed methylation of 4-aminoacetophenone. A. Mass spectrum of 4-aminoacetophenone control peak. B. Mass spectrum of methyl-4-aminoacetophenone peak.

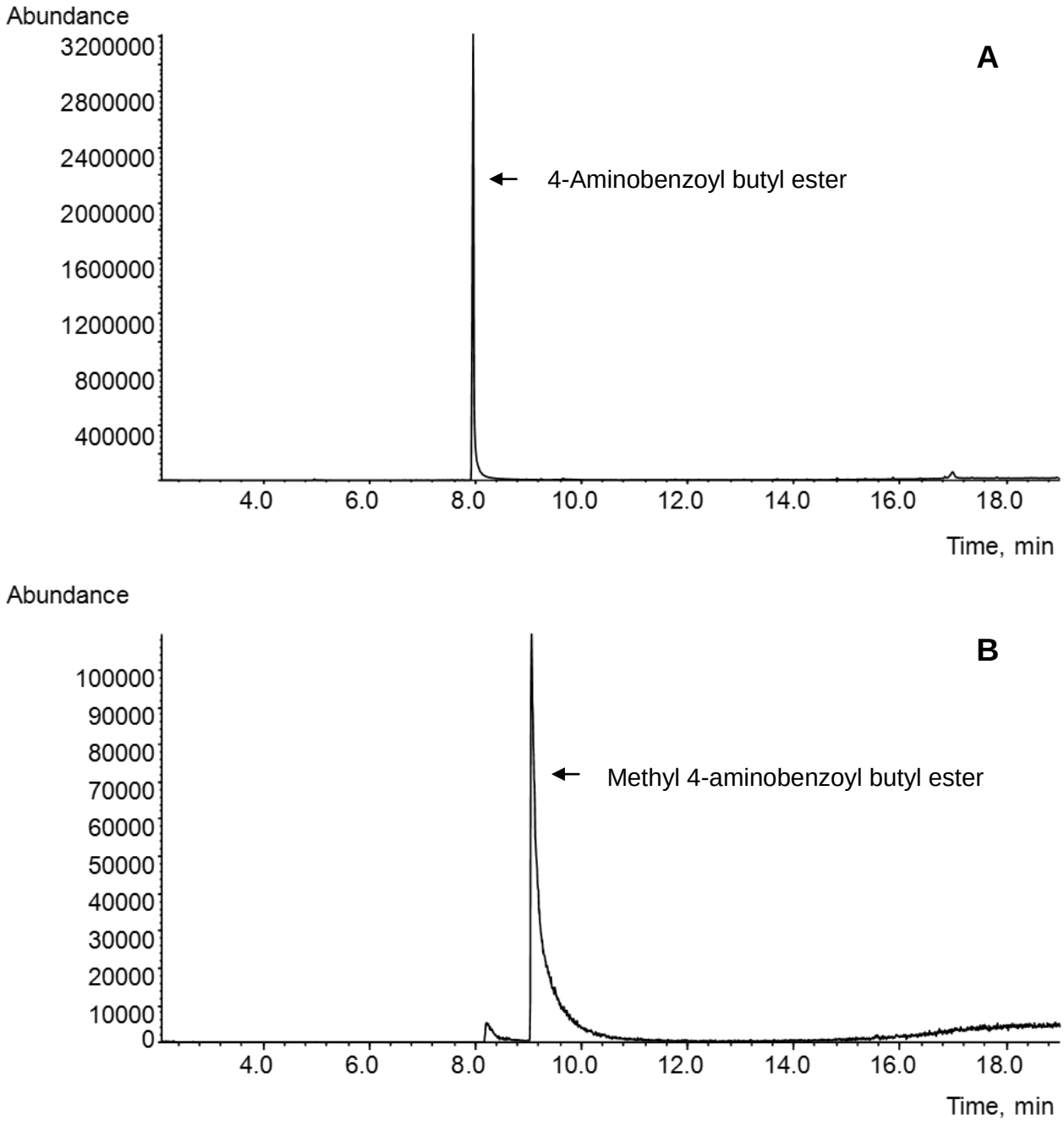


Fig. S14 GC-MS of AceS-catalysed methylation of 4-aminobenzoyl butyl ester. A. Chromatogram from analysis of 4-aminobenzoyl butyl ester control. B. Chromatogram from analysis of reaction mixture after 120 minute incubation.

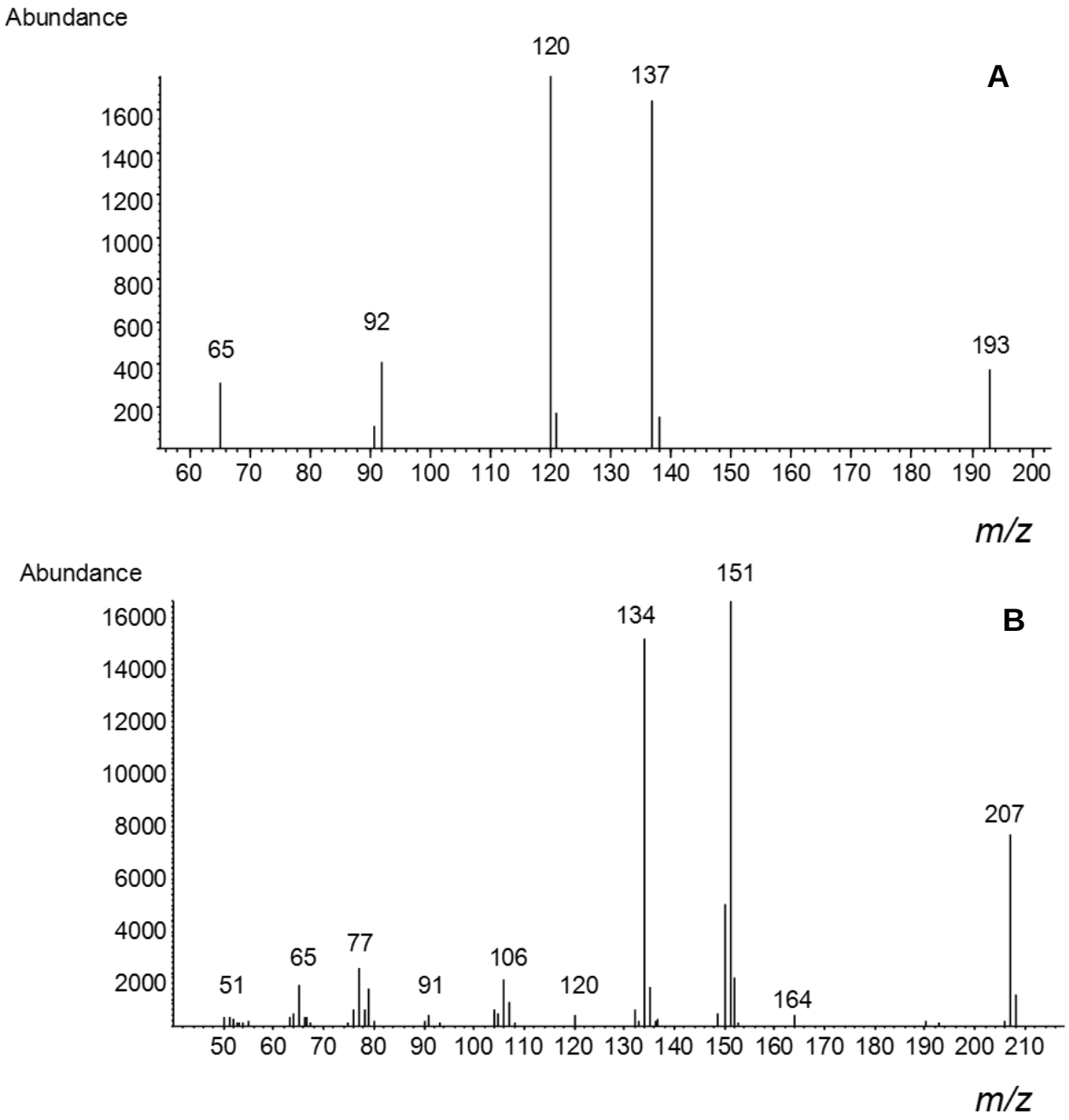


Fig. S15 GC-MS of AceS-catalysed methylation of 4-aminoacetophenone. A. Mass spectrum of 4-aminobenzoyl butyl ester control peak. B. Mass spectrum of methyl-4-aminobenzoyl butyl ester peak.