

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 polyenes mentioned in this work.

Table S1 67-121C biosynthetic genes.

Fig. S2 The 67-121C PKS.

Fig. S3 A- and B-type KR domains.

Fig. S4 DH domains give *trans* double bonds when paired with B-type KR and *cis* double bonds when paired with A-type KR.

Fig. S5 Role of KR domains in specifying methyl stereochemistry.

Fig. S6 The ER determines methyl stereochemistry when the β -ketone is fully reduced.

Fig. S7 Stereospecificity motifs in *C. caeruleus* (Ace) polyene polyketide synthase domains.

Table S2 Amino acids revealing KR type in stereospecificity motifs.

Fig. S8 LCMS analysis of polyenes from *C. caeruleus*.

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.

Table S3 Masses of candicidins detected in Fig S10.

Fig. S11. SDS-PAGE analysis of purified AceS methylase.

Fig. S12 Chromatograms from GC-MS analysis of (A) 4-aminoacetophenone and (B) methylated 4-aminoacetophenone.

Fig. S13 Mass spectra of (A) 4-aminoacetophenone and (B) methylated 4-aminoacetophenone

Fig. S14 Chromatograms from GC-MS analysis of (A) 4-aminobenzoyl butyl ester and (B) methylated 4-aminobenzoyl butyl ester.

Fig. S15 Mass spectra of (A) 4-aminobenzoyl butyl ester and (B) methylated 4-aminobenzoyl butyl ester.

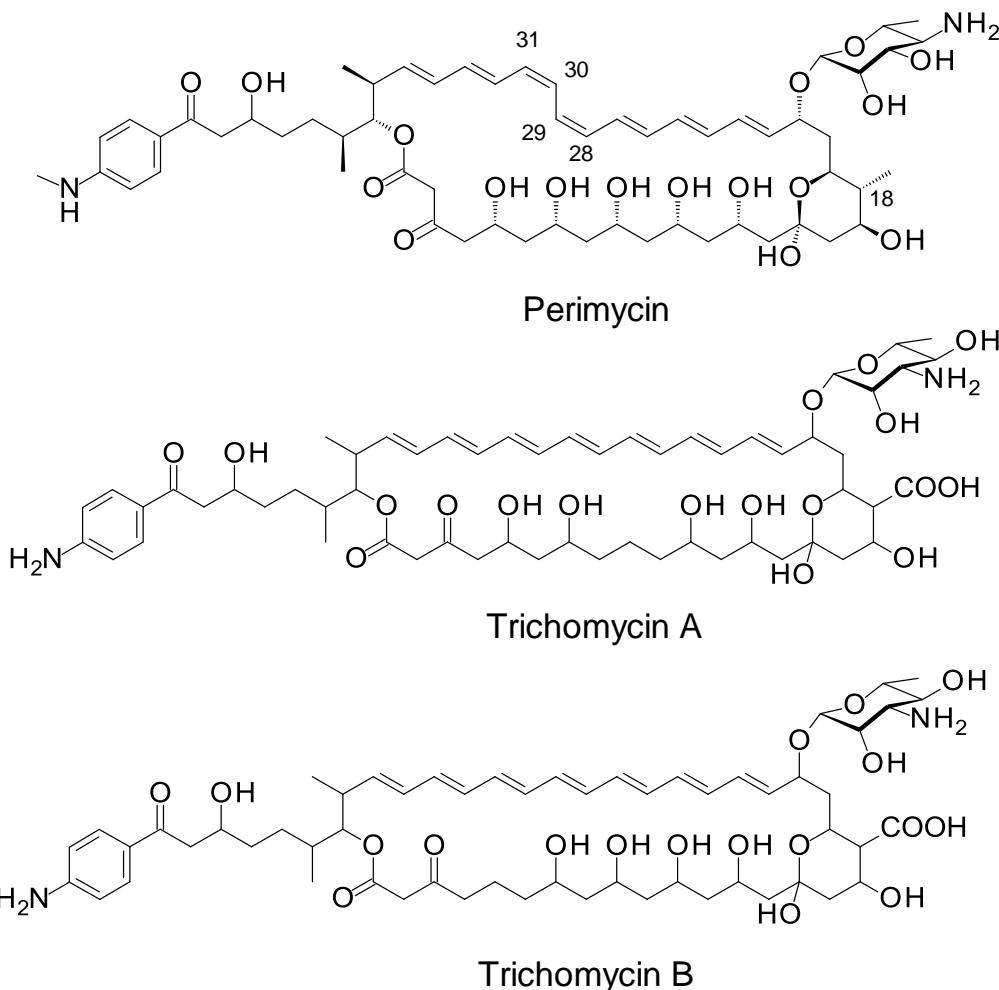


Fig. S1 Additional aromatic heptaenes mentioned in this work. Perimycin A differs from other members of the partricin group in that the methyl branch at C18 is not oxidised to a carboxyl group, and the sugar is D-perosamine rather than D-mycosamine.

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 Mycosaminytransferase	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

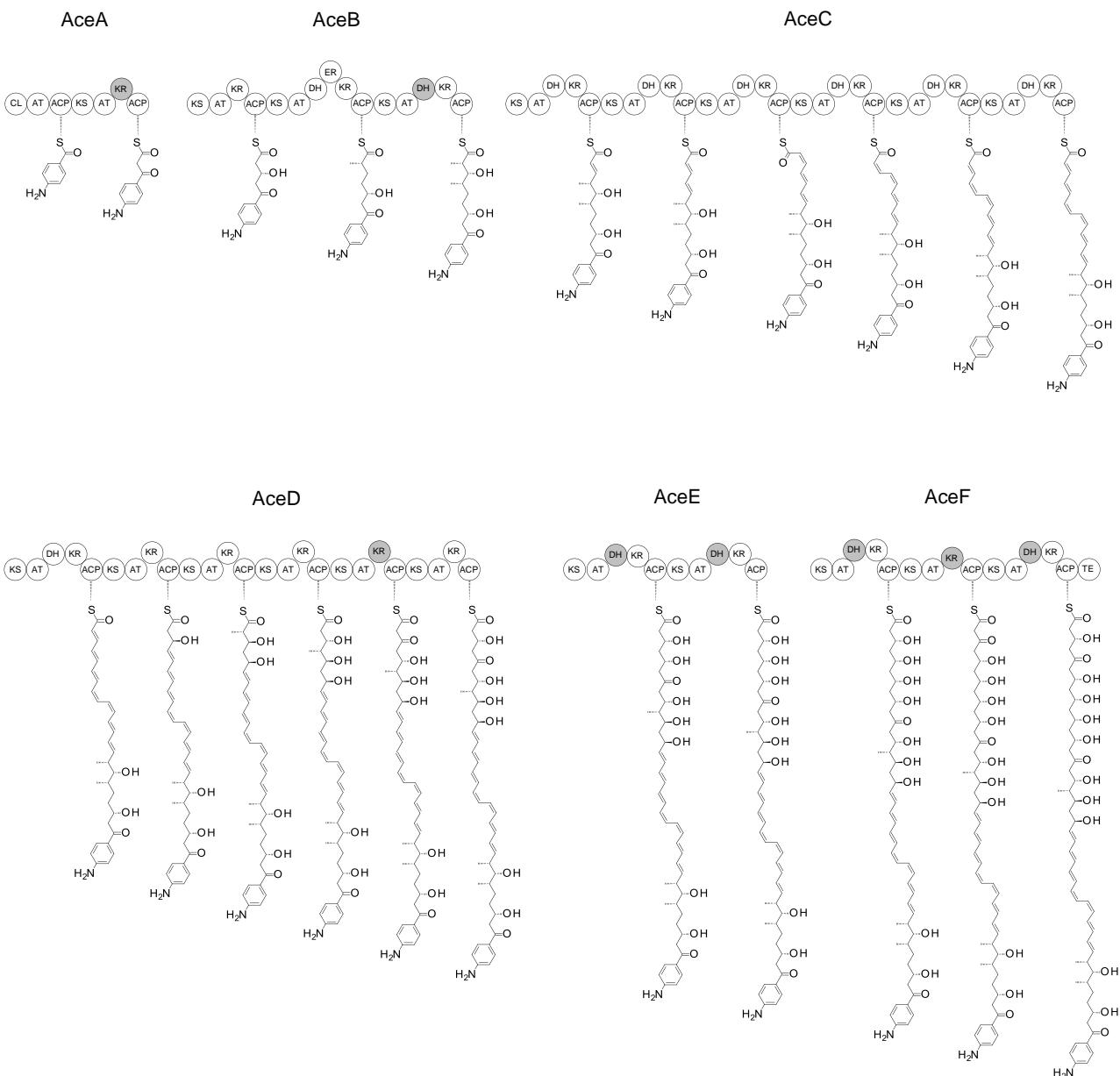


Fig. S2 The *C. caeruleus* (Ace) polyketide synthase. Inactive domains are shaded in grey.

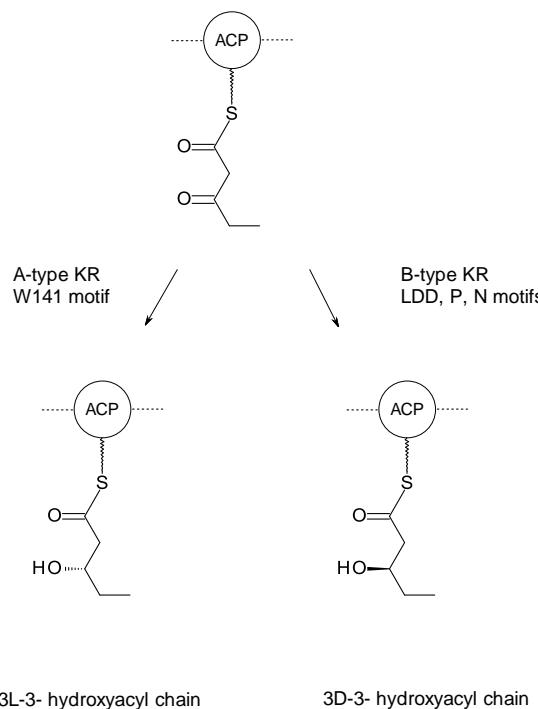


Fig. S3 A and B-type KR domains specify 3L ("3S") and 3D ("3R") alcohol stereochemistry. A 3-ketopentanoyl-ACP thioester substrate is shown as an example. A type KR domains give 3L-3-hydroxyacyl chains, B-type KR domains give 3D-3-hydroxyacyl chains. A-type KR domains have a conserved tryptophan W-141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

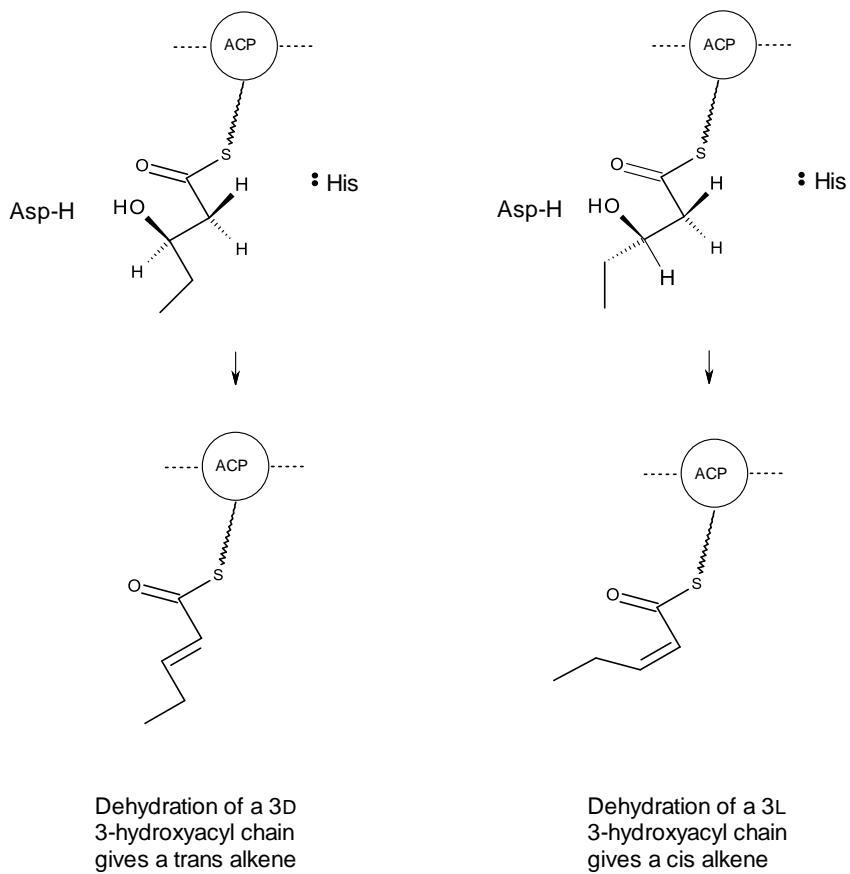


Fig. S4 DH domains give *trans* double bonds when paired with B-type KRs and *cis* double bonds when paired with A-type KRs. His and Asp residues in the DH active site act as base and acid catalysts during the syn elimination reaction.^{33, 35}

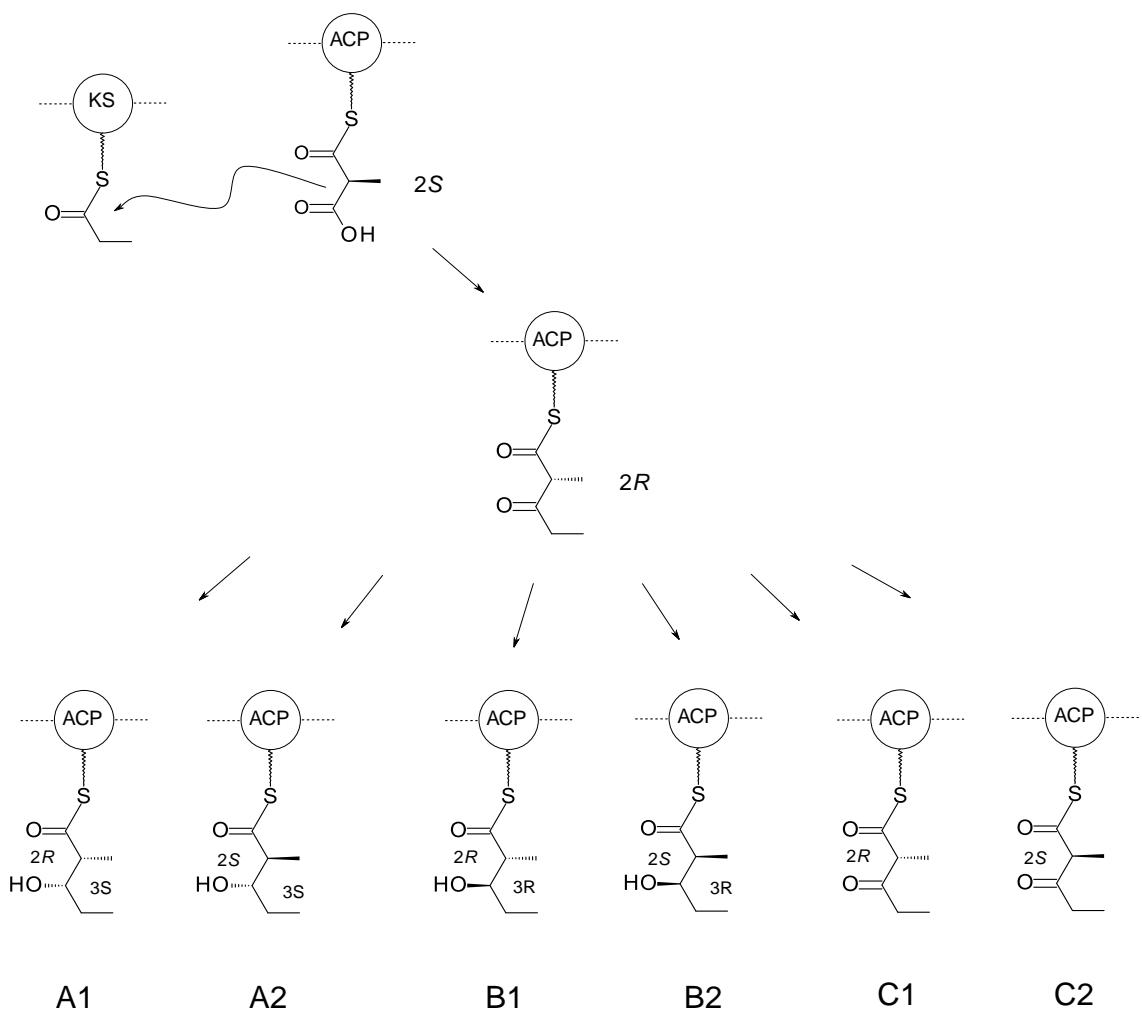


Fig. S5 Role of KR domains in specifying methyl stereochemistry. PKS KR domains process their 2R 2-methyl-3-ketopentanoyl-ACP substrates in different ways (a 2R 2-methyl-3-ketopentanoyl-ACP thioester is shown here as an example). A1 and B1 KRs reduce this substrate as shown. A2 and B2 KRs epimerise C-2 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

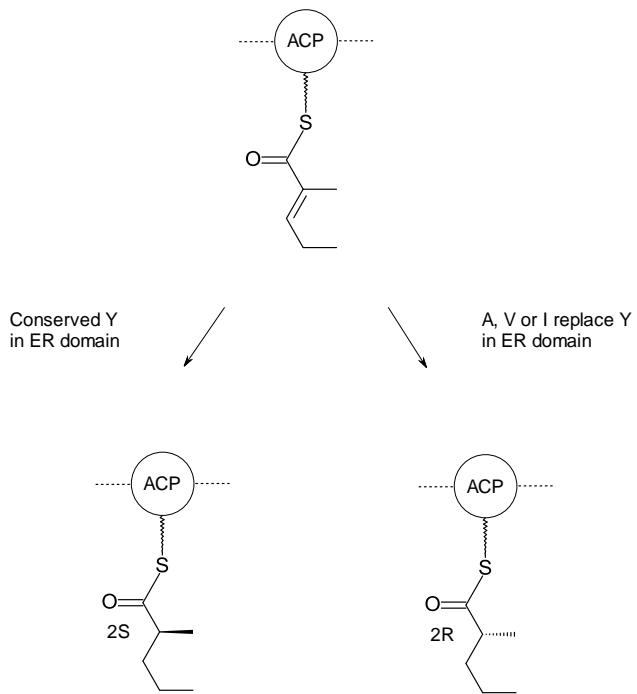


Fig. S6 The ER determines methyl stereochemistry when the β -ketone is fully processed. A conserved tyrosine correlates with stereochemical outcome.

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	HAAGV LDD GLLTTLTPAKLDAVLRAKAQAAANLDDLT--GDLDMFVLFSSIAGSVGNHGQANYAAAN						B-type
AceB KR3	HAAGQ LDD GTVASLTPDRIRAVMRPKADAARHLDELTRGHDLAEMVYFSSAAGVFGSPGQGNYAAAN						B-type
AceB KR4	HAAGV LDD QLIESLTPQRQLDAVLRPKADAAIHLDELTRDRDLRQFVLFSSFAGVAGGMAQANYAAAN						B1-type
AceC KR5	HSAGV LDD GVIGSLTPERLATVLRPKVDAAWNLTATLVRDLDASFVLFSSVSGLFGGPGQGSYSAAN						B-type
AceC KR6	HCAGV LDD GVIGSLTRERLATVLPKVDAAWNLTATLGRDLDASFILFSSVAGVFGAAAGQGNYAAGN						B-type
AceC KR7	HAAGV GDN GLITALT PERLDAVLPKADAAWYLHELTADMDLTAFVLISSVGGVLTAGQGNYAAAN						A-type
AceC KR8	HAAGV GDN GLITALT PERLDAVLPKADAAWYLHELTADMDLTAFVMFSSAGGTVLGGQGNYAAAN						A-type
AceC KR9	HAAGV LDD GVIESLTPERADRLQPKITAAWNHLHAATRDRDLSAFVLFSSVAGLLGNPGQASYAAAN						B-type
AceC KR10	HAAGV LDD GVIGSLTPDRLDLAVLPKVDAAWNHLHATKDDV--FVLFSSMAGLLGNPGQASYAAGN						B-type
AceD KR11	HAAGV LDD GVVESLTPQRSLTVLRPKADAVWNLHRA--AGDVAGFVVFFSFSGTAGAAGQGNYAAAN						B-type
AceD KR12	HAAGV QAG PLTAATLDEVAATVSAKMTGAAHLDLSLEGHDLDLFLVSSIAGV W GSAGQSYGAAN						A-type
AceD KR13	HTAAV IEL SIEATSLDAFDRVMHAKVTGARLLDELLGDDLDDF-VLYSSTAGM W GSQ H AYVAAN						A2-type
AceD KR14	HTAGI VDD GVIDALTPQRFAAVQRAKMDATRSLHELT-PDARAF-VLFSSTAGVLGAAGQGNYAAAN						B-type
AceD KR15	Deleted						
AceD KR16	HAAGV LDD GILDGLTAAQFATVFRAKVTSALLDELTAGRDLTVFALFSSASAAGVNPQQANYAAAN						B-type
AceE KR17	HTAGV LDD GVITALNPDRLATVLRPKVDAAWNHLHAATKDDA--FVLFSSISGIMGSAGQANYAAGN						B-type
AceE KR18	HTAGV LDD GVITALNPDRLATVLRPKVDAAWNHLHAATKDDA--FVLFSSISGIMGSAGQANYAAGN						B-type
AceF KR19	HAAGI LDD GILTSLTPQRSLSAVLEPKVDGAWNHLHATASRHDAFVLFSSISGVTGTAGQANYAAGN						B-type
AceF KR20	HAASAV DH GVVADLTADRLRLVVDAKVRPAIILDELTAGLDLDAFVLFSSVSGSPGRAIAAVG						C-type
AceF KR21	HIAGV LDD AVLTSLTPDRMERVLRPKVDAWNHLHELTCDMGLAAFVSSGAGIMGNPGQGNYAAAN						B-type

ER domain

AceB ER3 (2S) AGLNFRDVNLVLMGY**P**GGARYLGSEAAGVVVEADDVTTIAPGDRVGMVAGGFGTHAIA

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 (2S, 3S)-2methyl-3-hydroxyacyl intermediate (fingerprint H residue is magenta). KR4 is predicted to generate a (2R, 3R)-2-methyl-3-hydroxyacyl intermediate (fingerprint residue in magenta is A not P). The ER3 domain is predicted to give a (2S)-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

Peak 1

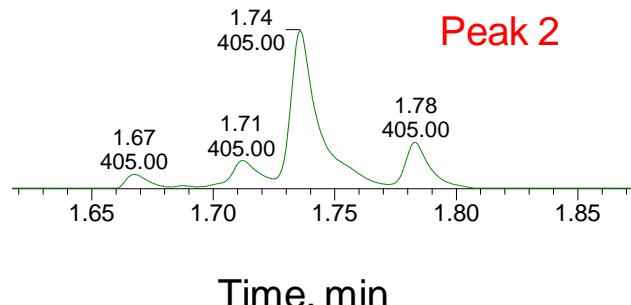
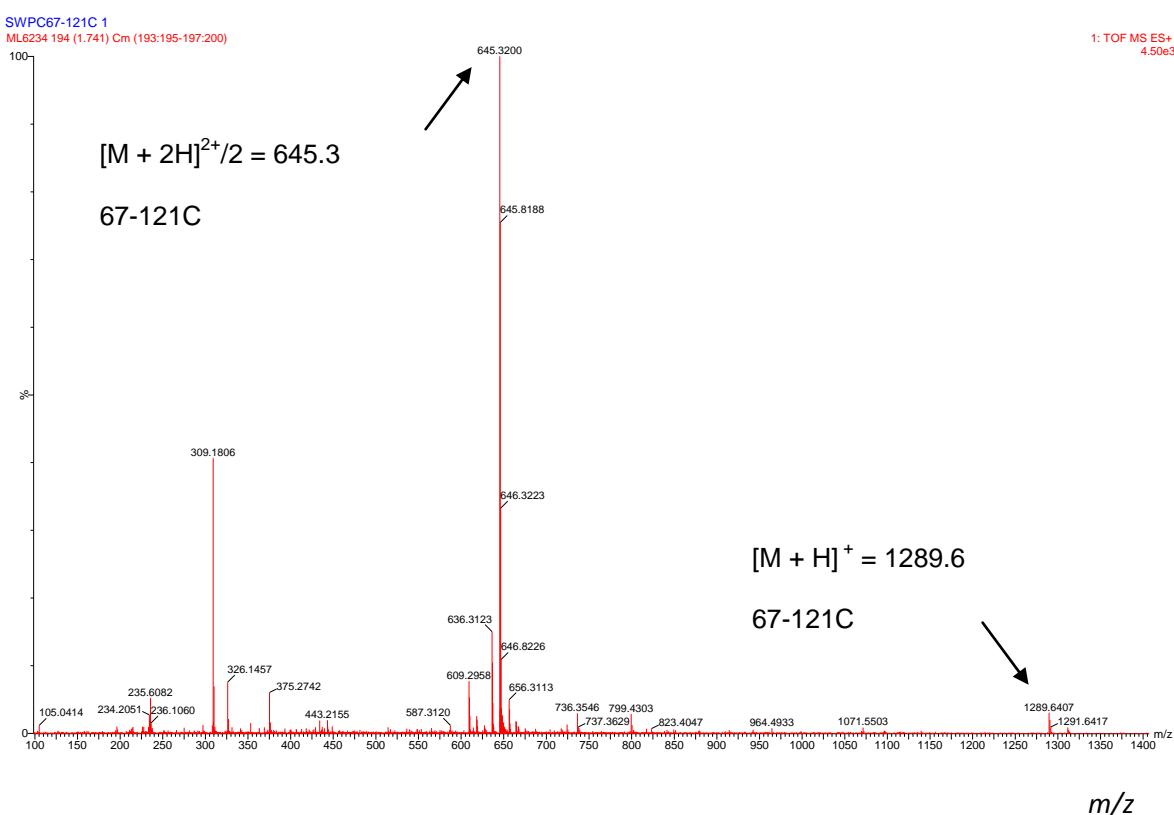
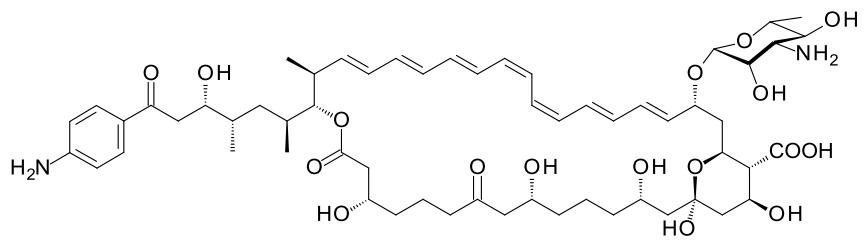
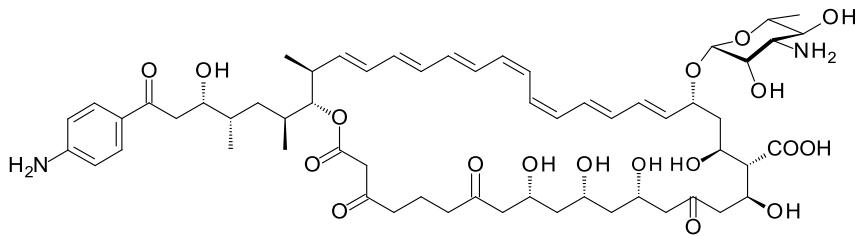
**B**

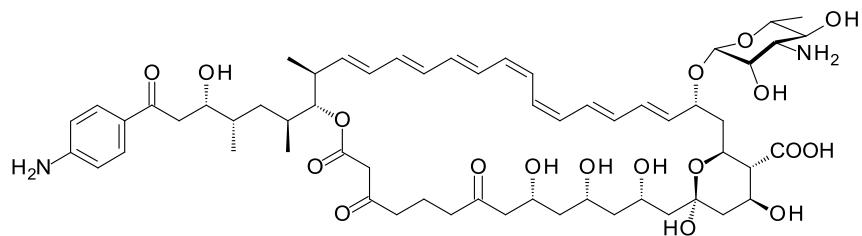
Fig. S8 LCMS analysis of polyenes from *C. caeruleus*. A. Section of HPLC chromatogram showing heptaene peaks. B. Mass spectrum of major peak 1 showing doubly protonated and singly protonated 67-121C ions.



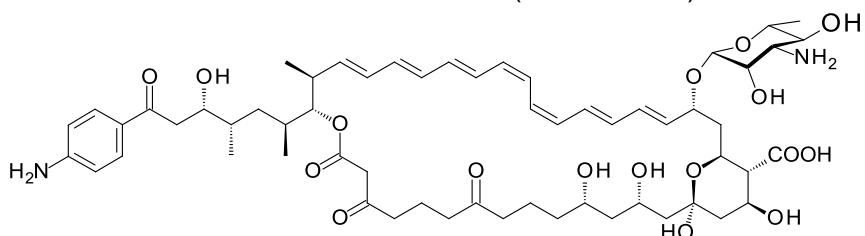
Candidin I (candidin A3)



Candidin II



Candidin III (candidin D)



Candidin IV (candidin A1)

Fig. S9 Main components of the candidin complex.

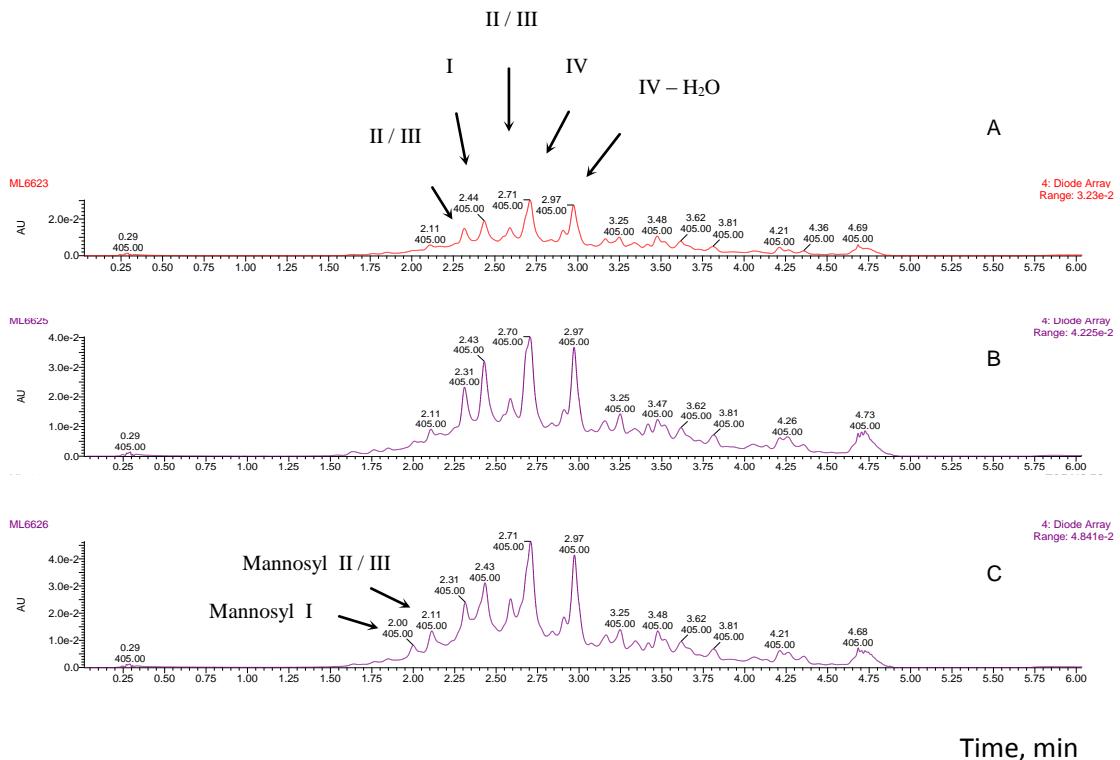


Fig. S10 LC-MS analysis of candididins from (A) *S. albidoflavus* pIAGO, (B) *S. albidoflavus* pIAGO-aceS, (C) *S. albidoflavus* pIAGO-pegA+aceS. The compounds were identified from mass spectra, candididins I, II, III and IV, and mannosylated forms in (C). Candididins II and III have the same mass and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

Table S3 Masses of candididins detected in Fig S10.

Candididin	Mass	[M + H] ⁺ observed
Candididin I	1110.6	1111.6
Candididin II	1108.6	1109.6
Candididin III	1108.6	1109.6
Candididin IV	1092.6	1093.6
Mannosyl-candididin I	1272.6	1273.6
Mannosyl-candididin II	1270.6	1271.6
Mannosyl-candididin III	1270.6	1271.6
Mannosyl-candididin IV	1254.6	Not detected

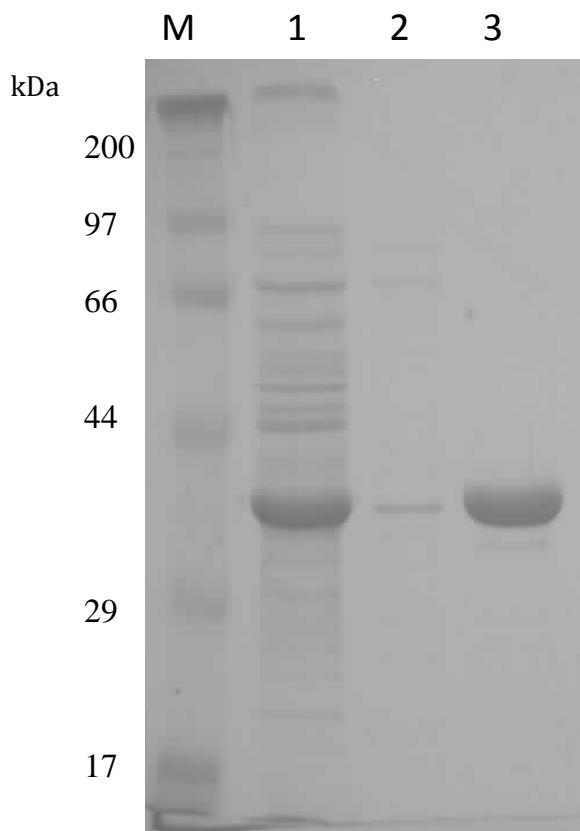


Fig. S11 SDS-PAGE analysis of purified AceS methyltransferase. Lanes: M = Protein molecular weight markers; 1 = soluble fraction from *E. coli* BL32 DE3 pET28-MetN, 2 = AceS after purification on a Ni-NTA column; 3 = concentrated purified AceS protein.

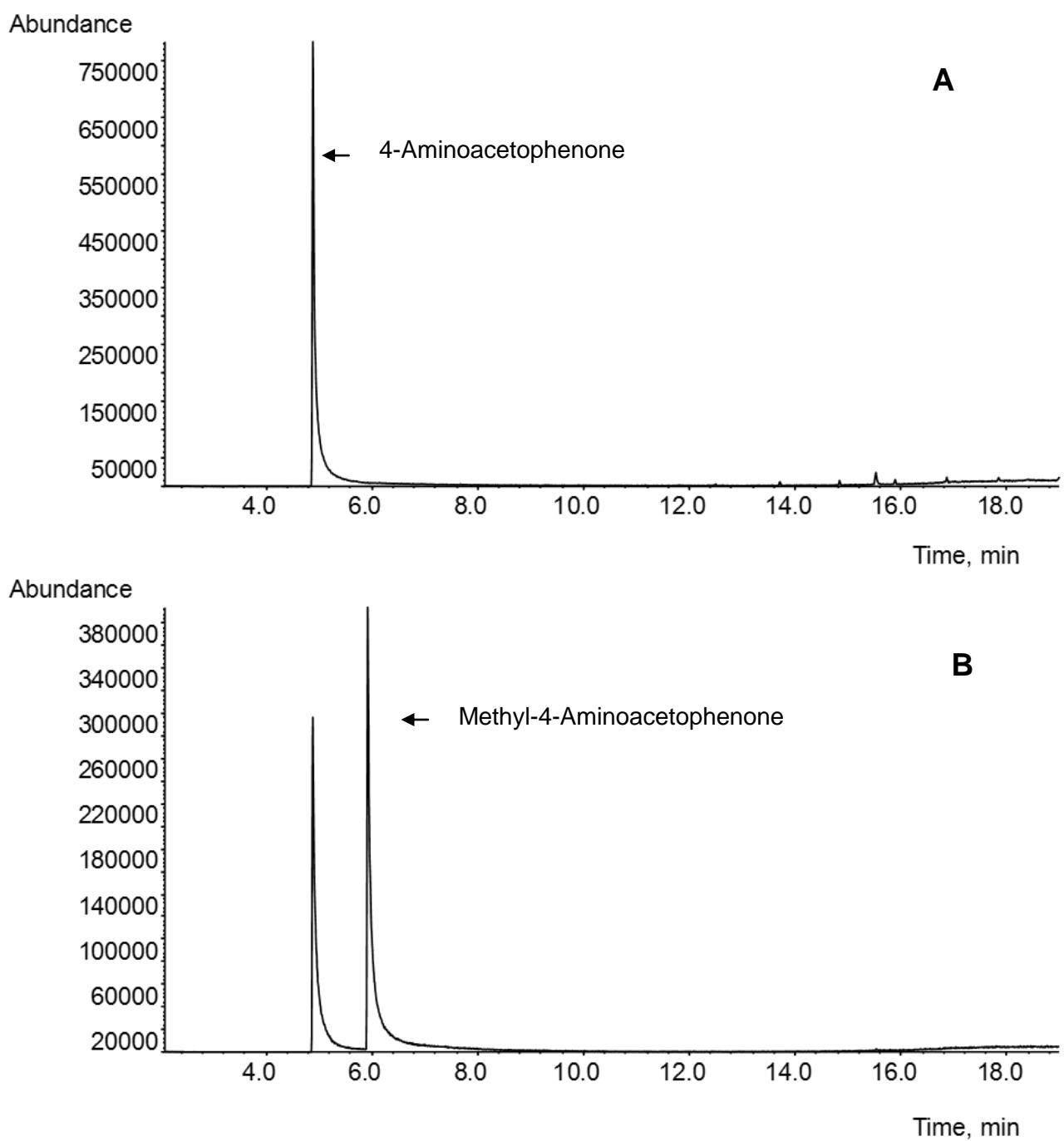


Fig. S12 GC-MS of AceS-catalysed methylation of 4-aminoacetophenone. A. Chromatogram from analysis of 4-aminoacetophenone control. B. Chromatogram from analysis of reaction mixture after 120 minute incubation.

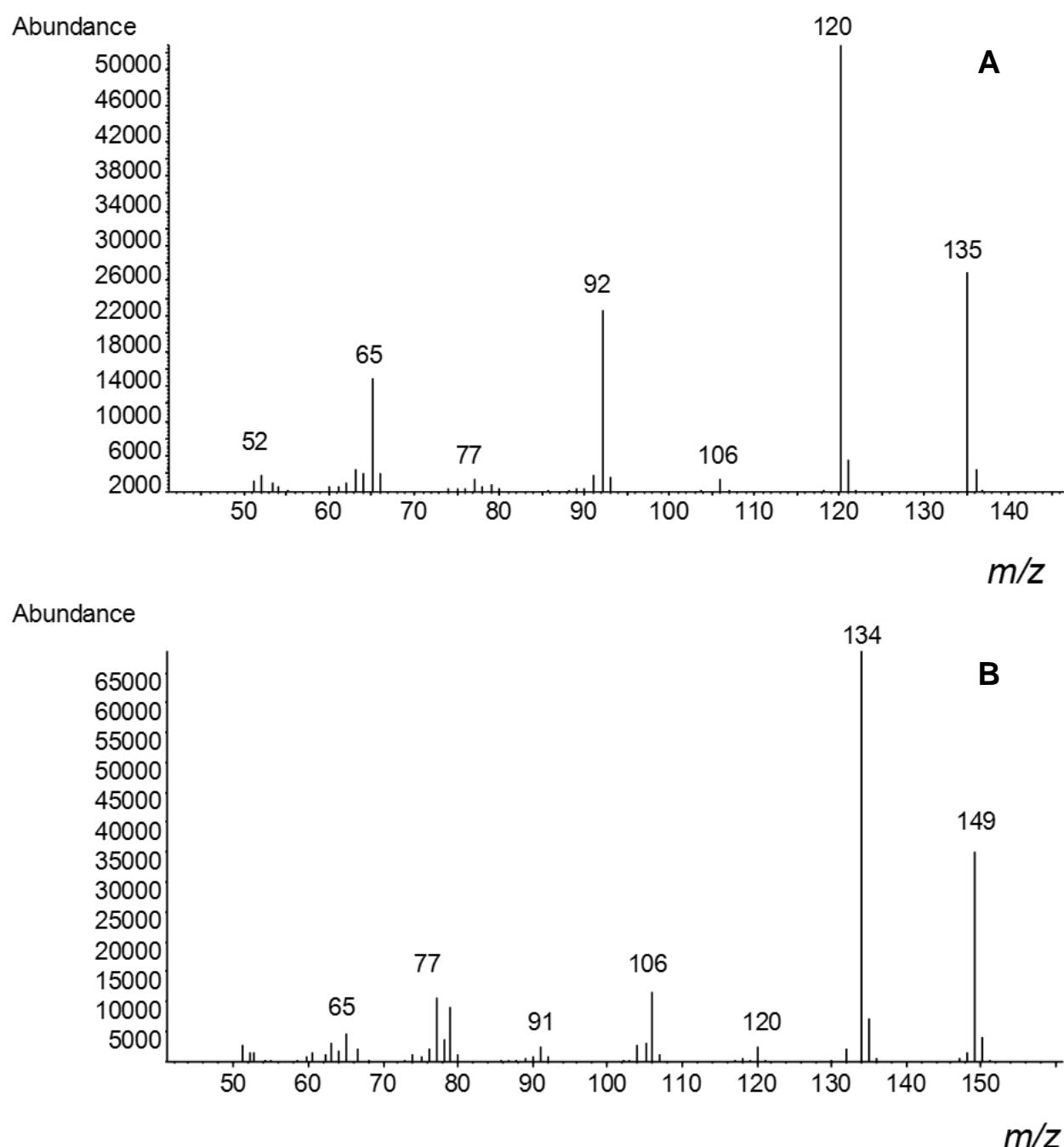


Fig. S13 GC-MS of AceS-catalysed methylation of 4-aminoacetophenone. A. Mass spectrum of 4-aminoacetophenone control peak. B. Mass spectrum of methyl-4-aminoacetophenone peak.

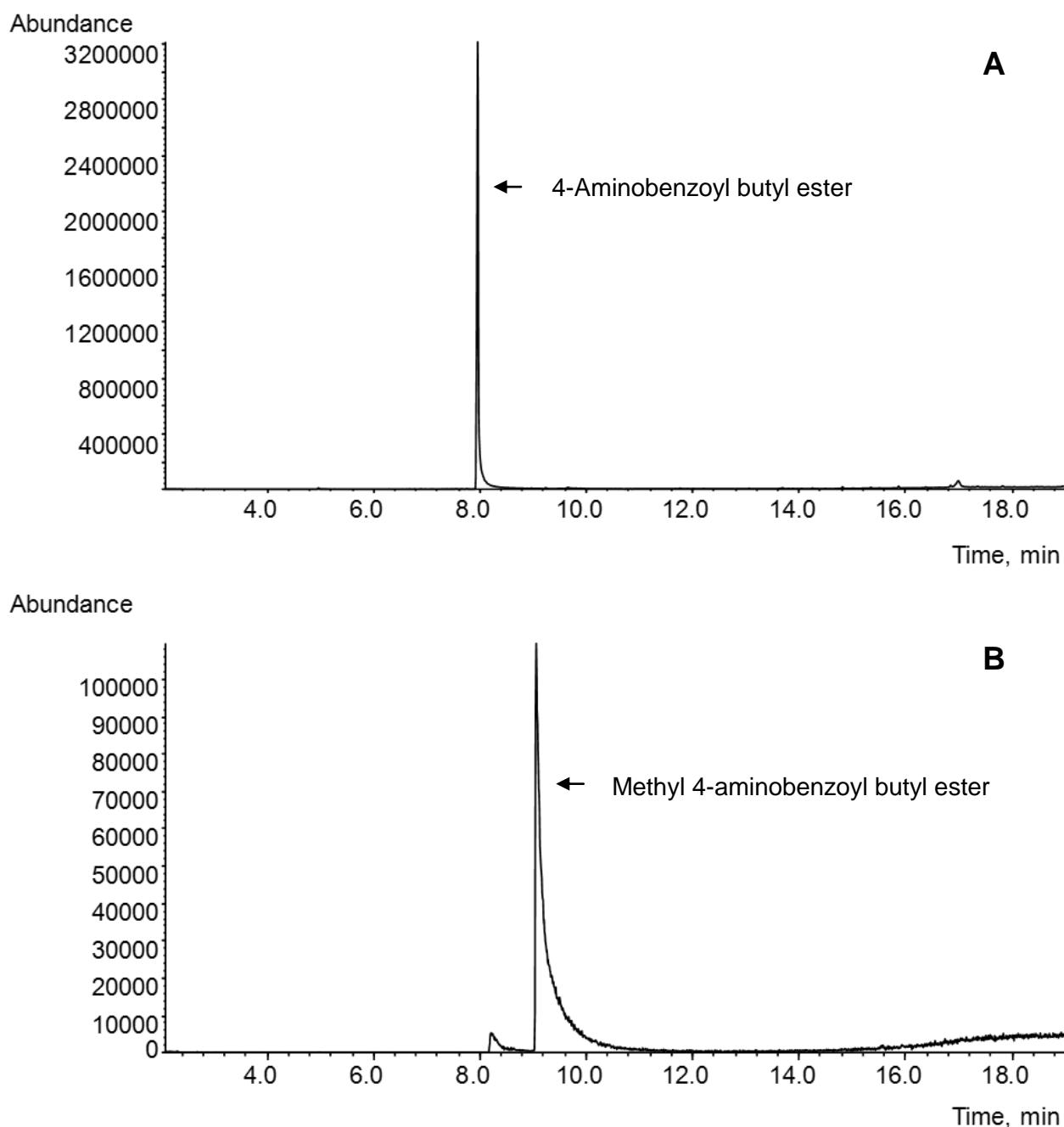


Fig. S14 GC-MS of AceS-catalysed methylation of 4-aminobenzoyl butyl ester. A. Chromatogram from analysis of 4-aminobenzoyl butyl ester control. B. Chromatogram from analysis of reaction mixture after 120 minute incubation.

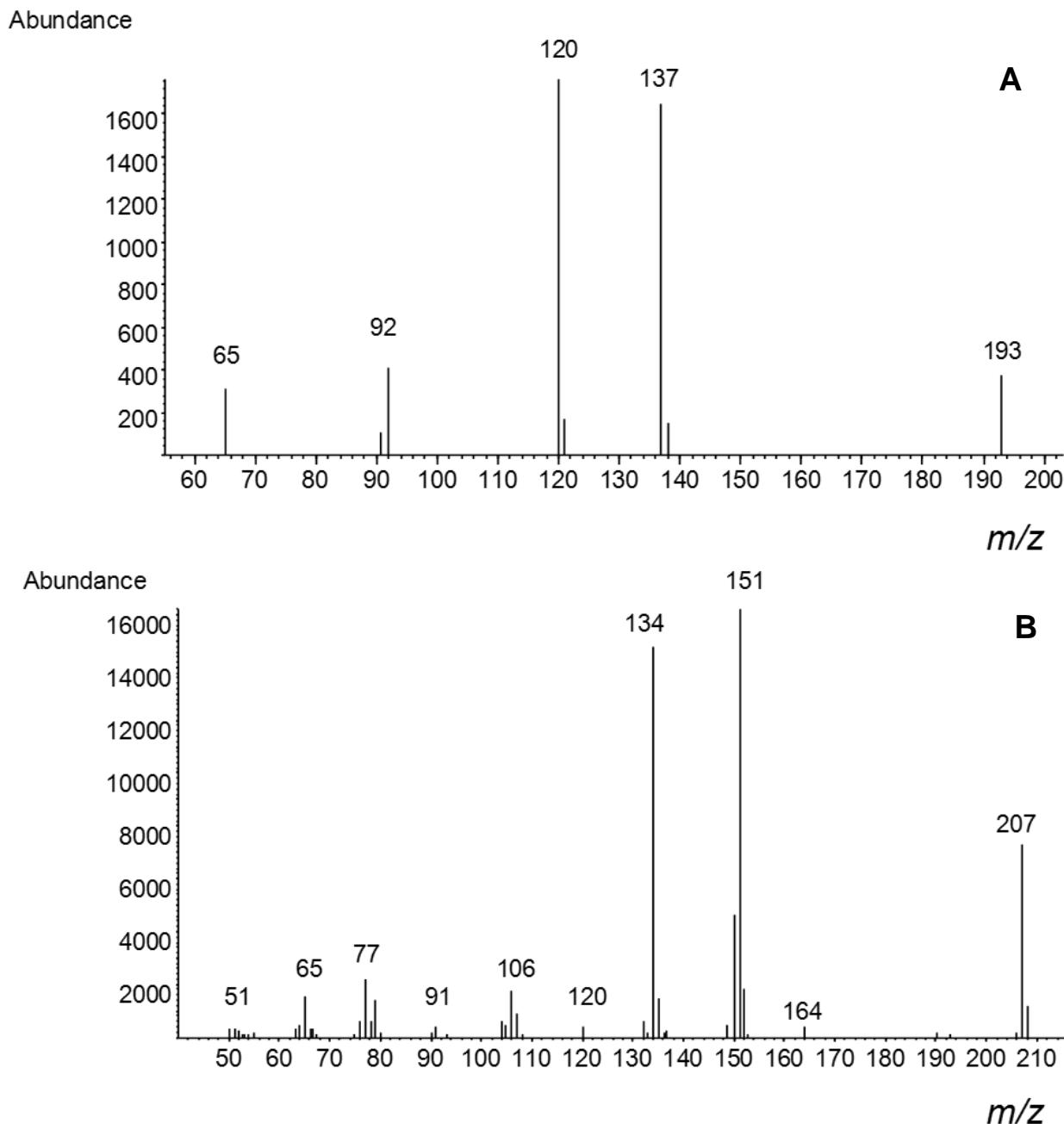


Fig. S15 GC-MS of AceS-catalysed methylation of 4-aminoacetophenone. A. Mass spectrum of 4-aminobenzoyl butyl ester control peak. B. Mass spectrum of methyl-4-aminobenzoyl butyl ester peak.