Supplementary Information

A dual transacylation mechanism for polyketide synthase chain release in enacyloxin antibiotic biosynthesis

Joleen Masschelein[1,†,‡], Paulina K. Sydor[1,†,‡], Christian Hobson[1], Rhiannon Howe[1], Cerith Jones[2,‡], Douglas M. Roberts[1], Zhong Ling Yap[2], Julian Parkhill[3], Eshwar Mahenthiralingam[2] and Gregory L. Challis[1,4,5,*]

[1]Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom

[2]Cardiff School of Biosciences, Cardiff University, Cardiff CF10 3AX, United Kingdom

[3]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom

[4]Warwick Integrative Synthetic Biology Centre, University of Warwick, Coventry CV4 7AL, United Kingdom

[5]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia

[†]These authors contributed equally to this work.

[‡]Current addresses: Rega Institute for Medical Research, KU Leuven, Herestraat 49, PO Box 1041, 3000 Leuven, Belgium (J.M.); Syngenta, Jealott's Hill International Research Centre, Bracknell, Berkshire, RG42 6EY, UK (P.K.S.); Faculty of Computing, Engineering and Science, University of South Wales, Pontypridd, CF37 1DL, United Kingdom (C.J.).

*Author to whom correspondence should be addressed.

TABLE OF CONTENTS

Methods

All reagents and chemicals were purchased from Sigma Aldrich unless otherwise stated.

Genome sequencing. Genomic DNA was prepared from *B. ambifaria* BCC0203 as previously described[1] and prepared for sequencing using a DNA template preparation kit 2.0 (Pacific Biosciences). The genome sequence was assembled from data obtained from two SMRT cells with a PacBio RSII sequencer using SMRTanalysis v2.3.0 (Pacific Biosciences).

Gene deletion and complementation. In-frame deletions in *bamb_5915* and *bamb_5917* were introduced via double homologous recombination using the suicide plasmid pGPI and the I-*Sce*I expression plasmid pDAI.[2] Briefly, the sequences (500-1000 bp) flanking the gene regions targeted for deletion were amplified from *B. ambifaria* BCC0203 genomic DNA using Q5 DNA polymerase (NEB) and the primers listed in **Supplementary Table 1**. Restriction sites were introduced at the 5'-end of the *bamb_5915* primers to allow for directional cloning of the PCR products into pGPI. The amplified regions flanking *bamb_5917* were cloned into pGPI via Gibson Assembly. Constructs were mobilized into *E. coli* SY327 by electroporation (*bamb_5915*) or using chemically competent cells (*bamb_5917*), and transformants were selected on LB agar plates supplemented with trimethoprim (50 μg/mL). Plasmids were purified from trimethoprim-resistant colonies using the GeneJET Plasmid Miniprep kit (Thermo Scientific; *bamb_5915*), or the Qiagen Miniprep Kit (Qiagen; *bamb_5917*), and correct assembly of the mutagenesis constructs was confirmed by Sanger sequencing (GATC Biotech (*bamb_5915*); Eurofins (*bamb_5917*)). Validated constructs were transferred into *B. ambifaria* BCC0203 via triparental mating[2] and transconjugants were selected using trimethoprim (200 μg/ml) and polymyxin (600 U/ml). Single *B. ambifaria* mutants were selected and correct integration of the mutagenesis plasmids into the genome was confirmed by colony PCR. Next, the pDAI plasmid was introduced into the *B. ambifaria* single crossover mutants by triparental mating using *E. coli* SY327 (pDAI) and *E. coli* HB101 (pRK2013) as the donor and helper strain, respectively.[2] Transconjugants were selected on LB agar plates containing tetracycline (200 μg/ml) and polymyxin (600 U/ml). Single *B. ambifaria* mutants were selected and correct gene deletion was confirmed by colony PCR and Sanger sequencing. Finally, the deletion mutants were cured of pDAI by plating the cells on M9 minimal medium[3] containing 10% sucrose. To examine the effect of the gene deletions on enacyloxin biosynthesis, mutant strains were grown at 30°C on solid minimal medium containing glycerol as a sole carbon source (BSM-G).[4] Following incubation for 3 days, the cells were scraped off and ethyl acetate extracts of the agar were analyzed by UHPLC-ESI-Q-TOF-MS as described below.

For genetic complementation, *bamb_5915* and *bamb_5917* were amplified from *B. ambifaria* BCC0203 genomic DNA using Q5 DNA polymerase (NEB) and the primers listed in **Supplementary Table 1**. The appropriate PCR products were purified, digested with *Eco*RI and *Xba*I, and cloned into similarly digested expression vector pMLBAD.[5] The resulting constructs were introduced into *E. coli* SY327 using electroporation or chemically competent cells (see

above), and transformants were selected on LB plates supplemented with trimethoprim (50 μg/ml). Plasmids were purified from trimethoprim-resistant colonies using the GeneJET Plasmid Miniprep kit as described above and the validity of the expression constructs was confirmed by Sanger sequencing. Correct constructs were transferred into the corresponding *B. ambifaria* deletion mutants via triparental conjugation.[2] Single mutants carrying complementation plasmids were selected using trimethoprim (200 μg/ml) and polymyxin (600 U/ml). Their ability to restore enacyloxin production was confirmed by UHPLC-ESI-Q-TOF-MS analysis of ethyl acetate extracts from BSM-G agar-grown cultures. Control experiments were performed with *B. ambifaria* mutants carrying empty pMLBAD plasmids.

Gene cloning and site-directed mutagenesis. Protein coding sequences were amplified by PCR using the primers and templates listed in **Supplementary Table 1**. Domain boundaries were defined by sequence alignments and secondary structure predictions using Phyre[2].[6] A CACC sequence was introduced at the 5'-end of the forward primers to allow for directional cloning of the blunt-ended PCR products into pET151/D-TOPO. The PCR reaction was performed with Expand High Fidelity DNA Polymerase (Roche) according to the manufacturer's instructions. PCR products were purified using the GeneJET Gel Extraction kit (Thermo Scientific) and ligated with the linearized expression vector using the Champion pET151 Directional TOPO Expression kit (Invitrogen) according to the manufacturer's guidelines. The ligation mixtures were then used to transform One Shot TOP10 chemically competent *E. coli* cells (Life Technologies). Transformants were selected on LB agar plates supplemented with ampicillin (50 μg/mL). Plasmids were purified from ampicillin-resistant colonies using the GeneJET Plasmid Miniprep kit (Thermo Scientific) and the sequences of the cloned genes were confirmed by Sanger sequencing (GATC Biotech). One correct clone was used to transform BL21 Star(DE3) or C43(DE3) chemically competent *E. coli* cells for expression of the coding sequences as N-terminal His_6-tagged fusion proteins (**Supplementary Table 2**).

The H205A and C1988A mutations in Bamb_5915 and the KS^0 domain of Bamb_5919, respectively, were introduced using the QuickChange XL Site Directed Mutagenesis Kit (Agilent Technologies), according to the manufacturer's instructions. The primers are listed in **Supplementary Table 1**. The presence of the desired mutations was confirmed by Sanger sequencing (GATC Biotech). CD spectroscopy showed that the mutations did not significantly affect the secondary structure of the proteins (**Supplementary Fig. 15**)

Overproduction, purification and analysis of enacyloxin biosynthetic proteins. For overproduction of the enacyloxin biosynthetic proteins, the *E. coli* BL21 Star (DE3) or C43 (DE3) cells carrying the appropriate expression constructs (**Supplementary Table 1**) were cultured in LB medium supplemented with ampicillin (50 μg/mL) at 37 °C with shaking at 180 rpm. Incubation was continued until the optical density of the cultures at 600 nm reached 0.5-0.6, at which time isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM to induce expression. The cell cultures were then incubated overnight at 15 °C with shaking at 180 rpm.

For purification of the His_6-tagged proteins, the *E. coli* cells were harvested by centrifugation (5000 g, 15 min, 4 °C) and cell pellets were resuspended in 20 mL of binding buffer (20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 20 mM imidazole, 10% glycerol). The cells were subsequently lysed with a high-pressure cell disruptor (Constant Systems Ltd. TS-Series Cabinet) at 20300 psi. After removal of the cellular debris by centrifugation (15000 g, 30 min, 4 °C), the cell-free lysate was applied to a 1 mL HiTrap HP affinity column (GE Healthcare) equilibrated with binding buffer. Unbound proteins were removed by washing the column with 20 mL binding buffer. The recombinant His_6-tagged proteins were eluted with 4 mL elution buffer (20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 200 mM imidazole, 10% glycerol). Fractions were analyzed by SDS-PAGE (**Supplementary Fig. 4**) and those containing the His_6-tagged proteins were pooled, concentrated and exchanged into storage buffer (20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10% glycerol) using Amicon Ultra centrifugal filters with a 10 or 30 kDa molecular weight cut-off (MWCO) membrane (Millipore) (**Supplementary Table 2**). The proteins were aliquoted, flash frozen in liquid N_2 and stored at -80 °C. Protein concentrations were determined by measuring the absorbance at 280 nm on a NanoDrop spectrophotometer (Thermo Scientific) and using the calculated extinction coefficients listed in **Supplementary Table 2**. UHPLC-ESI-Q-TOF-MS (see below) was used to confirm the molecular weights of the purified proteins (**Supplementary Fig. 4**).

***In vitro* conversion of *apo*-ACPs/PCPs to *holo*- or acylated *holo*-ACPs/PCPs.** *Apo*-carrier proteins (the Bamb_5919_ACP domain, the Bamb_5919_ACP-KS^0 di-domain, Bamb_5917 and the Bamb_5917_PCP domain) were converted into their *holo*- or acylated *holo*-forms using the promiscuous phosphopantetheinyl transferase (PPTase) Sfp from *B. subtilis*, which was overproduced and purified as reported previously.[7] (2*E*, 4*E*)-2, 4-pentadienoyl-CoA was generated in situ from the corresponding pantetheine thioester (synthesised as described below) using purified recombinant PanK, PPAT and DPCK.[8] The *apo*-proteins (200 μM) were combined with $MgCl_2$ (12.5 mM), Sfp (10 μM) and CoA, or acyl-CoA (0.5 mM), respectively, in a total volume of 200 μL. Reactions were carried out in buffer containing 20 mM Tris-HCl (pH 7.5) and 100 mM NaCl. Following incubation at 30 °C for 45 min, excess (acyl-)CoA was removed using Amicon Ultra centrifugal filters with a 5 kDa MWCO membrane (Millipore). Conversion of the proteins to the *holo* and acetylated *holo*-forms, respectively, was verified by UHPLC-ESI-Q-TOF-MS analysis (**Supplementary Fig. 5 and 6**).

Analysis of acetyl transfer between the ACP and KS^0 domains of Bamb_5919. 250 μM of the acetylated Bamb_5919 ACP domain was combined with 25 μM of the Bamb_5919 KS^0 domain, or its C1988A mutant, in buffer containing 25 mM Tris (pH 7.5) and 0.5 M NaCl. Following overnight incubation at 25 °C, the reactions were diluted 2.5-fold with deionised water and the extent of KS^0 domain acetylation was determined using UHPLC-ESI-Q-TOF-MS, as described below. Each experiment was performed in triplicate.

Analysis of KS^0 domain-catalysed acetyl transfer from the Bamb_5919 ACP domain to the Bamb_5917 PCP domain. The acetylated Bamb_5919 ACP domain (100 μM) and *holo*-

Bamb_5917 or its excised PCP domain (100 μM) were combined with the Bamb_5919 KS^0 domain (25 μM) and reaction buffer (25 mM Tris (pH 7.5), 0.5 M NaCl) in a total volume of 100 μl. The Bamb_5919 KS^0 domain was omitted, or replaced with the C1988A mutant (25 μM), in negative control reactions. When the acetylated Bamb_5919 ACP-KS^0 di-domain was used in place of the isolated KS^0 and ACP domains, a concentration 100 μM of was used. All reactions were incubated at 30°C for 3 hours, then diluted 10-fold with deionised water prior to UHPLC-ESI-Q-TOF-MS analysis (see below). Each experiment was carried out in triplicate.

Carrier protein, acyl donor and acyl acceptor specificity of Bamb_5915. To investigate the carrier protein specificity of Bamb_5915, the acetylated Bamb_5919 ACP and Bamb_5917 PCP domains and acetylated full-length Bamb_5917 (200 μM) were separately incubated with Bamb_5915 (20 μM) and DHCCA or AHCCA (1 mM), in a total volume of 100 μl. Reactions were carried out in buffer containing 20 mM Tris-HCl (pH 7.5) and 100 mM NaCl. Bamb_5915 was omitted from negative control reactions. Following incubation for 1 h at 30 °C, the reactions were stopped by the addition of two volumes of methanol. Precipitates were removed by centrifugation (13000 g, 1 min) and the supernatants were analysed by UHPLC-ESI-QTOF-MS (see below). All assays were performed in triplicate.

The acyl acceptor specificity of Bamb_5915 was examined by incubating the enzyme (20 μM) with the acetylated Bamb_5917 PCP domain or full-length protein (200 μM) and a range of DHCCA analogues (1 mM), in a total volume of 100 μl of buffer (20 mM Tris-HCl (pH 7.5), 100 mM NaCl). Bamb_5915 was omitted from negative control reactions. After incubation for 1 h at 30 °C, the proteins were precipitated by the addition of two volumes of methanol. Precipitates were removed by centrifugation (13000 g, 1 min) and the supernatants were analysed by UHPLC-ESI-QTOF-MS (see below).

The acyl donor specificity of Bamb_5915 was investigated similarly, except that a range of NAC thioesters (1.5 mM) and AHCCA (1 mM) were used instead of the acetylated Bamb_5917 PCP domain/full-length protein and the DHCCA analogues, respectively.

***In vitro* reconstitution of chain release from the enacyloxin PKS**. The acetylated Bamb_5919 ACP domain (25 μM) was combined with the *holo*-Bamb_5917 PCP domain/full-length protein (25 μM), the Bamb_5919 KS^0 domain (25 μM), Bamb_5915 (20 μM) and AHCCA (1 mM) in a total volume of 100 μl. Reactions were carried out in buffer containing 20 mM Tris-HCl (pH 7.5) and 100mM NaCl. To evaluate the role played by the Bamb_5919 KS^0 domain in the overall process, it was omitted from the reaction or replaced with the C1988A mutant. In addition, the acetylated Bamb_5919_ACP-KS^0 di-domain was employed in place of the isolated ACP and KS^0 domains at a concentration of 25 μM. A negative control reaction from which Bamb_5915 was omitted was also carried out. All reactions were incubated for 1h at 30 °C and stopped by the addition of two volumes of methanol. Precipitates were removed by centrifugation (13000 g, 1 min) and the supernatants were analysed by UHPLC-ESI-QTOF-MS (see below). All experiments were performed in triplicate.

UHPLC-ESI-Q-TOF-MS analyses. Small molecules were analyzed using a Dionex UltiMate 3000 UHPLC connected to a Zorbax Eclipse Plus column (C_{18}, 100 × 2.1 mm, 1.8 μm) coupled to a Bruker maXis impact mass spectrometer. Mobile phases consisted of water (A) and acetonitrile (B), each supplemented with 0.1% formic acid. The column was eluted with a linear gradient of 5 - 100% B over 35 min, employing a flow rate of 0.2 mL/min. The mass spectrometer was operated in positive ion mode with a scan range of 50-3000 *m/z*. Source conditions were: end plate offset at −500 V; capillary at −4500 V; nebulizer gas (N_2) at 1.6 bar; dry gas (N_2) at 8 L min−1; dry temperature at 180 °C. Ion transfer conditions were: ion funnel RF at 200 Vpp; multiple RF at 200 Vpp; quadrupole low mass at 55 *m/z*; collision energy at 5.0 eV; collision RF at 600 Vpp; ion cooler RF at 50–350 Vpp; transfer time at 121 μs; pre-pulse storage time at 1 μs. Calibration was performed with 1 mM sodium formate through a loop injection of 20 μL at the start of each run.

UHPLC-ESI-Q-TOF-MS analyses of intact proteins were conducted using a Dionex UltiMate 3000 RS UHPLC connected to an ACE 3 C_4-300 reverse phase column (Advanced Chromatography Technologies, Aberdeen, UK; 100 × 2.1 mm) coupled to a Bruker maXis II mass spectrometer. Proteins were eluted with a linear gradient of 5 - 100% MeCN containing 0.1% formic acid over 30 min. A flow rate of 0.2 ml/min was employed. Source conditions were: capillary at -4500 V, ion polarity positive, time-of-flight tube at 9860 V and detector at 3500 V.

CD spectroscopy. The CD spectra of Bamb_5915, the Bamb_5919 KS^0 domain and their mutated variants were measured from 180 to 260 nm in buffer consisting of 50 mM potassium phosphate (pH 7.4) and 50 mM NaCl, using a JASCO J-1500 spectrometer and a 1 mm path length quartz cuvette. The spectra were recorded at 20°C using a protein concentration of 0.2 mg/ml.

Synthesis of substrates and substrate analogues

General procedures

Room temperature refers to ambient temperature (20 - 22 °C), 5 ˚C refers to a cold water bath and 0 °C refers to an ice slush bath. Heated experiments were conducted using thermostatically controlled oil baths. All commercially available solvents and chemicals were used without any further purification. NMR spectra were recorded on a Bruker Avance AV-300 MHz spectrometer or a Bruker HD-500 MHz spectrometer equipped with a DCI cryoprobe at room temperature (298 K). Chemical shifts are reported in parts per million (ppm) referenced to $CHCl_3$ (δ_H: 7.26 ppm and δ_C: 77.0 ppm). Coupling constants (*J*) are rounded to the nearest 0.5 Hertz (Hz). Multiplicities are given as multiplet (m), singlet (s), doublet (d), triplet (t), quartet (q), quintet (quin.), sextet (sext.), septet (sept.), octet (oct.) and nonet (non.). ^1H and ^{13}C assignments were established on the basis of COSY, DEPT, HMQC and HMBC correlations. Infra-red spectra were recorded using either a Perkin Elmer Spectrum 100 FT-IR spectrometer or an Alpha Bruker Platunium ATR single reflection diamond ATR module.

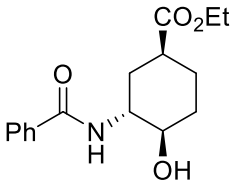
Melting points were recorded on a Stuart scientific melting point apparatus and are uncorrected. Silica column chromatography was performed on 40-60 Å silica gel. Thin layer chromatography (TLC) was carried out on aluminium sheets coated with 0.2 mm silica gel 60 F_{254}. Visualisation was effected by UV light (254 nm) or by potassium permanganate solution followed by heating. Low resolution mass spectra (LRMS) were recorded using an Agilent 6130B single Quad (ESI). High resolution mass spectra (HRMS) were obtained using a Bruker MaXis mass spectrometer.

Synthesis of (±)-(1R,3R)-3-hydroxycyclohexanecarboxylic acid and (±)-(1S,3R)-3-hydroxycyclohexanecarboxylic acid

These compounds were synthesized according to literature procedures.[9]

Synthesis of (±)-(1S,3R,4S)-3-amino-4-hydroxycyclohexane carboxylic acid (AHCCA), (±)-(1S,3R,4S)-3-acetamido-4-hydroxycyclohexane-1-carboxylic acid and (±)-(1S,3R,4R)-3-amino-4-hydroxycyclohexane-1-carboxylic acid

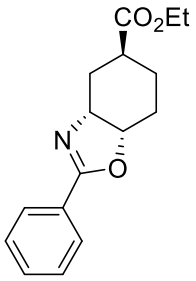
(±)-Ethyl (1*S*,3*R*,4*R*)-3-benzamido-4-hydroxycyclohexane-1-carboxylate



To a solution of ethyl (±)-(1*S*,3*S*,6*R*)-7-oxabicyclo[4.1.0]heptane-3-carboxylate[10] (1.15 g, 6.76 mmol) in EtOH (10 mL) was added 28 % aqueous ammonia (20 mL). The reaction was stirred at 45 ^{o}C overnight and concentrated *in vacuo* to afford the crude amino alcohol as a yellow oil. The oil was dissolved in EtOH (20 mL) and benzoic anhydride (1.68 g, 7.44 mmol) was added in an ice-water bath. After 15 min the reaction was allowed to warm to room temperature and was stirred for a further 2 h. The reaction was then concentrated *in vacuo*, and the crude oil purified by silica chromatography (EtOAc : Petroleum ether, 30 : 70) to afford the product as a white solid (1.75 g, 89 %, m.p. 138-140 $^{\circ}$C).

ν_{max}/cm^{-1} (neat) 3423 (NH), 3310 (OH), 2944 (ArH), 1720, 1627 (C=O), 1025 (C-O); δ_H (500 MHz; $CDCl_3$) 7.76 (2H, d, J 7.5, Ar*H*), 7.51 (1H, t, J 7, Ar*H*), 7.43 (2H, t, J 7.5, Ar*H*), 7.18 (1H, br. d, J 6.5, N*H*), 4.19 (2H, q, J 7.5, OCH_2CH$_3$), 4.12-4.05 (1H, m, C*H*N), 3.71 (1H, br. s, O*H*), 3.56 (1H, td, J 9 and 4.5, C*H*OH), 2.72 (1H, quin., J 4, C*H*CO$_2$), 2.50-2.44 (1H, m, CH_2CHN), 2.22-2.14 (1H, m, CH_2CHCO$_2$), 1.99-1.91 (1H, m, CH_2CHOH), 1.73 (1H, br. s, O*H*), 1.63-1.54 (3H, m, CH_2CHN, CH_2CHCO$_2$ and CH_2CHOH), 1.29 (3H, t, J 7, CH_3); δ_C (125 MHz, $CDCl_3$) 174.0 (*C*O$_2$CH$_2$), 169.3 (*C*ON), 134.2 (Ar$C_{quart.}$), 132.0 (Ar*C*), 128.8 (Ar*C*), 127.2 (Ar*C*), 74.2 (*C*HO), 60.9 (*C*H$_2$CH$_3$), 53.1 (*C*HN), 38.9 (*C*HCO$_2$), 31.5 (*C*H$_2$CHN), 30.7 (*C*H$_2$CHOH), 25.0 (*C*H$_2$CHCO$_2$), 14.4 (*C*H$_3$); HRMS (ESI) calc. for $C_{16}H_{21}NNaO_4^+$: 314.1363, found: 314.1363.

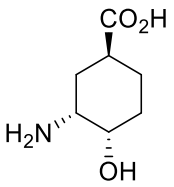
(±)-Ethyl (3a*R*,5*S*,7a*S*)-2-phenyl-3a,4,5,6,7,7a-hexahydrobenzo[d]oxazole-5-carboxylate



To a solution of (±)-ethyl (1*S*,3*R*,4*R*)-3-benzamido-4-hydroxycyclohexane-1-carboxylate (1.5 g, 5.15 mmol) in $CHCl_3$ (30 mL) was added $SOCl_2$ (1.5 mL, 9.82 mmol) dropwise under an inert atmosphere. The reaction was allowed to stir for 1 h at 40 ^oC , saturated $NaHCO_3$ (15 mL) was added and the mixture was stirred for a further ten minutes. The organic layer was separated and the aqueous layer was extracted with $CHCl_3$ (2 x 20 mL). The combined organics were dried ($MgSO_4$), filtered and concentrated *in vacuo* to afford the product as a yellow oil (1.21 g, 86 %).

ν_{max}/cm^{-1} (neat) 2941 (ArH), 1720 (C=O), 1631 (C=N), 1025 (C-O); δ_H (500 MHz; $CDCl_3$) 7.94 (2H, d, J 7.5, Ar*H*), 7.48 (1H, t, J 7.5, Ar*H*), 7.41 (2H, t, J 7.5, Ar*H*), 4.81 (1H, dt., J 9 and 5.5, C*H*OCN), 4.39 (1H, quin., J 4.5, C*H*N), 4.13 (2H, qd, J 7 and 2, OC*H*$_2$CH$_3$), 2.59 (1H, dtd, J 11.5, 7 and 5, C*H*CO_2), 2.24 (1H, dt, J 14.5 and 4, C*H*$_2$CHN), 2.01-1.92 (2H, m, C*H*$_2$CHN and C*H*$_2$CHO), 1.88-1.73 (2H, m, C*H*$_2$CHO and C*H*$_2$CHCO$_2$), 1.68-1.60 (1H, m, C*H*$_2$CHCO$_2$), 1.24 (3H, t, J 7, C*H*$_3$); δ_C (125 MHz, $CDCl_3$) 176.0 (*C*O$_2$CH$_2$), 164.5 (*C*ON), 131.6 (Ar*C*), 128.5 (Ar*C*), 128.3 (Ar*C*), 128.0 (Ar*C*$_{quart.}$), 77.7 (*C*HO), 63.2 (*C*HN), 60.6 (*C*H$_2$CH$_3$), 35.7 (*C*HCO$_2$), 28.6 (*C*H$_2$CHN), 25.0 (*C*H$_2$CHO), 20.9 (*C*H$_2$CHCO$_2$), 14.3 (*C*H$_3$); HRMS (ESI) calc. for $C_{16}H_{20}NO_3^+$: 274.1438, found: 274.1439.

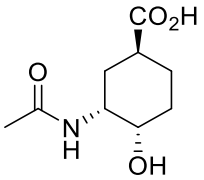
(±)-(1*S*,3*R*,4*S*)-3-amino-4-hydroxycyclohexane-1-carboxylic acid



(±)-Ethyl (3a*R*,5*S*,7a*S*)-2-phenyl-3a,4,5,6,7,7a-hexahydrobenzo[d]oxazole-5-carboxylate (1.20 g, 4.41 mmol) was dissolved in aqueous HCl (6 M, 50 mL) and heated to reflux overnight. The reaction was concentrated *in vacuo*, diluted with water and washed with Et_2O (5 x 50 mL). The aqueous phase was subsequently concentrated *in vacuo* to afford (±)-(1*S*,3*R*,4*S*)-3-amino-4-hydroxycyclohexane-1-carboxylic acid as a white solid (0.65 g, 92 %, m.p. 233-235 $^{\circ}C$).

ν_{max}/cm^{-1} (neat) 3283 (NH_2), 3134 (OH), 2904 (CO_2H), 1697 (C=O), 1040 (C-O); δ_H (500 MHz; D_2O) 4.06 (1H, dt, J 6.5 and 3, C*H*OH), 3.62 (1H, dt, J 8 and 4, C*H*NH_2), 2.79 (1H, quin., J 6 C*H*CO_2H), 2.10-1.91 (3H, m, CHC*H*$_2$CHNH_2, C*H*$_2$CHCO_2H), 1.85-1.67 (3H, m, C*H*$_2$CHCO_2H, C*H*$_2$CHOH); δ_C (125 MHz, D_2O) 178.9 (*C*O_2H), 65.5 (*C*HOH), 50.2 (*C*HNH), 37.1 (*C*HCO_2H), 27.1 (*C*H_2CHOH), 26.3 (*C*H_2CHNH_2), 22.0 (*C*H_2CHCO_2H); HRMS (ESI) calc. for $C_7H_{14}NO_3^+$: 160.0967, found: 160.0968.

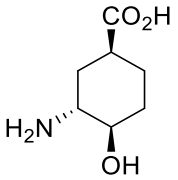
(±)-(1*S*,3*R*,4*S*)-3-acetamido-4-hydroxycyclohexane-1-carboxylic acid



To a solution of (±)-(1*S*,3*R*,4*S*)-3-amino-4-hydroxycyclohexane-1-carboxylic acid (100 mg, 0.63 mmol) in H_2O (2 mL) was added acetic anhydride (75 µL, 0.85 mmol), followed by sodium acetate (200 mg, 2.52 mmol). The resulting mixture was stirred at room temperature for 4 h, then acidified using aqueous HCl (2 M, 2 mL). Concentration *in vacuo* afforded the crude product as a white tacky solid. This was resuspended in MeOH (10 mL), filtered and concentrated *in vacuo* to afford (±)-(1*S*,3*R*,4*S*)-3-acetamido-4-hydroxycyclohexane-1-carboxylic acid as a viscous oil (92 mg, 73 %).

ν_{max}/cm^{-1} (neat) 3136 (OH), 2935 (NH), 1699, 1618 (C=O), 1023 (C-O); δ_H (500 MHz; D_2O) 5.12 (1H, dt, J 6 and 3, C*H*OH), 3.77 (1H, dt, J 9.5 and 3.5, C*H*NH), 2.82 (1H, quin., J 5, C*H*CO_2H), 2.19-2.14 (1H, m, C*H*$_2$CHNH), 2.16 (3H, s, C*H*$_3$CON), 2.07 (1H, ddd, J 14, 10 and 5, C*H*$_2$CHNH), 1.96-1.75 (4H, m, C*H*$_2$CHCO_2H, C*H*$_2$CHOH); δ_C (125 MHz, D_2O) 177.4 (*C*O_2H), 171.2 (*C*ONH), 67.7 (*C*HOH), 46.5 (*C*HNH), 35.9 (*C*HCO_2H), 24.8 (*C*H_2CHNH), 22.5 (*C*H_2CHOH), 19.8 (*C*H_2CHCO_2H), 18.4 *C*H_3; HRMS (ESI) calc. for $C_9H_{16}NO_4^+$: 202.1074, found: 202.1075).

(±)-(1*S*,3*R*,4*R*)-3-amino-4-hydroxycyclohexane-1-carboxylic acid



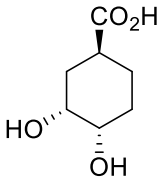
(±)-Ethyl (1*S*,3*R*,4*R*)-3-((*tert*-butoxycarbonyl)amino)-4-hydroxycyclohexane-1-carboxylate[11] (1.35 g, 4.70 mmol) was dissolved in aqueous HCl (6 M, 50 mL) and heated to reflux overnight. After cooling, the mixture was concentrated *in vacuo* to afford (±)-(1*S*,3*R*,4*R*)-3-amino-4-hydroxycyclohexane-1-carboxylic acid as a viscous oil (0.73 g, 98 %).

ν_{max}/cm^{-1} (neat) 3277 (NH_2), 3145 (OH), 2912 (CO_2H), 1698 (C=O), 1037 (C-O); δ_H (500 MHz; D_2O) 4.12-4.09 (1H, m, C*H*OH), 3.40 (1H, dt, J 12.5 and 3.5, C*H*NH_2), 2.55 (1H, tt, J 12.0 and 3.5, C*H*CO_2H), 2.06-1.93 (2H, m, CHC*H*$_2$CHNH_2, C*H*$_2$CHOH), 1.83-1.76 (2H, m, C*H*$_2$CHCO_2H,

CH_2CHOH), 1.67-1.57 (2H, m, CHCH_2CHNH_2, CH_2CHCO_2H); δ_C (125 MHz, D_2O) 178.6 (*C*O_2H), 64.3 (*C*HOH), 51.6 (*C*HNH), 40.5 (*C*HCO_2H), 29.8 (*C*H_2CHOH), 25.9 (*C*H_2CHNH_2), 21.1 (*C*H_2CHCO_2H); HRMS (ESI) calc. for $C_7H_{14}NO_3^+$: 160.0967, found: 160.0968.

Synthesis of (±)-(1S,3R,4S)-3,4-dihydroxycyclohexane carboxylic acid (DHCCA), (±)-(1R,3R,4S)-3,4-3,4-dihydroxycyclohexane carboxylic acid, (±)-(1R,3R,4R,5R)-3,4,5-trihydroxycyclohexane-1-carboxylic acid and (±)-(1S,3R,4S,5R)-3,4,5-trihydroxycyclohexane-1-carboxylic acid

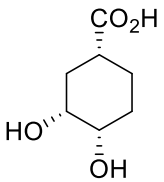
(±)-(1*S*,3*R*,4*S*)-3,4-dihydroxycyclohexane-1-carboxylic acid



To a solution of the acetonide of (±)-(1*S*,3*R*,4*S*)-3,4-dihydroxycyclohexane-1-carboxylic acid[12] (350 mg, 1.64 mmol) in MeOH (30 mL) was added TsOH (35 mg, 0.20 mmol). The resulting mixture was stirred at room temperature for 1 h, then concentrated *in vacuo* and partitioned between EtOAc (20 mL) and saturated $NaHCO_3$ (20 mL). The layers were separated and the aqueous phase was extracted with EtOAc (2 x 20 mL). The combined organics were washed with brine, dried ($MgSO_4$), filtered and concentrated *in vacuo* to afford a viscous oil, which was dissolved in THF (10 mL). H_2O (5 mL) and LiOH (40 mg, 2 mmol) were added and the resulting mixture was stirred at room temperature overnight. The THF was removed *in vacuo* and the resulting suspension was acidified and extracted with EtOAc (5 x 10 mL). The combined organics were washed with brine, dried ($MgSO_4$), filtered and concentrated *in vacuo* to afford the (±)-(1*S*,3*R*,4*S*)-3,4-dihydroxycyclohexane-1-carboxylic acid as a viscous oil (192 mg, 73 %).

ν_{max}/cm^{-1} (neat) 3147 (OH), 2935 (CO_2H), 1718 (C=O), 1066 (C-O); δ_H (500 MHz; D_2O) 4.05-4.01 (1H, br. m, CHCH_2C*H*OH), 3.69 (1H, ddd, J 11, 4 and 3, C*H*OH), 2.55 (1H, tt, J 12 and 3.5, C*H*CO_2H), 2.06-1.99 (1H, m, CHCOHC*H*$_2$CHCO_2H), 1.96-1.90 (1H, m, C*H*$_2$CHCO_2H), 1.76-1.60 (3H, m, CHCOHC*H*$_2$CHCO_2H, C*H*$_2$CHOH), 1.47 (1H, qd, J 12.5 and 4, C*H*$_2$CHCO_2H); δ_C (125 MHz, D_2O) 183.1 (*C*O_2H), 71.2 (*C*HOH), 69.2 (CHCH_2*C*HOH), 38.4 (*C*HCO_2H), 33.9 (CH*C*H_2CHOH), 26.3 (*C*H_2CH_2CHCO_2H), 27.0 (*C*H_2CHCO_2H); HRMS (ESI) calc. for $C_7H_{13}O_4^+$: 161.0814, found: 161.0813.

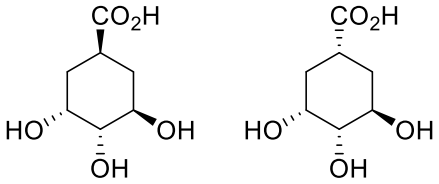
(±)-(1*R*,3*R*,4*S*)-3,4-dihydroxycyclohexane-1-carboxylic acid



A solution of 3,4-dihydroxybenzoic acid (1.5g, 9.8mmol) and rhodium on alumina (5 wt%, 0.2 g) in water (60 ml) was hydrogenated for 2 days at 140 °C and 80 atm. The mixture was cooled to room temperature and filtered. The resulting solution was lyophilized to yield an off-white solid, which purified by reverse-phase HPLC (isocratic 100% water) to yield (±)-(1R,3R,4S)-3,4-dihydroxycyclohexane-1-carboxylic acid as a white solid (630 mg, 42%).

ν_{max}/cm^{-1} (neat) 3152 (OH), 2940 (CO_2H), 1716 (C=O), 1067 (C-O); δ_H (500 MHz; D_2O) 3.96-3.93 (1H, br m, $CHCH_2CHOH$), 3.70 (1H, ddd, J 11, 4 and 3, $CHOH$), 2.37 (1H, tt, J 11 and 3.5, $CHCO_2H$), 1.93-1.84 (2H, m, $CHCOHCH_2CHCO_2H$, CH_2CHCO_2H), 1.76-1.53 (3H, m, $CHCOHCH_2CHCO_2H$, CH_2CHOH); δ_C (125 MHz, D_2O) 183.0 (CO_2H), 71.2 ($CHOH$), 69.2 ($CHCH_2CHOH$), 43.3 ($CHCO_2H$), 31.0 ($CHCH_2CHOH$), 29.7 ($CH_2CH_2CHCO_2H$), 22.8 (CH_2CHCO_2H); HRMS (ESI) calc. for $C_7H_{13}O_4^+$: 161.0814, found: 161.0817.

(±)-(1*R*,3*R*,4*R*,5*R*)-3,4,5-trihydroxycyclohexane-1-carboxylic acid and (±)-(1*S*,3*R*,4*S*,5*R*)-3,4,5-trihydroxycyclohexane-1-carboxylic acid



To a stirred solution of shikimic acid (100 mg, 0.57 mmol) in MeOH (4 mL) was added palladium on carbon (20 mg). The resulting mixture was placed under a H_2 atmosphere and stirred at room temperature for 16 h. It was then filtered through Celite and concentrated *in vacuo* to afford a mixture of the two products, which was separated by semi-preparative reverse phase HPLC to afford (1*R*,3*R*,4*R*,5*R*)-3,4,5-trihydroxycyclohexane-1-carboxylic acid (5 mg, 5 %) and (1*S*,3*R*,4*S*,5*R*)-3,4,5-trihydroxycyclohexane-1-carboxylic acid (4 mg, 4 %) as colourless oils.

(1*R*,3*R*,4*R*,5*R*)-3,4,5-trihydroxycyclohexane-1-carboxylic acid: ν_{max}/cm^{-1} (neat) 3261 (OH), 1701 (CO); δ_H (500 MHz; CD_3OD) 4.06 (1H, q, J 3.5, $CHOHCHOHCHOH$), 3.75 (1H, ddd, J 11.0, 9.5 and 4.5, $CHOHCHOHCHOH$), 3.27 (1H, dd, J 9.0 and 3.0, $CHOHCHOHCHOH$), 2.82 (1H, tt, J 12.5 and 3.5, $CHCO_2H$), 2.16-2.19 (1H, m, $CH_2CHCO_2HCH_2$), 2.03 (1H, dq, J 14.0 and 3.5, $CH_2CHCO_2HCH_2$), 1.61 (1H, ddd, J 14.5, 12.5 and 2.0, $CH_2CHCO_2HCH_2$), 1.41 (1H, q, J 12.5, $CH_2CHCO_2HCH_2$); δ_C (125 MHz, CD_3OD) 177.2 (CO_2H), 77.3 ($CHOHCHOHCHOH$), 70.4 ($CHOHCHOHCHOH$), 70.1 ($CHOHCHOHCHOH$), 37.2 ($CHCO_2$), 36.3 ($CH_2CHCO_2CH_2$), 34.9 ($CH_2CHCO_2CH_2$); HRMS (ESI) cald. for $C_7H_{12}NaO_5$: 199.0582, found: 199.0581.

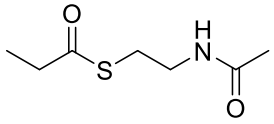
(1*S*,3*R*,4*S*,5*R*)-3,4,5-trihydroxycyclohexane-1-carboxylic acid: ν_{max}/cm^{-1} (neat) 3268 (OH), 1711 (CO); δ_H (500 MHz; CD_3OD) 4.08 (1H, td, J 5.5 and 3.5, $CHOHCHOHCHOH$), 4.00 (1H, ddd, J 8.5, 5.5 and 3.0, $CHOHCHOHCHOH$), 3.72 (1H, dd, J 5.5 and 3.0, $CHOHCHOHCHOH$), 2.76 (1H, tt, J 10.0 and 4.5, $CHCO_2H$), 2.02-1.87 (3H, m, $CH_2CHCO_2HCH_2$), 1.75 (1H, dt, J 14.0 and 5.0, $CH_2CHCO_2HCH_2$); δ_C (125 MHz, CD_3OD) 176.9 (CO_2H), 77.1 ($CHOHCHOHCHOH$), 70.5

(C*H*OHCHOHC*H*OH), 69.9 (C*H*OHCHOHC*H*OH), 37.4 (*C*HCO$_2$), 36.2 (C*H*$_2$CHCO$_2$CH$_2$), 34.7 (CH$_2$CHCO$_2$C*H*$_2$); HRMS (ESI) cald. for $C_7H_{12}NaO_5$: 199.0582, found: 199.0578.

Synthesis of N-acetyl cysteine thioesters - general procedure.

To a solution of N-(2-mercaptoethyl)acetamide (100 mg, 0.84 mmol) and triethylamine (107 µl, 0.77 mmol) in CH_2Cl_2 (5 mL) was added the acid chloride (0.7 mmol) dropwise at 0 °C. The resulting mixture was stirred at room temperature overnight and quenched by the addition of 2 M HCl. It was then extracted with CH_2Cl_2 (3 x 10 mL), washed with brine (10 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo*. The resdiue was purified by silica chromatography (EtOAc) to afford the product as a colorless oil.

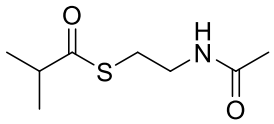
S-(2-acetamidoethyl) propanethioate[13]



S-(2-acetamidoethyl) propanethioate was synthesized using the general procedure described above using propionyl chloride (61 µl) to afford the product as a colourless oil (112 mg, 91 %).

δ_H (500 MHz; $CDCl_3$) 5.92 (1H, br. s, N*H*), 3.42 (2H, q, J 6, C*H*$_2$NH), 3.01 (2H, t, J 6.5, C*H*$_2$S), 2.59 (2H, q, J 7.5, CH$_3$C*H*$_2$), 1.96 (3H, s, COC*H*$_3$), 1.17 (3H, t, J 7.5, C*H*$_3$CH$_2$); δ_C (125 MHz, $CDCl_3$) 201.0 (*C*OS), 170.5 (*C*ON), 39.8 (*C*H$_2$NH), 37.6 (*C*H$_2$CH$_3$), 28.5 (*C*H$_2$S), 23.3 (CO*C*H$_3$), 9.8 (CH$_2$*C*H$_3$); HRMS (ESI) calc. for $C_7H_{13}NNaO_2S^+$: 198.0565, found: 198.0567.

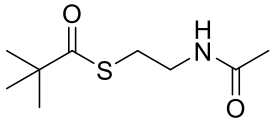
S-(2-acetamidoethyl) 2-methylpropanethioate[13]



S-(2-acetamidoethyl) 2-methylpropanethioate was synthesized using the general procedure described above using isobutyryl chloride (85 µl) to afford the product as a colourless oil (92 mg, 70 %).

δ_H (500 MHz; $CDCl_3$) 5.96 (1H, br. s, N*H*), 3.41 (2H, q, J 6.5, C*H*$_2$NH), 2.99 (2H, t, J 6.5, C*H*$_2$S), 2.59 (1H, septet, J 7, (CH$_3$)$_2$C*H*), 1.94 (3H, s, COC*H*$_3$), 1.18 (6H, d, J 7, (C*H*$_3$)$_2$CH); δ_C (125 MHz, $CDCl_3$) 205.0 (*C*OS), 170.4 (*C*ON), 43.3 (*C*H(CH$_3$)$_2$), 39.9 (*C*H$_2$NH), 28.2 (*C*H$_2$S), 23.3 (CO*C*H$_3$), 19.5 (CH(*C*H$_3$)$_2$); HRMS (ESI) calc. for $C_8H_{15}NNaO_2S^+$: 212.0721, found: 212.0722.

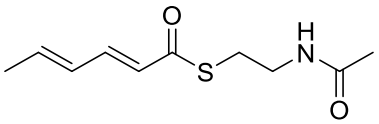
S-(2-acetamidoethyl) 2,2-dimethylpropanethioate



S-(2-acetamidoethyl) 2,2-dimethylpropanethioate was synthesized using the general procedure described above using pivaloyl chloride (86 µl) to afford the product as a white solid (78 mg, 55 %, m.p. 45-47 ^{o}C).

ν_{max}/cm^{-1} (neat) 3251 (NH), 1678, 1640 (C=O); δ_H (500 MHz; $CDCl_3$) 5.79 (1H, br. s, N*H*), 3.42 (2H, q, J 6, C*H*$_2$NH), 3.00 (2H, t, J 6.5, C*H*$_2$S), 1.96 (3H, s, COC*H*$_3$), 1.24 (9H, s, (C(C*H*$_3$)$_3$); δ_C (125 MHz, $CDCl_3$) 207.8 (*C*OS), 170.3 (*C*ON), 46.7 (*C*(CH$_3$)$_3$), 39.9 (*C*H$_2$NH), 28.2 (*C*H$_2$S), 27.5 (*C*(CH$_3$)$_3$), 23.4 (CO*C*H$_3$); HRMS (ESI) calc. for $C_9H_{11}NNaO_2S^+$: 226.0878, found: 226.0880.

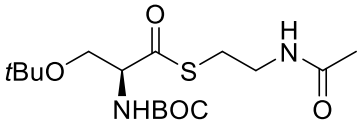
<u>*S*-(2-acetamidoethyl) (2*E*,4*E*)-hexa-2,4-dienethioate</u>[14]



To a solution of sorbic acid (76 mg, 0.67 mmol), N-(2-mercaptoethyl)acetamide (88 mg, 0.74 mmol, 1.4 equiv.) and DMAP (18 mg, 0.16 mmol, 0.3 equiv.) in CH_2Cl_2 (5 mL), was added EDC (141 mg, 0.75 mmol, 1.4 equiv.) at 0 ^{o}C. The resulting mixture was stirred at room temperature overnight. After quenching by the addition of 2 M HCl, the mixture was extracted with CH_2Cl_2 (3 x 10 mL), washed with brine (10 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo*. The residue was purified by silica chromatography (EtOAc) to afford the product as a white solid (121 mg, 85 %, m.p. 167-169 ^{o}C).

δ_H (500 MHz; $CDCl_3$) 7.20 (1H, dd, J 15.5 and 11, COSCHC*H*), 6.25 (1H, dq, J 15 and 6.5, CH$_3$C*H*), 6.16 (1H, dd, J 14.5 and 11, CH$_3$CHC*H*), 6.08 (1H, d., J 15.5, COSC*H*), 5.93 (1H, br. s, N*H*), 3.46 (2H, q, J 6, C*H*$_2$NH), 3.10 (2H, t, J 6.5, C*H*$_2$S), 1.96 (3H, s, COC*H*$_3$), 1.88 (3H, d, J 6.5, CHC*H*$_3$); δ_C (125 MHz, $CDCl_3$) 190.6 (*C*OS), 170.4 (*C*ON), 142.1 (COSCH*C*H), 142.0 (CH$_3$*C*H), 129.7 (CH$_3$CH*C*H), 125.8 (COS*C*H), 77.7 (*C*HCO), 40.1 (*C*H$_2$NH), 28.5 (*C*H$_2$S), 23.4 (CO*C*H$_3$), 19.1 (CH*C*H$_3$); HRMS (ESI) calc. for $C_{10}H_{15}NNaO_2S^+$: 236.0716, found: 236.0718.

<u>*N*-(*tert*-butoxycarbonyl)-*O*-(*tert*-butyl)-L-serine *N*-acetyl cysteamine thioester</u>[15]

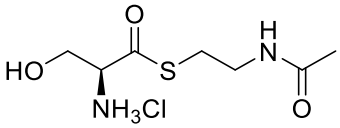


To a stirred solution of *N*-(*tert*-butoxycarbonyl)-*O*-(*tert*-butyl)-L-serine (500 mg, 1.9 mmol) in DMF (10 mL) under an argon atmosphere was added *N*-acetyl cysteamine (266 mg, 1.9 mmol), K_2CO_3 (756 mg, 5.7 mmol) and BOP (960 mg, 2.3 mmol). The resulting mixture was stirred for 3 hours and diluted with ethyl acetate (20 mL). The organics were washed with NH_4Cl (sat., 2 × 10 mL) and $NaHCO_3$ (sat., 10 ml), dried ($MgSO_4$), and concentrated *in vacuo*. The resulting

oil was filtered through silica (25% hexane in ethyl acetate) to yield the product as a white solid (588 mg, 84 %).

δ_H (300 MHz, $CDCl_3$): 5.97 (1H, br s), 5.46 (1H, br d *J* 8.0), 4.39-4.31 (1H, m), 3.85 (1H, br d *J* 9.0), 3.52 (1H, dd *J* 9.0, 3.0), 3.46-3.36 (2H, m), 3.02 (2H, t *J* 6.5), 1.95 (3H, s), 1.48 (9H, s), 1.13 (9H, s); δ_C (75 MHz, $CDCl_3$): 201.8, 170.4, 155.7, 80.8, 73.7, 62.0, 61.1, 39.7, 28.7, 28.5, 27.4, 23.2; MS ES^+: 385 (100 %, $[M+Na]^+$), 363 (11 %, $[M+H]^+$).

L-Serine *N*-acetyl cysteamine thioester hydrochloride[15]

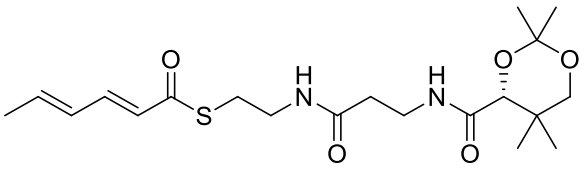


A solution of HCl (dioxane, 2 M) was added to *N*-(*tert*-butoxycarbonyl)-*O*-(*tert*-butyl)-L-serine *N*-acetyl cysteamine thioester (500 mg, 1.4 mmol) on ice with stirring. The resulting solution was stirred for 10 minutes at 0 °C and then at room temperature for 3 h. The solvent was removed *in vacuo* and the resulting solid was recrystallized (CH_3OH/CH_2Cl_2) to yield the desired product as a white solid (274 mg, 81 %).

δ_H (400 MHz, d_6-DMSO): 8.60 (3H, br s), 8.24 (1H, br t J 5.5), 4.32-4.25 (1H, m), 3.86 (2H, d J 3.5), 3.21 (2H, q, J 6.5), 3.05-2.99 (2H, m), 1.80 (3H, s); δ_C (100 MHz, d_6-DMSO): 194.8, 169.7, 60.7, 60.5, 37.8, 28.3, 22.5; MS ES^+: 435 (19 %, $[M_2-H_2Cl_2+Na]^+$), 413 (100 %, $[M_2-H_2Cl_2+H]^+$), 206 (17 %, $[M-Cl]^+$).

Synthesis of the pantetheine thioester of 2,4-hexadienoic acid

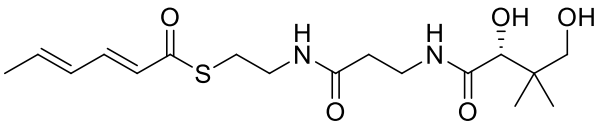
S-(2-(3-((*R*)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamido)propanamido)ethyl) (2*E*,4*E*)-hexa-2,4-dienethioate



A modification of the procedure reported by Roberts *et al.* was used.[16] To a solution of sorbic acid (61 mg, 0.54 mmol, 1.3 equiv.), (*R*)-*N*-(3-((2-mercaptoethyl)amino)-3-oxopropyl)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamide (184 mg, 0.58 mmol, 1.4 equiv.) and DMAP (14 mg, 0.12 mmol, 0.3 equiv.) in CH_2Cl_2 (5 mL), was added EDC (109 mg, 0.58 mmol, 1.4 equiv.) at 0 ^oC. The resulting mixture was stirred at room temperature for 17 h, then quenched by the addition of 2 M HCl, extracted with CH_2Cl_2 (3 x 10 mL), washed with brine (10 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo*. The residue was purified by silica gel chromatography (EtOAc) to afford the product as a colourless oil (199 mg, 90 %).

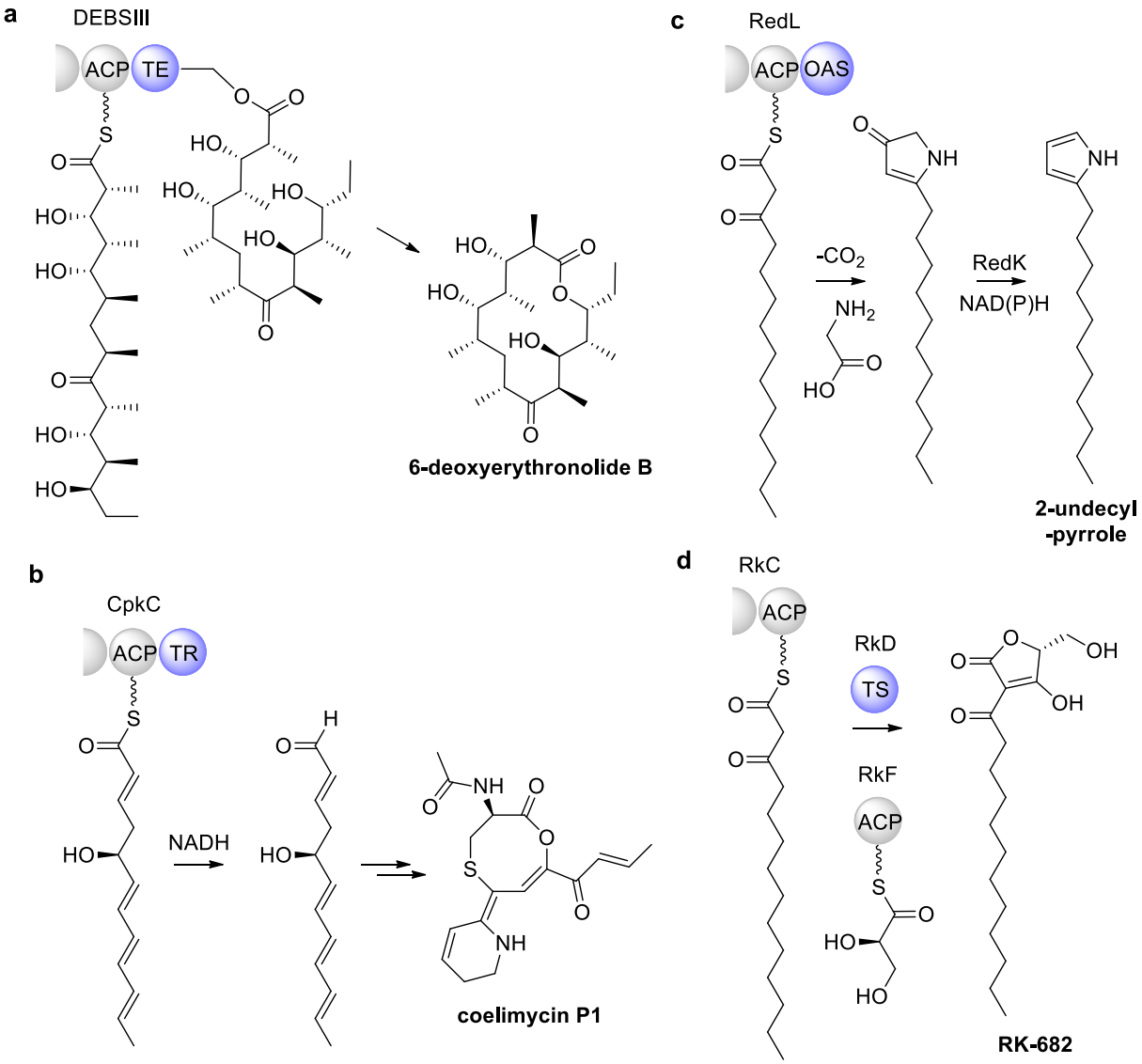
ν_{max}/cm^{-1} (neat) 3328 (NH), 2940 (C=C-H), 1671, 1512 (C=O); δ_H (500 MHz; $CDCl_3$) 7.19 (1H, dd, J 15 and 10.5, C*H*CHCOS), 7.03 (1H, br. t, J 6.0, N*H*), 6.24 (1H, dq, J 15.0 and 7.0, CH_3CH), 6.23 (1H, br. m, N*H*CH_2CH_2S), 6.15 (1H, ddd, J 15.0, 11.0 and 1.0, CH_3CHCH), 6.06 (1H, d, J 15.0, C*H*COS), 4.06 (1H, s, C*H*CONH), 3.67 (1H, d J 11.5, $CH_2OC(CH_3)_2$), 3.60-3.40 (4H, m, NHCH_2, CH_2CH_2S), 3.26 (1H, d J 11.5, $CH_2OC(CH_3)_2$), 3.08 (2H, t, J 6.5, CH_2S), 2.41 (2H, t, J 6.5, CH_2CONH), 1.87 (3H, d, J 7.0, $CHCH_3$), 1.45 (3H, s, $OC(CH_3)_2$), 1.41 (3H, s, $OC(CH_3)_2$), 1.02 (3H, s, $CH_2C(CH_3)_2$), 0.96 (3H, s, $CH_2C(CH_3)_2$); δ_C (125 MHz, $CDCl_3$) 190.3 (CO_2S), 171.4 (CH_2CONH), 170.2 ($CHCONH$), 142.1 (CH_3CH), 142.0 (CHCHCOS), 129.7 (CH_3CHCH), 125.8 (CH*C*HCOS), 99.2 ($OC(CH_3)_2$), 77.3 (*C*H), 71.6 ($CH_2OC(CH_3)_2$), 40.0 (CH_2CH_2S), 36.1 (CH_2CONH), 34.9 (CH_2NH), 33.1 ($CH_2C(CH_3)_2$), 29.7 ($OC(CH_3)_2$), 28.5 (CH_2S), 22.3 ($CH_2C(CH_3)_2$), 19.1 ($OC(CH_3)_2$), 18.9 ($CH_2C(CH_3)_2$), 14.1 ($CHCH_3$)**;** HRMS (ESI) calc. for $C_{20}H_{32}N_2NaO_5S$ [M + Na]$^+$: 435.1924, found: 435.1923..

<u>*S*-(2-(3-((*R*)-2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl) (2*E*,4*E*)-hexa-2,4-dienethioate</u>



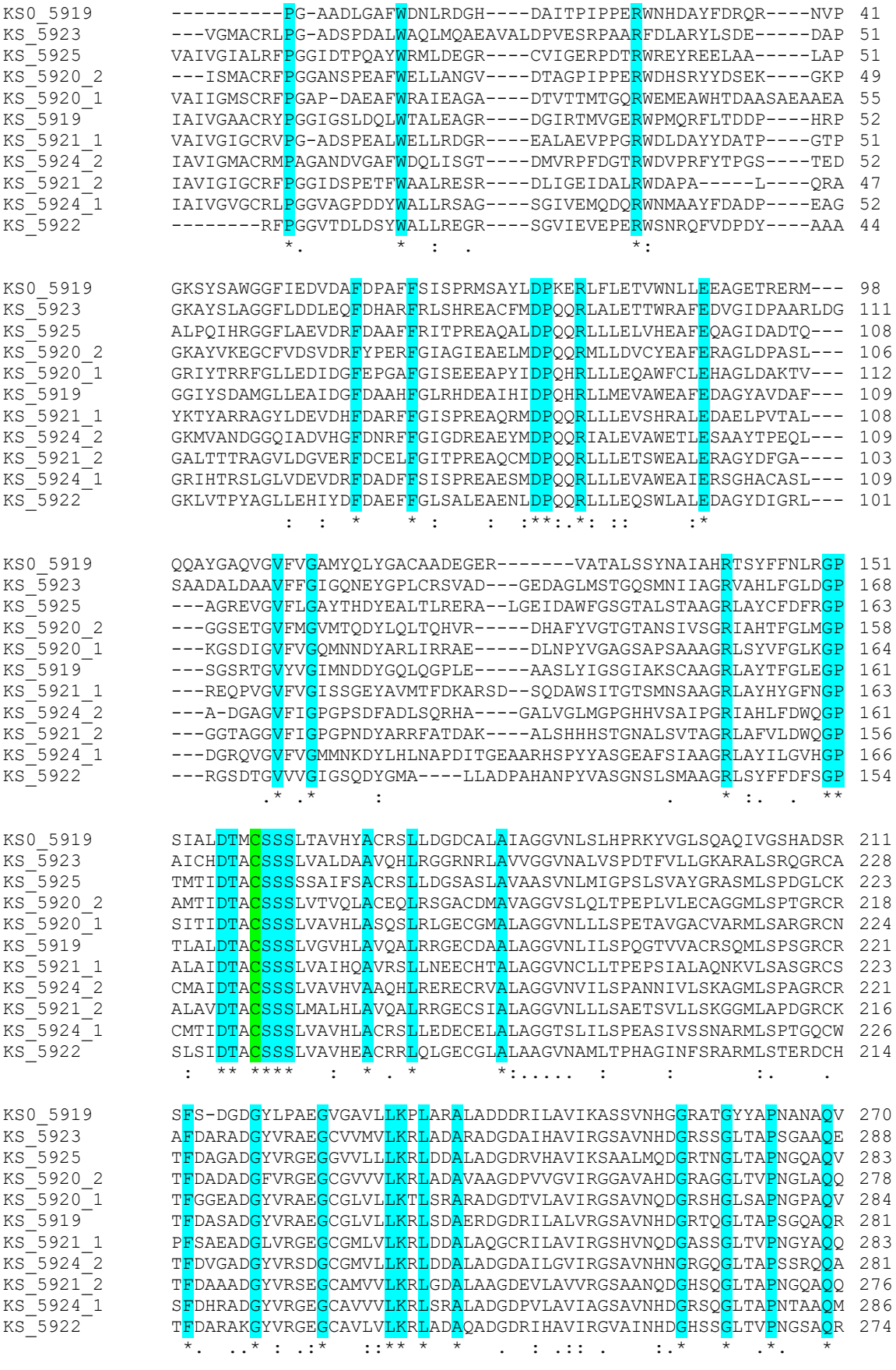
S-(2-(3-((*R*)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamido)propanamido)ethyl) (2*E*,4*E*)-hexa-2,4-dienethioate (180 mg, 0.45 mmol, 1.0 equiv.) was stirred in AcOH : H_2O (2 : 1, 3 mL), for 16 h at room temperature. The solvent was removed *in vacuo* and the residue was purified using silica gel chromatography (CH_2Cl_2 : MeOH, 85 : 15) to give the product as a colourless oil (42 mg, 25 %).

ν_{max}/cm^{-1} (neat) 3340 (OH), 2969 (NH), 2940 (C=C-H), 1651, 1540 (C=O); δ_H (500 MHz; CD_3OD) 7.21 (1H, dd, J 15.0 and 10.0, C*H*CHCOS), 6.30 (1H, dq, J 15.0 and 6.5, CH_3CH), 6.24 (1H, dd, J 15.0 and 10.0, CH_3CHCH), 6.15 (1H, d, J 15.0, C*H*COS), 3.89 (1H, s, C*H*CONH), 3.53-3.33 (6H, m, NHCH_2, CH_2CH_2S, CH_2OH), 3.08 (2H, t, J 7.0, CH_2S), 2.41 (2H, t, J 6.5, CH_2CONH), 1.87 (3H, d, J 6.0, $CHCH_3$), 0.92 (6H, s, $C(CH_3)_2$); δ_C (125 MHz, CD_3OD) 191.2 (CO_2S), 176.1 (CH_2CONH), 173.9 ($CHCONH$), 142.8 (CH_3CH), 142.8 (CHCHCOS), 130.8 (CH_3CHCH), 126.9 (CHCHCOS), 77.3 (*C*H), 70.3 (CH_2OH), 40.4 ($CH_2C(CH_3)_2$), 40.2 (CH_2CH_2S), 36.4 (CH_2CONH), 36.3 (CH_2NH), 29.0 (CH_2S), 21.3 ($CH_2C(CH_3)_2$), 20.9 ($CH_2C(CH_3)_2$), 14.2 ($CHCH_3$)**;** HRMS (ESI) calc. for $C_{17}H_{28}N_2NaO_5S$ [M + Na]$^+$: 395.1611, found: 395.1605.



Supplementary Fig. 1: Type I PKSs employ various chain release mechanisms resulting in structurally diverse products. (**a**) Thioesterase (TE) domains catalyze product release via macrolactonization (e.g. to yield 6-deoxyerythronolide B, a key intermediate in erythromycin biosynthesis), macrolactamization or hydrolysis. (**b**) Thioester reductase (TR) domain-mediated chain release results in an aldehyde, which can undergo transamination to form an amine (e.g. in the biosynthesis of coelimycin P1 and several other polyketide alkaloids) or further reduction to an alcohol. (**c**) Chain release by α-oxoamine synthase (OAS) domains involves decarboxylative condensation of an amino acid with the ACP-bound polyketide thioester. The resulting α-oxoamine can undergo cyclodehydration and reduction to form a 2-alkylpyrrole (e.g. 2-undecylpyrrole, a key intermediate in the biosynthesis of streptorubin B and other prodiginine alkaloids). (**d**) Tetronate synthase (TS)-mediated chain release involves condensation of the polyketide chain with an ACP-bound glyceryl thioester, resulting in the formation of a 2-acyl-4-hydroxymethyl-tetronic acid (e.g. in the biosynthesis of RK-682 and numerous other (spiro)tetronate-containing natural products).

a


KS0_5919 ----------PG-AADLGAFWDNLRDGH----DAITPIPPERWNHDAYFDRQR----NVP 41
KS_5923 ---VGMACRLPG-ADSPDALWAQLMQAEAVALDPVESRPAARFDLARYLSDE-----DAP 51
KS_5925 VAIVGIALRFPGGIDTPQAYWRMLDEGR----CVIGERPDTRWREYREELAA-----LAP 51
KS_5920_2 ---ISMACRFPGGANSPEAFWELLANGV----DTAGPIPPERWDHSRYYDSEK----GKP 49
KS_5920_1 VAIIGMSCRFPGAP-DAEAFWRAIEAGA----DTVTTMTGQRWEMEAWHTDAASAEAAEA 55
KS_5919 IAIVGAACRYPGGIGSLDQLWTALEAGR----DGIRTMVGERWPMQRFLTDDP----HRP 52
KS_5921_1 VAIVGIGCRVPG-ADSPEALWELLRDGR----EALAEVPPGRWDLDAYYDATP----GTP 51
KS_5924_2 IAVIGMACRMPAGANDVGAFWDQLISGT----DMVRPFDGTRWDVPRFYTPGS----TED 52
KS_5921_2 IAVIGIGCRFPGGIDSPETFWAALRESR----DLIGEIDALRWDAPA-----L----QRA 47
KS_5924_1 IAIVGVGCRLPGGVAGPDDYWALLRSAG----SGIVEMQDQRWNMAAYFDADP----EAG 52
KS_5922 --------RFPGGVTDLDSYWALLREGR----SGVIEVEPERWSNRQFVDPDY----AAA 44
GKSYSAWGGFIEDVDAFDPAFFSISPRMSAYLDPKERLFLETVWNLLEEAGETRERM--- 98
KS_5923 GKAYSLAGGFLDDLEQFDHARFRLSHREACFMDPQQRLALETTWRAFEDVGIDPAARLDG 111
KS_5925 ALPQIHRGGFLAEVDRFDAAFFRITPREAQALDPQQRLLLELVHEAFEQAGIDADTQ--- 108
KS_5920_2 GKAYVKEGCFVDSVDRFYPERFGIAGIEAELMDPQQRMLLDVCYEAFERAGLDPASL--- 106
KS_5920_1 GRIYTRRFGLLEDIDGFEPGAFGISEEEAPYIDPQHRLLLEQAWFCLEHAGLDAKTV--- 112
KS_5919 GGIYSDAMGLLEAIDGFDAAHFGLRHDEAIHIDPQHRLLMEVAWEAFEDAGYAVDAF--- 109
KS_5921_1 YKTYARRAGYLDEVDHFDARFFGISPREAQRMDPQQRLLLEVSHRALEDAELPVTAL--- 108
KS_5924_2 GKMVANDGGQIADVHGFDNRFFGIGDREAEYMDPQQRIALEVAWETLESAAYTPEQL--- 109
KS_5921_2 GALTTTRAGVLDGVERFDCELFGITPREAQCMDPQQRLLLETSWEALERAGYDFGA---- 103
KS_5924_1 GRIHTRSLGLVDEVDRFDADFFSISPREAESMDPQQRLLLEVAWEAIERSGHACASL--- 109
KS_5922 GKLVTPYAGLLEHIYDFDAEFFGLSALEAENLDPQQRLLLEQSWLALEDAGYDIGRL--- 101
QQAYGAQVGVFVGAMYQLYGACAADEGER-------VATALSSYNAIAHRTSYFFNLRGP 151
KS_5923 SAADALDAAVFFGIGQNEYGPLCRSVAD---GEDAGLMSTGQSMNIIAGRVAHLFGLDGP 168
KS_5925 ---AGREVGVFLGAYTHDYEALTLRERA--LGEIDAWFGSGTALSTAAGRLAYCFDFRGP 163
KS_5920_2 ---GGSETGVFMGVMTQDYLQLTQHVR-----DHAFYVGTGTANSIVSGRIAHTFGLMGP 158
KS_5920_1 ---KGSDIGVFVGQMNNDYARLIRRAE-----DLNPYVGAGSAPSAAAGRLSYVFGLKGP 164
KS_5919 ---SGSRTGVYVGIMNDDYGQLQGPLE-----AASLYIGSGIAKSCAAGRLAYTFGLEGP 161
KS_5921_1 ---REQPVGVFVGISSGEYAVMTFDKARSD--SQDAWSITGTSMNSAAGRLAYHYGFNGP 163
KS_5924_2 ---A-DGAGVFIGPGPSDFADLSQRHA----GALVGLMGPGHHVSAIPGRIAHLFDWQGP 161
KS_5921_2 ---GGTAGGVFIGPGPNDYARRFATDAK----ALSHHHSTGNALSVTAGRLAFVLDWQGP 156
KS_5924_1 ---DGRQVGVFVGMMNKDYLHLNAPDITGEAARHSPYYASGEAFSIAAGRLAYILGVHGP 166
KS_5922 ---RGSDTGVVVGIGSQDYGMA----LLADPAHANPYVASGNSLSMAAGRLSYFFDFSGP 154
SIALDTMCSSSLTAVHYACRSLLDGDCALAIAGGVNLSLHPRKYVGLSQAQIVGSHADSR 211
KS_5923 AICHDTACSSSLVALDAAVQHLRGGRNRLAVVGGVNALVSPDTFVLLGKARALSRQGRCA 228
KS_5925 TMTIDTACSSSSSAIFSACRSLLDGSASLAVAASVNLMIGPSLSVAYGRASMLSPDGLCK 223
KS_5920_2 AMTIDTACSSSLVTVQLACEQLRSGACDMAVAGGVSLQLTPEPLVLECAGGMLSPTGRCR 218
KS_5920_1 SITIDTACSSSLVAVHLASQSLRLGECGMALAGGVNLLLSPETAVGACVARMLSARGRCN 224
KS_5919 TLALDTACSSSLVGVHLAVQALRRGECDAALAGGVNLILSPQGTVVACRSQMLSPSGRCR 221
KS_5921_1 ALAIDTACSSSLVAIHQAVRSLLNEECHTALAGGVNCLLTPEPSIALAQNKVLSASGRCS 223
KS_5924_2 CMAIDTACSSSLVAVHVAAQHLRERECRVALAGGVNVILSPANNIVLSKAGMLSPAGRCR 221
KS_5921_2 ALAVDTACSSSLMALHLAVQALRRGECSIALAGGVNLLLSAETSVLLSKGGMLAPDGRCK 216
KS_5924_1 CMTIDTACSSSLVAVHLACRSLLEDECELALAGGTSLILSPEASIVSSNARMLSPTGQCW 226
KS_5922 SLSIDTACSSSLVAVHEACRRLQLGECGLALAAGVNAMLTPHAGINFSRARMLSTERDCH 214
SFS-DGDGYLPAEGVGAVLLKPLARALADDDRILAVIKASSVNHGGRATGYYAPNANAQV 270
KS_5923 AFDARADGYVRAEGCVVMVLKRLADARADGDAIHAVIRGSAVNHDGRSSGLTAPSGAAQE 288
KS_5925 TFDAGADGYVRGEGGVVLLLKRLDDALADGDRVHAVIKSAALMQDGRTNGLTAPNGQAQV 283
KS_5920_2 TFDADADGFVRGEGCGVVVLKRLADAVAAGDPVVGVIRGGAVAHDGRAGGLTVPNGLAQQ 278
KS_5920_1 TFGGEADGYVRAEGCGLVLLKTLSRARADGDTVLAVIRGSAVNQDGRSHGLSAPNGPAQV 284
KS_5919 TFDASADGYVRAEGCGLVLLKRLSDAERDGDRILALVRGSAVNHDGRTQGLTAPSGQAQR 281
KS_5921_1 PFSAEADGLVRGEGCGMLVLKRLDDALAQGCRILAVIRGSHVNQDGASSGLTVPNGYAQQ 283
KS_5924_2 TFDVGADGYVRSDGCGMVLLKRLDDALADGDAILGVIRGSAVNHNGRGQGLTAPSSRQQA 281
KS_5921_2 TFDAAADGYVRSEGCAMVVLKRLGDALAAGDEVLAVVRGSAANQDGHSQGLTAPNGQAQQ 276
KS_5924_1 SFDHRADGYVRGEGCAVVVLKRLSRALADGDPVLAVIAGSAVNHDGRSQGLTAPNTAAQM 286
KS_5922 TFDARAKGYVRGEGCAVLVLKRLADAQADGDRIHAVIRGVAINHDGHSSGLTVPNGSAQR 274

```
KS0_5919        DLMEASFRKAGVSPESIDYIEAAANGTSLGDAVELRALARVFDGTARDGRRVPIGTVKSN  330
KS_5923         RVMRAALRDAGVAAHEVALVEAHGTGTALGDPIEYHALRAVYADDAPRATPLVLGALKSF  348
KS_5925         DVIRRALAQAGCDPADIDYVEAHGTGTRLGDPVEIQALHEAYCAGVERAAPLSVGSVKTN  343
KS_5920_2       RVLEKALADAGIARERVSYVEAHGTGTHLGDPIELNALQAVYGRTP-RDTPLLLGSVKTN  337
KS_5920_1       QVMRDALARARLDPAEVGYLETHGTGTPLGDPVEVQAIDTVYGRAEGRRSPLALGAVKAN  344
KS_5919         RVIAAALADAGVAAAEVGFVECHGTGTALGDPIELRALEASYVLEAGERAPLVVGALKSN  341
KS_5921_1       ALIATALKRARLAPGAIGYVEAHGTGTALGDPIEIKALQQALGAGREAGRPVLIGALKAH  343
KS_5924_2       RLIEAALARAGTLPSEIRYVEAHGTGTPLGDPIEMAALKATYGAHRDAADPLYVGAVKSA  341
KS_5921_2       RVLRNALADAALDPARVGLLEAHGTGTPLGDPIEFAAARAVYGEAPGREAPLWIGSVKTN  336
KS_5924_1       ALMREALRGAKLDAARIRYVEAHGTGTPLGDPIEMNSIQAVYGEARDEASPLVIGSVKTQ  346
KS_5922         AVIRAALRRAGVAPAEVDYAEAHGTGTRLGDPIEAHAIADVYGEAREAGRPLVIGAVKAN  334
                 ::  ::  *      :   *  ..** *** :*  :              : :*::*:

KS0_5919        IGHPEAASGIAQLTKVILQMQHETLVPSIKTEPVNPNLDLAHTPFRLLSRQAAWPSDPAR  390
KS_5923         IGHTEAASGLAGLLKLVLSLRARIAPAQRHYVTPNPFIETSE-RIEIPRGA--RALGGDG  405
KS_5925         LGHTEAVSGMAGLVKVVLSMQHRRVPAHLHLNQPSPLLRLDERNIEIARQARDWQATPGR  403
KS_5920_2       IGHAEAAAGIAGLIKVLLAMRHETLPPHLHYRRANPNFDWTRGALEVVGQ--RRGWHAAA  395
KS_5920_1       MGHGESAAGIAGLIKLVQLLRHDSLPPVAHLDALNPHFDGLSDQLLFPKGA-AAAWPQGR  403
KS_5919         LGHMESAAGIGGLHKAIQVVRHRRVPRNLHFETLNPQIRVDLERLRIA-AE-AVAMPERE  399
KS_5921_1       IGHLEAASGVAGVIKTVLALRHRLLPAQINLGTPTPHVDWSSGGVAVVSESTPIAYGPDA  403
KS_5924_2       IGHTESAAGVAGLIKVLLMMRHRMIPPTLHLNTLNPHLEIDPRTIRIPTAPQPLLAREDG  401
KS_5921_2       LGHAEAAAGIAGFIKAVLCLRHEMIVPHLHFTRLNPEIELDEAAMRIPGA---TAAWRGA  393
KS_5924_1       IGHTEACAGVAGLIKLALCVAHDRVVPQRNFERLNPHITLRDGVRLALR-DEPFGGEA-G  404
KS_5922         LGHLEAAAGLAGLIKAMLVVRHGEAPPQPGFETLNPAIGWDTAKFKVVRQPTPLRPADGR  394
                :** *: :*:. . *    :              .* .

KS0_5919        PRRATVSSFGASGANAHLI-----   409
KS_5923         RVLGAVSAFGFNGTNAHVIVER--   427
KS_5925         PRRAGISSFGFSGSNTHLIVEE--   425
KS_5920_2       PLVAGVSSFGLSGTNAHLLVEQ--   417
KS_5920_1       PSVAALSSFGYTGTNAHLLL----   423
KS_5919         RALAGVSSFGFSGTNAHVIVE---   420
KS_5921_1       PFYAGVSSFGFSGTNAHLILQD--   425
KS_5924_2       TLSCAVSSFGFSGTNAHLIVAAPP   425
KS_5921_2       GRYAAVSSFGFSGTN---------   408
KS_5924_1       ARYGAVNSFGFSGTNAHLIVRDLP   428
KS_5922         PWLAGVSSFGFSGTNAHAIV----   414
                     :.:** .*:*
```

b

```
KS_GbnD5_1       ----VIGLAGRYPGAATLEAFWEAVVAARPATGALPADQWSRHVGEDDAE----ATAAAT 52
KS_GbnD4_4       --IAVVGMACHFPAAQDIDAFWRNLRDGRDCIGEVPASRWSVARHYGGPDY---ADGKSV 55
KS0_5919         ------------PGAADLGAFWDNLRDGHDAITPIPPERWNHDAYFDRQRN---VPGKSY 45
KS_GbnD5_2       --IAIVGMSGRFPQARDLDAFWRNLRDGRDCISEIPASRWDLAHYYDDGAA---EPGRIH 55
KS_GbnD1_4       --IAIVGLGGRYPQANDLDAFWDNLESGRDAITEIPPERWALTGFYDAGKD---RRGMSY 55
KS_GbnD2_2       --IAVIGLAGRYPQAADLDAFWENLSTGRDCITEIPSTRWDHEAYFDARKG---QPGKSY 55
KS_GbnD2_1       QAVAVIGLAGRYPQAADLDAFWENLSTGRDCITEIPSTRWDHEVYFDARKG---QPGKSY 57
KS_GbnD3_2       --IAIIGLSGRYPRARTLDDYWDNLRTGRDCITEIPPERWSLEGFYDPQADCSPGNGKSY 58
KS_GbnD1_3       --VAIIGLAGRYPQAPDLNAYWENLREGRDCITEIPAERWSLDGFYCEDVEQAVAQGRSY 58
KS_GbnD6_2       --IAIIGLAGRYPQAPDLNAYWENLREGRDCITEIPAERWSLDGFYCEDVERAVSEGLSY 58
KS_GbnD6_3       --IAIIGLAGRYPQAPDLNAYWENLREGRDCVTEIPAERWSLDGFYCEDVERAVSEGLSY 58
KS_GbnD1_1       ----------RFPQAASPEAFWDNLLAGRNAITSVPDGRWSDAR-------------SAA 37
KS_GbnD4_1       --VAIVGFSGRFPRARSLDAFWTNLLEQRDCIDEIPAERWDWRAIHGDPHE---PGNRTT 55
KS_GbnD3_3       -PIAIIGVSVRTAGANDAGELWELLRAGRQAIGEVPASRWDWRPYFSGPGE---ASNRIA 56
KS_GbnD4_2       --VAIIGVSVRTAGANDAGELWELLRSGRRAIGEVPASRWDWRPYFNGPGE---ASNRIA 55
KS_GbnD6_4       --VAIIGVSVRTAGANDAGELWELLRSGRRAIGEVPASRWDWRPYFSGPGE---ASNRIA 55
KS_GbnD4_3       DGIAVVGMACRCAGAQDPAAFWKLVARGEIHLDSVAARRPAWGEYLAAHEI---D---AQ 54
KS_GbnD2_3       --IAVIGGAGRYPGAPDLDSYWRNLAAGVDSVGEVPAWRWSGQPYADDEA--------LY 50
KS_GbnD1_2       --IAVVGMSGAFPKSPDLARFWDNLAAGRDCVSEVPASRWDVEAYC-DGSG---AAGRSA 54
KS_GbnD3_1       DPIAVVGMSGRFARAADLDALWAHLANGDDLVGPVTRWPKPAGR-----------ASGDA 49
KS_GbnD6_1       --IAIIGMSGRFGAAGDIDAFWRCVADGVSLIGPLPPERAALFRSLEAGEA---GKLAHR 55
                              :      *  :         :

KS_GbnD5_1       LGHGGALADADAFDALFFGMTPADAAAVDPQARLLLETAWHACEDAACLPATLA------ 106
KS_GbnD4_4       THQGGFLDDIESFDPGYFGIPEAVAPGVDPLARQWLEVSAEALADAGYTRQDVW------ 109
KS0_5919         SAWGGFIEDVDAFDPAFFSISPRMSAYLDPKERLFLETVWNLLEEAGETRERMQ---QAY 102
```

```
KS_GbnD5_2      SKWGGFIDGVDEFDPLFFRISPLEAEMMDPQERLFLQAAWETIEDAGYTRAALADAGRAP 115
KS_GbnD1_4      SKWGGFLDGVDEFDPLFFNISPREAQLIDPQERLFLQCAYHTLEDAGHTRESLG------ 109
KS_GbnD2_2      SKWGGFLDGVDEFDPFFFNISPREAQLMDPQERLFLQCAYHALEDAGHTRASLG------ 109
KS_GbnD2_1      SKWGGFLDGVDEFDPSFFSISPREAQLMDPQERLFLQCAYHALEDAGHTRASLG------ 111
KS_GbnD3_2      CKWGGFLEGFAEFDPLFFGISPREAESMDPQERLFVETCWEVIEDAGYTRRTLR---ERH 115
KS_GbnD1_3      AKWGGFVESFAQFDPLFFNLSPRDAEDIDPQERLFLETCWHVIEDAGYTRDRIA---RRH 115
KS_GbnD6_2      SKWGGFIEGFADFDPLFFNLSPREAEGIDPQERLFMQTCWHVIEDAGYTRDSLA---RRH 115
KS_GbnD6_3      SKWGGFIEGFADFDPLFFNLSPREAEGIDPQERLFMQTCWHVIEDAGYTRDSLA---RRH 115
KS_GbnD1_1      AFRGGFIDHADCFDAGFFRISPKEAEFMDPQQRVLLEVLWHTLEDARVRPSSLA------ 91
KS_GbnD4_1      IRWGGFIDEIAAFDSQFFGISPREAELMDPQQRLLMEHVWAAMEDAGYSAKSIA------ 109
KS_GbnD3_3      TNRGAFIEGLDGFDPLFFEISPREAQWMDPRQRLILEEAWRAFEDAGYAGERLR------ 110
KS_GbnD4_2      THRGAFIEGLDGFDPLFFEISPREAQWMDPRQRLILEEAWRAFEDAGYAGERLR------ 109
KS_GbnD6_4      THRGAFIEGLDGFDPLFFEISPREAQWMDPRQRLILEEAWRAFEDAGYAGERLR------ 109
KS_GbnD4_3      SLRAGFADDIDAFDPLFFDISPVQAEQMDASQRVLLEGIHAAIEDAGYDPASLA------ 108
KS_GbnD2_3      CRRLGALDDIEYFDPLFFNISPAEAELIDPQHRLFLQESHRAFEDAGYAGAGLD------ 104
KS_GbnD1_2      SRWLGLLDEADRFDPLFFRISPVEAAAMDPQQRLFLEHAWACVESAGIDPRRLA------ 108
KS_GbnD3_1      AWFGGLLDDIDRFDARFFNISATEASVMDPQQRLFLEEAWQALEAAGYAGEALD------ 103
KS_GbnD6_1      G-WSSHLEAIEAFDPLFFNISGVEARHMDPQQRLFLEESWRALEDAGLTPARLD------ 108
                    .       **  :* :    :  :*   *  ::        *      :

KS_GbnD5_1      RGRTGVYIGVMNDDYTWVAAEHL--ARTGSYASAGSYAHELANRLSHQFDLRGPSMTVES 164
KS_GbnD4_4      GRRVGVFCGSRTSNYSSRLAQLD--SK----AVIGVGQNFISAHLSHCYHLGGPNVVVDT 163
KS0_5919        GAQVGVFVGAMYQL---YGACAA--DEGERVATALSSYNAIAHRTSYFFNLRGPSIALDT 157
KS_GbnD5_2      HQRVGVYVGVMYSEYQLYGFQQQ--AERR-IFGLPGNTAGVANRVSYFFDFHGPSMSLDT 172
KS_GbnD1_4      TDRVGVFVGVMYEEYPLLGVQTQ--RGGGGGVALNGNPSSIANRVSYCFDFRGPSMAVDT 167
KS_GbnD2_2      AVRVGVFVGVMYEEYHLLSDPAG--GDLAQTYIPGGYLSSVANRVSYFGNFRGPSFGVDT 167
KS_GbnD2_1      AARVGVFVGVMYEEYPYHGSPSQ--GTTQ-PQALGGSSASIANRVSYAFNLNGPSIAMDT 168
KS_GbnD3_2      DGRIGVYAGITKTGFDLYGPALW--ARGD-DTYPHTSFCSLANRVSYLFDFNGPSLAFDT 172
KS_GbnD1_3      GSRVGVFVGVTKTGFDLHGPGAR--RDDP-LRFPHTSFSSVANRVSYFLNLHGPSLPIDT 172
KS_GbnD6_2      GGRVGVFAGITKTGFELHGPELW--RRGV-PAFPRTSFSSVANRVSYFLNLHGPSLPIDT 172
KS_GbnD6_3      GGRVGVFAGITKTGFELHGPELW--RRGV-PAFPRTSFSSVANRVSYFLNLHGPSLPIDT 172
KS_GbnD1_1      GRKVGVFMGVCNNDYVDLLAEFGVGDSSGLYGSTGTSSAILSNRVSFVFDFRGPSVTVDT 151
KS_GbnD4_1      GSRTGVFLAIGPGGYRQSA-----SQPIESYSATGAVPSMAPNRISFLLNLHGPSEPVET 164
KS_GbnD3_3      GSRCGVFVGVEEGVPGEAA----------DGLATSHHNGILAARISYVLDLKGPNLAINT 160
KS_GbnD4_2      GSRCGVFIGVEEGVPGEAA----------DGLATSHHNGILAARISYVLDLKGPNLAINT 159
KS_GbnD6_4      GSRCGVFIGVEEGVPGEAA----------DGLATSHHNGILAARISYVLDLKGPNLAINT 159
KS_GbnD4_3      AREVGTFIGSMGVAGA---------DSLSHHAMLGNDGAILSSRIAYHLNLSGPAMTVNT 159
KS_GbnD2_3      GLNCGVYLGIMGCEYAQIVMR----A--GRGSATGASAAIAAGRVAYHYNLNGPAIPVDT 158
KS_GbnD1_2      GSACGVFAGCVHGDYHRGLPA----ASLTAQDLMGASSSILAARIAYALDLKGPCLSIDT 164
KS_GbnD3_1      AARCGVYVGYNGADYQTLVD-----EHAPAQAMWGNAGSILSARIAYCLNLQGPAVTIDT 158
KS_GbnD6_1      GSRTGVYAGAADSKYGSYVR-----EDEVGHSFWGTAPSILPARVAYFLNLKGPAVTIDT 163
                    *.: .                                  : :.  .: **   .::

KS_GbnD5_1      ACSSSLLALHLARTALLAGECDLALAGGVNLSLHRSKYLLLAGLGLMSADGRERTFDIGA 224
KS_GbnD4_4      ACASALTAIHLAAQSLRSGESSLAIAGGVDILLDEGPFLLLSSARILSAGGRCRTFDEGA 223
KS0_5919        MCSSSLTAVHYACRSLLDGDCALAIAGGVNLSLHPRKYVGLSQAQIVGSHADSRSFSD-G 216
KS_GbnD5_2      MCSSSLTAIHLACRGIAYGECGVALAGGVNLSIHPNKYAVLSQARIISTKGRCESFGKGG 232
KS_GbnD1_4      MCSSSLTAIHLARESLLRGECDAALAGGVNVSIHPNKYLSLSQGHFASSDGRCRSFGAGG 227
KS_GbnD2_2      MCSSSLTALHLACQSLRLGECELALAGGVNVSIHPNKYLGLSQGQFASSEGRCRSFGAGG 227
KS_GbnD2_1      MCSSSLTALHLACQSLRLGECELALAGGVNVSIHPNKYLGLSQGQFASSEGRCRSFGAGG 228
KS_GbnD3_2      MCSSSLVAIHEACAHLLAGECELAIAGGVNLYLHPSNYVALCAHRMLSPDGRCRSFGAGA 232
KS_GbnD1_3      MCSSSLTAIHEACEHLLRDECELAIAGGVNLYLHPSNYVELCRHRMLSPRGRCRSFGEGG 232
KS_GbnD6_2      MCSSSLTAIHEACEHLLRDECELAIAGGVNLYLHPSSYVMLCLSRMLSPRGRCRSFGEGG 232
KS_GbnD6_3      MCSSSLTAIHEACEHLLRDECELAIAGGVNLYLHPSSYVMLCLSRMLSPRGRCRSFGEGG 232
KS_GbnD1_1      ACSSSLVAVLMAARAIRSGECEAALAGGVNICWAASRFLAFAQAGMLSKDGVCRTFDAHA 211
KS_GbnD4_1      ACSSSLVAVHRAMRAIEAGDCEQAFVGGVNTIVSPELHICFSKAGMLSPDGRCKTFSREA 224
KS_GbnD3_3      ACSSGLVAVHAACQSVQRDECELALAGGVNVLNSPLMYLALTQGGMLSPDGECYTFDARA 220
KS_GbnD4_2      ACSSGLVAVHTACQSVQRGECELALAGGVNVLNSPLTYVALTQGGMLSSSGECHAFDARA 219
KS_GbnD6_4      ACSSGLVAVHTACQSVQRGECELALAGGVNVLNSPLTYVALTQGGMLSSSGECHAFDARA 219
KS_GbnD4_3      ACSSALVAISLACDKLRAGELDMAIAGGITLYTQPASFVMMRNAGMLSPSGACRPFDDGA 219
KS_GbnD2_3      ACSSSLVAIHLAALALERGEIDLALAGGVSLYLSAETYARMCEAGMLSRSGSCRSFDDGA 218
KS_GbnD1_2      ACSSSLVAVATACDSLAEGRSELALAGGVCVLAGPDLHIMASDAGMLSPTGRCHSFDSRA 224
KS_GbnD3_1      ACSSSLVSVHLACQALRAREIDAALAGGAFVQATPGFYLLAGRAGMLSPTGRCATFAAGA 218
KS_GbnD6_1      ACSSSLVAIHLGCESLRNHETDLILAGGAFLQATPHFGAAAGNAQMLAADGKCHTFDDRA 223
                  *:*.* ::  .   :        :.**                : .  .    *   .

KS_GbnD5_1      NGYVPGEGVGLALLKPMEAALADGDRIHGVIAGSATNHSGRAAGRYSPNLHALVEVIERG 284
KS_GbnD4_4      DGIGIGEGCGVLILKRLSDAVREGNKIYGVIDGSAVNADGSTMGITTPNPERQRELIELA 283
```

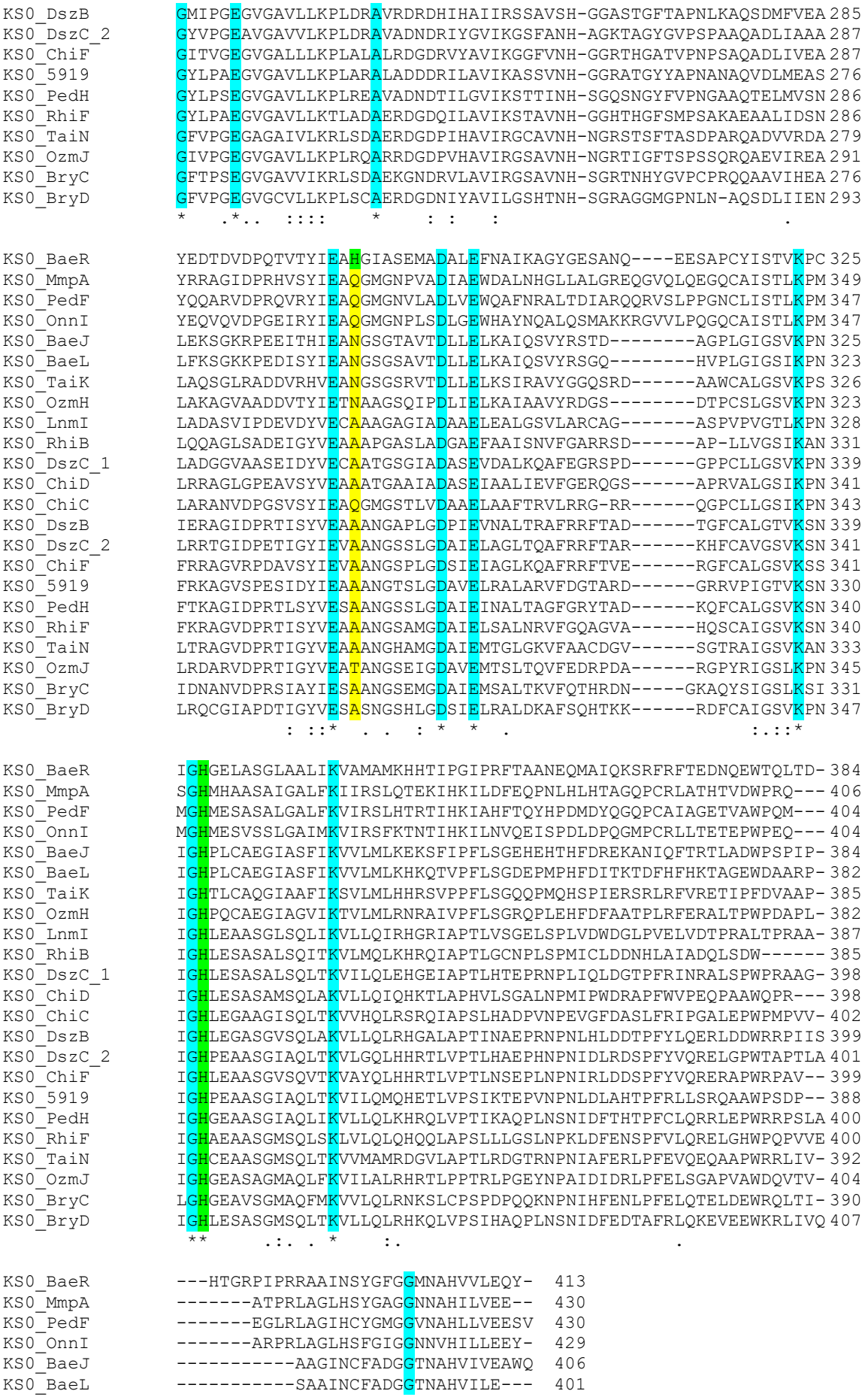
```
KS0_5919        DGYLPAEGVGAVLLKPLARALADDDRILAVIKASSVNHGGRATGYYAPNANAQVDLMEAS 276
KS_GbnD5_2      EGYIPGEGVGCVMLKPLARALADRDAIHGIVKGTALSHGGRANGYTVPNPGAQAATISMA 292
KS_GbnD1_4      DGYVPGEGVGAVLLRRIDDALADGDVIHALIRGSSINHGGKTNGYTVPNPGAQRELIEAA 287
KS_GbnD2_2      DGYVPSEGVGCVLLRPLAAAEAAGDRILGVIRASAINHGGRTNGYTVPNPNAQGELIAEA 287
KS_GbnD2_1      DGYVPSEGVGCVLLRPLAAAEAAGDRILGVIRASAINHGGRTNGYTVPNPNAQGELIAEA 288
KS_GbnD3_2      NGYVPGEAVGAVLLKPLAAAQADGDNIHGIVRGSAINHGGKTHGYTVPNPGAQGALIERA 292
KS_GbnD1_3      DGMVPGEGVGAVLLKRLGAAEADGDRIHGVIRASAINHGGKTNGFTVPNPNAQAELIVEA 292
KS_GbnD6_2      DGMVPGEGVGAVLLKRLSAAEADGDRIHGVIRASAINHGGKTNGYTVPNPGAQRELIVSA 292
KS_GbnD6_3      DGMVPGEGVGAVLLKRLSAAEADGDRIHGVIRASAINHGGKTNGYTVPNPGAQRELIVSA 292
KS_GbnD1_1      NGYVRGEGAGAVLLKPLRRALADGNEVYGVIRGGAINHGGRTRGLTVTNPDQQRQLLIDA 271
KS_GbnD4_1      NGYVRGEGVGAMLLRPLGRAIADGDTIHGVIRAVSVNHGGRAASLTAPNVQAQAALLTDA 284
KS_GbnD3_3      NGMVPGEAVAAVVLKDLARAQADGDPIHGVIRASGVNYDGRTNGITAPSGLSQQQLVEQL 280
KS_GbnD4_2      DGMVPGEAVAAVVLKDLARAQADGDPIHGVIRASGVNYDGRTNGITAPSARSQRALISEV 279
KS_GbnD6_4      DGMVPGEAVAAVVLKDLARAQADGDPIHGVIRASGVNYDGRTNGITAPSARSQRALISEV 279
KS_GbnD4_3      DGIVVGDGMGVVVLMPAERARAEGAHVYGVIRAIGTNQDGRTSGITAPSFLAQSRLETQV 279
KS_GbnD2_3      DGFVPGEGVGAVVLKRLDDALRDGDRVDAVIVASGINQDGRTNGITAPSMKSQAALLRTV 278
KS_GbnD1_2      DGFVPAEGVGVVLLKRLDDALRDGDPVHGVLSGWGVNQDGRTNGITAPSVASQAALQAGV 284
KS_GbnD3_1      DGFVPAEGVGVVMLKRLADALDDGDTIHGVIRGSGINQDGTTSGITAPSALSQERLQREV 278
KS_GbnD6_1      NGMVSGEAVAVLVLKRLDEARADGDRIHGVIVGSGINQDGATNGITAPSALAQQQLQTDV 283
                :*   .:. .  :*     *      : .:: . . . .* : .    .

KS_GbnD5_1      VASAGLAPEAIGHVETHGTGTQLGDPIEVQAIARALGAAGR-------------GDKRCT 331
KS_GbnD4_4      IADAAVDAASISYVEAHGTGTLIGDPIELRSLTAVLAPHHR-------------AARACG 330
KS0_5919        FRKAGVSPESIDYIEAAANGTSLGDAVELRALARVFDGTAR-------------DGRRVP 323
KS_GbnD5_2      IEQAGVAPRAISYIEAHGTGTSLGDPIEIAGLVHAFGELGA-------------SGQFCA 339
KS_GbnD1_4      LAAAGVRADEISYVEAHGTGTELGDPIEIAGLTQAFGDAHAA-------------AGRRCA 335
KS_GbnD2_2      LRASGVDARAISYLEAHGTGTALGDPIEIAGLVKAYGAWEGE-----PGEPGDARLEPCA 342
KS_GbnD2_1      LRASGVDARAISYLEAHGTGTALGDPIEIAGLVKAYGAWEGE-----PGEPGDARLEPCA 343
KS_GbnD3_2      LARAGVAARQLGYVEAHGTGTELGDPIEIAGLMRAFGADAM---------PA--EDRPCA 341
KS_GbnD1_3      LTRGGIDSAAVAYLEAHGTGTALGDPIEIAGLAKAFAQASR---------QTHERALPCA 343
KS_GbnD6_2      LAKSGIEARDIGYVEAHGTGTELGDPIEIAGLSQAYGDT---------------GGSSCA 337
KS_GbnD6_3      LAKSGIEARDIGYVEAHGTGTELGDPIEIAGLSQAYGDT---------------GGSSCA 337
KS_GbnD1_1      YRDAGVDPAMVGYIEVHGTGTSLGDPIELLGLKQAFESTHLAAPAGRARSRTRVEPGQCV 331
KS_GbnD4_1      YRRANLDIRTVSYVEAHGTGTSLGDPAEINGLKRAFATLYE-------QQQARVERVHCA 337
KS_GbnD3_3      YRRYRIDASEIGYVIAHGTGTRLGDPVEVNSLADAFRT-------------FTDRRGFCA 327
KS_GbnD4_2      LERGGVEAERIEAVLAHSVGSPLGDPIEARALCEALGEGL--------NEGLNEGTQTRV 331
KS_GbnD6_4      LERGGIEAERIEAVLAHSVGSPLGDPIEARALCEALGEGLGEGLNEGLNEGLNEGTQTRV 339
KS_GbnD4_3      YRAARLSPEALQYVEAHGTGTRLGDPVELHALTEAFRG-------------FTARRSFCA 326
KS_GbnD2_3      HARHGIDPASIGYVEAHGTGTRLGDPIELAALATAFGA-------------APRAGGPCA 325
KS_GbnD1_2      HRRFGIDPASISLVEAHGTGTQLGDPIELEALTASFGA-------------HAASRAGCA 331
KS_GbnD3_1      YDSWGIDVETIGLVEAHGTGTRLGDPVEHRALASAFRR-------------DTAKRGFCA 325
KS_GbnD6_1      YARFGIDPDGIQYVEAHGTGTSLGDPIEMRALTDAFRR-------------HTERRGYCA 330
                     :    :  : . . *: :**  *  .:

KS_GbnD5_1      LGS-RANFGHLESASGIGALTKTLLGMRRGLIPPCANLETLNPALRLDETALTIPREAMR 390
KS_GbnD4_4      VGSVKSNLGHLLSAAGAAGMVKVLLSLAHRALPPTLHCDTPNPRFDFEASPLYPVRELQA 390
KS0_5919        IGTVKSNIGHPEAASGIAQLTKVILQMQHETLVPSIKTEPVNPNLDLAHTPFRLLSRQAA 383
KS_GbnD5_2      IGSAKSNIGHGESAAGIAGVCKVLLQMRHRQLAPSLHSAELNPYIDFASSPFAVQRELGE 399
KS_GbnD1_4      IGSVKSNIGHAESAAGIAGLTKVLLQMRHGRLVPSLHADTLNPHIDFERTALRVQRELAD 395
KS_GbnD2_2      IGSVKSNIGHCESAAGIAGLTKVLLQMRHGKLAPSLHAQTLNPLIDFGRTPFRVQRELAP 402
KS_GbnD2_1      IGSVKSNIGHCESAAGIAGLTKVLLQMRHGKLAPSLHAQTLNPLIDFGRTPFRVQRELAP 403
KS_GbnD3_2      LGSVKSNIGHAESAAGIAAVTKVLLQFRHRQLVPTLHVDTPSPQIRFERTPLRLQTQLAD 401
KS_GbnD1_3      IGSVKSNIGHAESAAGIAGLTKVLLQMRHGELAPSLHADHTNPLIDFGATPFRVQRELAR 403
KS_GbnD6_2      IGSAKSNIGHAESAAGIAGLTKVLLQMRHGELVPSLHAQRPNPHIDFGRTPFRLQTALSA 397
KS_GbnD6_3      IGSAKSNIGHAESAAGIAGLTKVLLQMRHGELVPSLHAQRPNPHIDFGRTPFRLQTALSA 397
KS_GbnD1_1      LGAVKTNIGHLEGAAGIAGLIKVLLALRHHTIPANLNFQTLNPRIKLEGTRFRIPTETLA 391
KS_GbnD4_1      LGSVKSSIGHLEVAAGMAGLFKVLLAMRHGVLPGTLHCEDTNPYIELEGTPFEILKQNRA 397
KS_GbnD3_3      LGSFKPNLGHTFAASGVVSLVAMLAAIRHRQIPPSANHRSDNEFLDLPNSAFRLVEHCTR 387
KS_GbnD4_2      LGSVKPQIGHTFAASGVVNVIAMCASLRHELRLGIANHEVANPDLRIGDGALSLGAGAQP 391
KS_GbnD6_4      LGSVKPQIGHTFAASGVVNVIAMCASLRHGLRLGIANHEVANPDLRIGDGALSLGAGAQP 399
KS_GbnD4_3      IGSVKANIGHATAAAGALGLIKVLLAMRHATLPPVSGFAQPNRHVDFAASPFRVETQARP 386
KS_GbnD2_3      LGAVKTNLGHTSAAAGVAGLHKILLCLRHRELVPSLHFTRPNAHFEFAGSGLRLVSEREA 385
KS_GbnD1_2      LGSVKSNIGHSLAAAGISGLLKVLLALRHRQLPPSIHFVTPNPHLLLESSPFRINTALCD 391
KS_GbnD3_1      LGAIKTNLGHAAAAAGIAGLLKLLLALRHRQIPPSLHFDAPNPAIDFADSPFYVNTTLQA 385
KS_GbnD6_1      IGSVKTNVGHTVTAAGVTGVVKVLMAMRHRLLPALLNYEVPNRHIDFETSPFFVNTEALP 390
                :*: : ..**   *:*   :      : :           .  . :    :

KS_GbnD5_1      WTLPL---------ESRVSGIHAFGIGGSNVFMV-----     415
```

```
KS_GbnD4_4      WAGVD---------GVLRAGISAFGLGGHNAHLIVSD--     418
KS0_5919        WPSDP--------ARPRRATVSSFGASGANAHLI-----     409
KS_GbnD5_2      WR-------------------------------------     401
KS_GbnD1_4      WVRPRAT---------------------------------     402
KS_GbnD2_2      WRRPRVRVDGVEREMPRLAGISSFGAGGANAH-------     434
KS_GbnD2_1      WRRPRVRVDGVEREMPRLAGISSFGAGGANAH-------     435
KS_GbnD3_2      WDAPG--------GTPRLAGVSSFGAAGTNAHVVLQEY-     431
KS_GbnD1_3      WPRARATAAGIERELPRLAGISSFGAGGANAHLIVE---     439
KS_GbnD6_2      WPRPVCEEAGERRERPRIAAISSFGAGGANAHLIVEEY-     435
KS_GbnD6_3      WPRPVREEAGERRERPRIAAISSFGAGGSNAHLIVEEY-     435
KS_GbnD1_1      WPAPT---SGR-SRLRRRAGVSSFGYGGAYAHVVVE---     423
KS_GbnD4_1      WTRLS---DARGAPVPRRAGISSFGFGGVNAHLVVEEY-     432
KS_GbnD3_3      WESD----------APRVGAISAFGMSGTNAHLVIAEYV     416
KS_GbnD4_2      WPKRA---G-----VARCGLVSATGMSGTNACVVIEE--     420
KS_GbnD6_4      WPKRA---G-----VARCGLVSATGMSGTNACVVIEE--     428
KS_GbnD4_3      WLPGE---D-----GMRRAALSAFGFSGTNAHL------     411
KS_GbnD2_3      WATHG---D-----APRRAALSSFGYSGTNAHLVVEEYV     416
KS_GbnD1_2      WTVPA---G-----GLRRAAINSFGFSGTNAHLVVDEAP     422
KS_GbnD3_1      WNPPP---G-----MPRRAALSAFGFSGTNAHLVID---     413
KS_GbnD6_1      WTPAP---G-----GVLRAAVSSFGFSGTNAHLVLESH-     420
                *
```

c

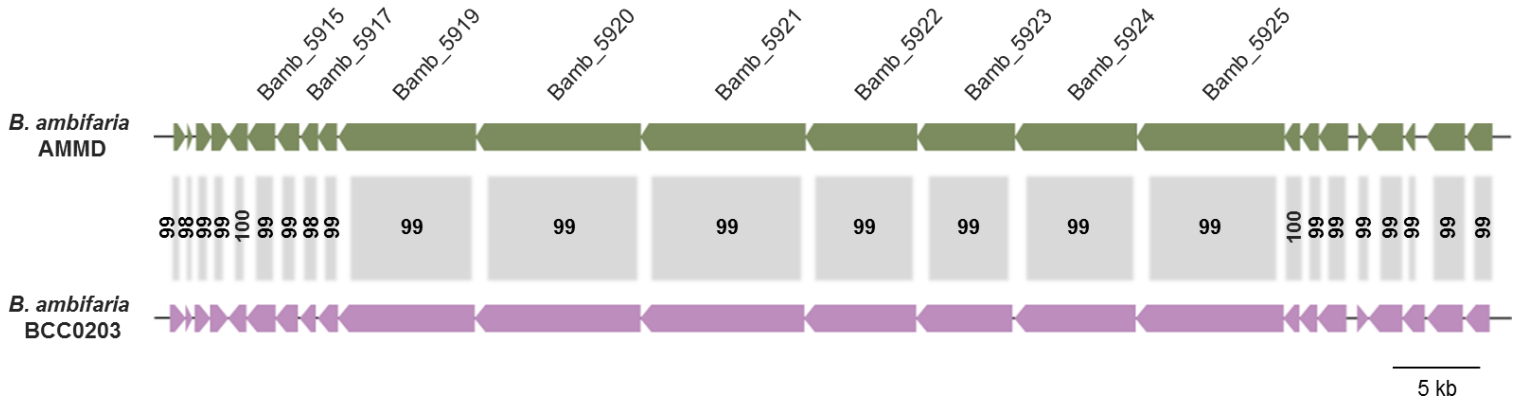
```
KS0_BaeR        ---IAIIGMSAQFPQSPDIQSFWEHIVNGDHCITEIPADRWDWRRYAGD-------ENDT 50
KS0_MmpA        PEPVAIIGLSANVAQSASVRQFWQALDDDRSLIEEIPATRFDFTSWYAGSNI---EEGKM 57
KS0_PedF        -EAIAIVGLSGYFPQSASVDEFWRHLDQDATLIEEIPDSRFDWRKVFDPTGE---RPGSS 56
KS0_OnnI        -EPIAIIGLSGSLPKSQTIAEFWRSLDQDLSLIEEIPRSRFNWEEVYDPDGK---DVDKM 56
KS0_BaeJ        -EDIAVVGMSCRFPGAASLEEYWSLLAEGRSAIRPVPAERWGFKT---------------- 44
KS0_BaeL        ---IAVVGMSCRFPGAESLEQYWDLLRSGRSAIGSVPAERFGYAN---------------- 42
KS0_TaiK        --DIAVVGLACRFPGAPSVDAYWALLRDGARGIGPAPRERFAQAD---------------- 43
KS0_OzmH        --DIAVVGMSGRFPGAPDLDAYWRLLSEGRSAIAPVPARRWADGT---------------- 43
KS0_LnmI        ------IGMAGRLPGAGDLDAFWDNLVSGRTAIGPAPASRPETAPS-------------- 40
KS0_RhiB        ---MAIIGISGRYPGAANPDELWQNLSAGRASIIPLSREALF------YGSDD----AGD 47
KS0_DszC_1      -----VIGLAGRYPGADTPRQLWRALRSGQSAVTRPPAGRFGASAPQGDEPRG----GGA 51
KS0_ChiD        ---IAIVGQSGRYPGAPDAAALWERLRRGERSIRPAPADRWDPAPLQATGPDK----GGI 53
KS0_ChiC        -DDIAILALDGRYPQARSPEELWENLRAGRECTREVPADRWDVSAYYDADPRR-AAAGRM 58
KS0_DszB        ---IAIIGVSGRYPQAEDLRALWARLQAGESCIEEIPAERWDKDRYFDP--QK-GRSGKS 54
KS0_DszC_2      -VDIAIVGLSGRYPGADTIDAFWSNLRQGRDSVTEVPADRWDAAAIFDP--E--GGPGKT 55
KS0_ChiF        ---FAIIGIGGRYPEAADVREFWENLKAGRSCIGEVPPHRWDGDAYYRP--D--GG-GAS 52
KS0_5919        -------------PGAADLGAFWDNLRDGHDAITPIPPERWNHDAYFDR--QR-NVPGKS 44
KS0_PedH        ---IAIIGMSGRFPFAPDLEAFWENLSQGCDCITEIPPTRWKHQEYFDP--EK-GKPGKT 54
KS0_RhiF        ---LAIVGISGRYPGAEDLEAFWHKLAGGEDLISEVPTQRWDHQAYFAD--QR-DRFDKT 54
KS0_TaiN        -EPVAIIGISGRYPGAYDVPAFWRNLLAGACAITEVPAERWDWRAHYRADAAEAAREGKS 59
KS0_OzmJ        PDAVAVIGMSGVFPGAPDPDGLWELLMAGRSAVTEVPGRRWDWREHYDPHPEGADVVGKS 60
KS0_BryC        ---MAVIGMSACYPSAKNLDQYWENLKCGKNCITEIPDDRWSIDGFFCPDVEEALSQGKS 57
KS0_BryD        -EPIAIIGLSGHYPQANSLDAYWENLKAGKDCIREIPDDRWSLDGFFHEDVEEAIAQGKS 59
                                :      *  :  .

KS0_BaeR        SLRW-GGFIDGVGEFDPLFFGISPKEASQMGPEQFLLLMHTWKAMEDAGLTNKALS---- 105
KS0_MmpA        RTRW-GGFIPAIDQFDPVFFGMLPAEARKMDPQQRLLLMSVRQTFEDAGYRHTDWK---- 112
KS0_PedF        CSKW-GGFIPDIRGFDPAFFNIPGAEAITLDPRQRLLLMSAYQTLNDAGYASQALR---- 111
KS0_OnnI        RTKW-GGFLRDIYGFDPHFFKILPRDAAVMDPRQRLLLMSVYQTLADAGYAPETFK---- 111
KS0_BaeJ        --PYYAGMLDGIHQFDPDFFLLAEEDVKAMDPQALAVLEECLNLWYHAGYSPDEIK---- 98
KS0_BaeL        --QYVAGLIDNMDHFDSEFFFIPENDAKAMDPQALAVLEESLKLWCHAGYSREEIK---- 96
KS0_TaiK        --RFCGGFLDAVGRFDPDHFGIAPGDARAMDPQALLLLELGVELFHHAGYRPEELR---- 97
KS0_OzmH        --KYTAGLLDL-EGFDPGHFHLSDADAAAMDPQALLLLEETLFAFCDAGYAPDELK---- 96
KS0_LnmI        GARATGGFLPHIDRFDSLLFHVSPQEAPALDPQARLMLESVWQCLDDAGHTADSLR---- 96
KS0_RhiB        SPQWAVGALAGKQLFDPLLFKITPAEAKTLDPQERLFLQAVWHCLESSGYTAASLR---- 103
KS0_DszC_1      SPGW-GGYLERLDRFDSLFFGISPAEAKLMDPQERLFIEVAWECLEDAGYTPEELR---- 106
KS0_ChiD        YCSS-GGFLDDVDRFDCLLFRMSPAEARSIDPQERLFLEAAWACLEAAGTTAERLN---- 108
KS0_ChiC        YCKW-GGFLDDIGRFDALFFQISPTEAASLDPSERLFLEIAWSTLERAGYARRRPQ---- 113
KS0_DszB        ESKW-GGFLRDVDQFDPLLFNIPPARARIMDPMQRLFLESVYETLEDAGYTRAMLS---- 109
KS0_DszC_2      RQRW-GGFLDRVDRFDALLFNISPREAAGMDPQERLFLEIAWCAFEDAVYTRERLAEEQA 114
KS0_ChiF        RSKW-GGFLEDVDRFDPLLFNISPLEAERLDPQLRLFLQTAWETFEDAGYPRRRLRVVQQ 111
KS0_5919        YSAW-GGFIEDVDAFDPAFFSISPRMSAYLDPKERLFLETVWNLLEEAGETRERMQ---Q 100
KS0_PedH        YCKW-GGFLESIDQFDPLFFKIPPAQAEVLDPQERLFLETVWNLLESSGYLGETLQ---R 110
KS0_RhiF        YCKW-GGFLDGVEDFDPLFFNLSPREAEIINPNDRLFIETCWNLLESAGLTRQRLK---Q 110
KS0_TaiN        YSKW-GGFVDDVGRFDPAFFGMTPQDAQHTDPQELLFLEMCWHALQDAGQTPALLP---G 115
```

```
KS0_OzmJ        HSKW-GAFLDGFDAFDPAVFGFTEQEARNTDPQVRLFLQECWKALEDAGIAPSKLP---S 116
KS0_BryC        YSKW-GGFLEDFAAFDPLFFNLSPRDAMRIDPQERIFLQECWRAFEDAGYVCSRLS---P 113
KS0_BryD        YSKW-GGFLEGFADFDPLFFNLSPREVMTIDPQERLFLQSAWEAVEDAGYTRAQLA---S 115
                      . :     **   * .        .*    .:         :

KS0_BaeR        --SRPTGVFVAAGN-----------SD-PNNGTAI-PSIIPNRISYALNLQGPSEYYEAA 150
KS0_MmpA        --GSATGVFIAAERNEYHLNLLQAQID-PGEGLDQAASMLANRVSHFYDLRGPSERIDAM 169
KS0_PedF        --QSKTGVFVALQDNEYLQLLADAGID-PGQWYAQ-TCLLANRISYFFDWRGTSEVVDAQ 167
KS0_OnnI        --KSKTGVFFSIQDNEYLQLLREGGVD-RGEGFGH-ASMIANRIAYFFDFRGPSEFVDAQ 167
KS0_BaeJ        --GDAIGVYLGGRSRHRPGEEKLLDAK--NPIVALGQNYLAANLSQYFDMRGPSVVLDTA 154
KS0_BaeL        --GIEAGVYIGGRSQHQPDPEILANTR--NPIVAGGQNYLAANVSQFFDLRGPSIVLDTA 152
KS0_TaiK        --GGAVGVFLGGRSQHAPDAALLAHAH--HPIVAVGQNYLAANLSRHFDLNGACALVDTA 153
KS0_OzmH        --GRGIGVYVGGRSRHVPDEATLGRSR--NPVVAVGQNYLAANLSHHFDLRGPSTVVDTA 152
KS0_LnmI        RSAGRVGVFIGSMWHDYRQQGADRWNGGDSAEVAATASDIANRVSHFFDFRGPSLAVDTS 156
KS0_RhiB        RQAERIGVFVGAMWGDYQHHRPT----EQGERATSFLSAIANRVSFFNDFNGPSVAFDTS 159
KS0_DszC_1      RAAPRVGVFVGAMWSDYQSVGLEAWQRDRRAKAVAFHSSIANRISYLFDLHGPSVAIDTS 166
KS0_ChiD        AQAGKVGVFVGVMWNDFQNEGVEGFREDHVARAVALHSSIANRVSHTFDFKGPSVAVDTS 168
KS0_ChiC        --SRSVGVFVGVNVGDYHLLALEEQARGRWVFSNPSFSAIANRVSYFFDFQGPSLAIDTQ 171
KS0_DszB        KDGGKVGVYVGAIYHHYAMLAADESTR--SLLLSAFGAHIANHVSHFFDLHGPCMAVDTT 167
KS0_DszC_2      RAGVGAGVFVGSMYQQYSMLARTPDAG--ASSS---FWSIANRVSYFFDLRGPSLAVDTA 169
KS0_ChiF        GATSGVGVFVGSMYQHYPFVAPDGATA--AQLSSFPGSAIANRVSHYFDLKGPSMLVDTA 169
KS0_5919        AYGAQVGVFVGAMYQLYGACAADEGER--VATALSSYNAIAHRTSYFFNLRGPSIALDTM 158
KS0_PedH        IAQSRVGVFVGSMSQQYHAFQADLTRE--SLVTMSSHSSIANRVSYFFDFQGPSVAVDTM 168
KS0_RhiF        QYQQQVGVFVGVMYQQYQAFEADFVRE--SLVSVTSYSAIANRVSYFFDFQGPSLAIDTM 168
KS0_TaiN        DVRRRAGVFAAITKH--------------YAFPPTSFASLANRVSHALDFGGKSLAIDTM 161
KS0_OzmJ        ETRGRIGVFAGGAKHGFTQL----GAEGRLEMPRTSFGDMVNRVSFQFDLGGPSKAVDTA 172
KS0_BryC        ELRHKTGVYGAMTK----------------INPNTSFASLVNRVSYIMDLHGPSVPVDSM 157
KS0_BryD        QFNKRVGVFAGITKTGFNLYAGDLNSQAELFYPYTSFSMLVNRVSYFLDLQGPSIPVDTM 175
                      **: .                              :  . :   :  * .   ::

KS0_BaeR        CTSTLVALHRAVQSIRHNECEQAVVGAANILQSPKGFIGFDSMGYLSKNGRAKSFQKDAD 210
KS0_MmpA        CAGGAVALHHAVTALRSGQINAAIVGACNLLLRPDVFVTLSQSGQMSPEPTVRSFGAGAD 229
KS0_PedF        CPGAAVAIHRAVSALRNGEIELALVGAANLLLRPEPFVLLSESGQLSESASVHSFGAQAQ 227
KS0_OnnI        CAGAAVALYRAVSTLRSGDITYAVVGAANLLLRAEPFAVLTRANQLSPTNCVNSFGKDAQ 227
KS0_BaeJ        CSSALVGMNMAVQALVTGEIKAAVVGGVSLFESEETHKLFEQRGILSKAQSFHVFDQRAD 214
KS0_BaeL        CSSALTGMNMAVQALRSGDIKAAVVGGVSLLNTDAAHRMFQERGLLNEKPAFHVFDKRSG 212
KS0_TaiK        CSSALVAMHSAVLALAAGEIDAAVVGGVSLLSSDAGHRLFEQRGLLAPDGAFHLFDERAN 213
KS0_OzmH        CSSALVALHHAAQALRSGDVEAAVVAGVTLLPDAGGHRLFDRRGLLNTGTEFHVFDRRAR 212
KS0_LnmI        CSSSFAALHLAVESLRRGECGAAVVGAVNLLAHPYHWGLLDGLELLAADAPPAAYAAEGS 216
KS0_RhiB        CSSAMTALHFACNSIRQGECQAAIVGGVNLISHPSHLELLTSLKLLSDDSQSYPFGRHAN 219
KS0_DszC_1      CSSGLTALHLASRSLRLGECDVALVGGVNLLGHPFHPDLLEGLNLTSRDDKTRAFGAGGS 226
KS0_ChiD        CSSAMTALHLACESIQRGECRAAIVGGVNLMTHPYHQGLLCSLGMVSESGFGNALGEDAT 228
KS0_ChiC        CSSSLTAIHLACESLLRGECEMALAGGVNLYPHPSRYVNLCQVKALSSTGQTRSFGAGGD 231
KS0_DszB        CASSLTAIHLACEGLLLGRTDLAIAGGVNLSLIPEKYLGLSQLQFMSGGALSRPFGD-SD 226
KS0_DszC_2      CASSLTALHLACESLRRGECCLALAGGVNLHLHPHKYVALDRLGLLGSGAASKSLGD-GD 228
KS0_ChiF        CSSSLTAIYMACESLARGECAMALAGGVNLSLHPQKYVIFSQMGLLGSKERSSSLGE-GD 228
KS0_5919        CSSSLTAVHYACRSLLDGDCALAIAGGVNLSLHPRKYVGLSQAQIVGSHADSRSFSD-GD 217
KS0_PedH        CSSALVAVHMACESLLRDDCKAAVAGGVNLSIHPKKYIGLSASQILGSHPDSSSFGQ-GD 227
KS0_RhiF        CSSSISAIHAAGEALRNGDCRLAIAGGVNLTLHPKKYIGLSIGKVLGSHASSRSFAD-GD 227
KS0_TaiN        CSSSLVAVNEAWEYLQR-DGRLAVVGGVNLYLDPQQYAHLSRFRFASSGPVCKAFGEGGD 220
KS0_OzmJ        CSSAHVALHEAVESIRSGRCDLALAGAVNLYLHPSTYVELATVGLLSDRDDCASFGAEAA 232
KS0_BryC        CSSSLVALHQACESLRQGTIDMALVGAVNLYLHQDIYLGMCQAKVISDSATPAIFGCDGK 217
KS0_BryD        CSSSLTAIHEACEHLHRQRCELAIAGGVNLYLHPSSYIHLCAGHILSKNNRCSAFGQGGD 235
                * .   .:  *   :       *:... .:         :                  .

KS0_BaeR        GFVRSEGAGVIIIKPLEAAIEDGDHIHMVIKGTGVSH-GGKGMSLHAPNPAGMKAAMKKA 269
KS0_MmpA        GYLRGEGVCSLLLKPLSKAEADGDHIYGLIRNTAVNYNGGDAASIAAPSVSAHSSLVQDC 289
KS0_PedF        GHLRAEGVCSLLLKPLTKALADGDPIYASIKHSAVNFNGQGGASIAAPNVDSHVDLIKSC 287
KS0_OnnI        GHLRAEGVVSLLLKPLSKAEADGDPIYALIKNTACNYNGQGGMSIAAPNVDSHAELIETC 287
KS0_BaeJ        GVVLGEGVGMVLLKTVSQAIEDGDSIYAVVKAASVNN-DGRTAGPATPSLEAQKAVMKTA 273
KS0_BaeL        GVVLGEGVGMVLLKTVSQAQKDGDTIHAVIKAAAMNN-DGRTAGPSAPNMQAQKDVMQSA 271
KS0_TaiK        GTVLSEGAGLVMLKPLAAARAHGDTIYAVLKGLAVNN-DGRTAGPSSPNFAAQQAVMRRA 272
KS0_OzmH        GFTPAEGVGVLLLKPLAAAEAAGDRVHAVLKGIAVNN-DGRTAGPATPNPAAQRGVMARA 271
KS0_LnmI        GWHPGEGVGVLLLRPADAARRAKDTVHGLIEGTRIGH-AGRAPRYGAPHTAALADSLARA 275
KS0_RhiB        GWVAGEGVGALLIRPLEDAMRDGDSILGVIRATAISH-SGKTFRYGAPNADSHALSMRRV 278
KS0_DszC_1      GWVPGEGVGAVLLRRLPEAEERGEHIRCVLKGTALAH-AGKAPRYGMPSTRAQAGSIRDA 285
KS0_ChiD        GWIPGEGVGAVLIRPADDAERSGDHIHALIKATAINH-TGATPRYGMPSAEAQAASIRDV 287
KS0_ChiC        GFVPGEGVGAVLLKPLRQALLDRDPILAVIKGSALNH-AGKTSGFMAPSPAAQADLLERA 290
```

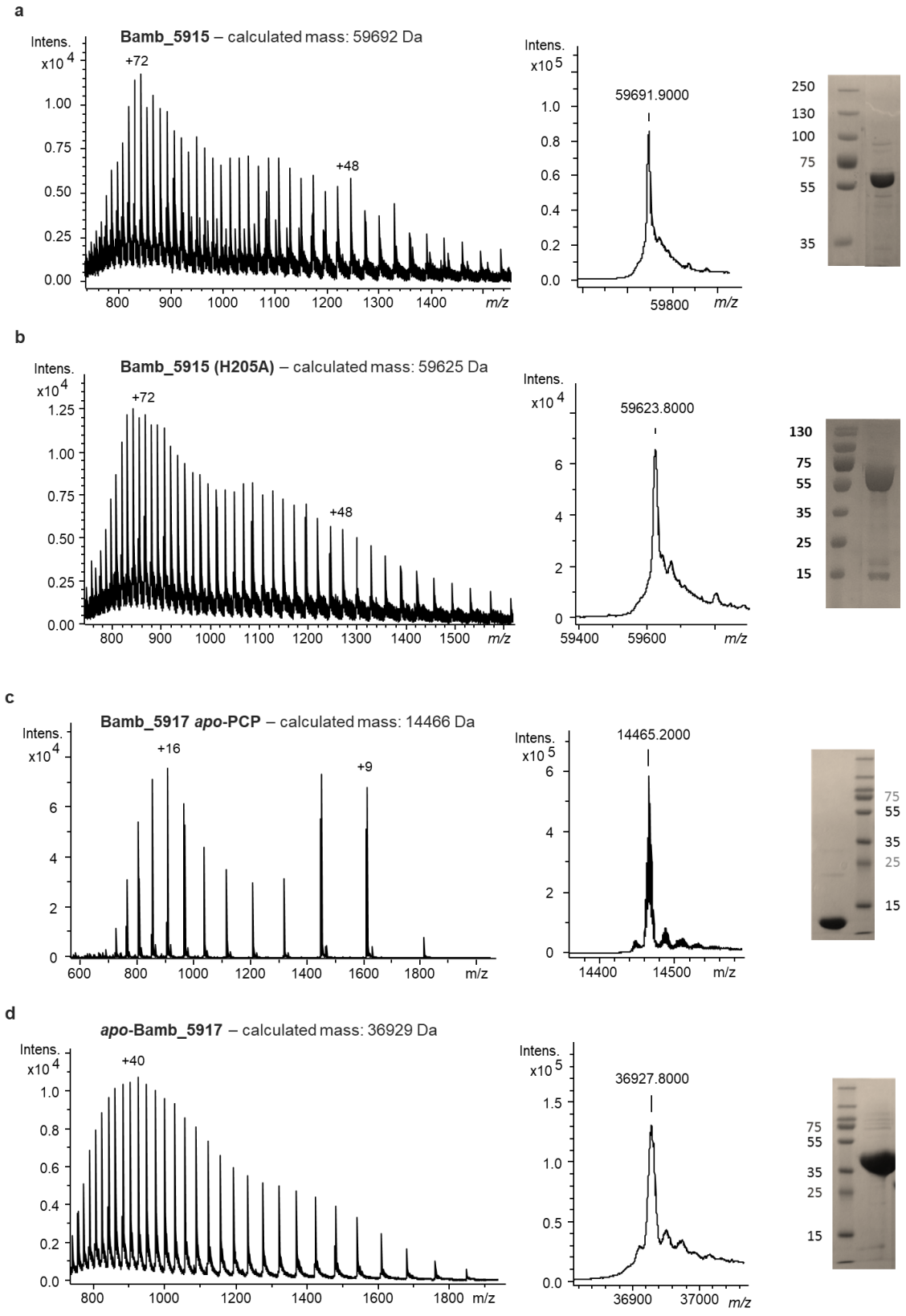

KS0_DszB GMIPGEGVGAVLLKPLDRAVRDRDHIHAIIRSSAVSH-GGASTGFTAPNLKAQSDMFVEA 285
KS0_DszC_2 GYVPGEAVGAVVLKPLDRAVADNDRIYGVIKGSFANH-AGKTAGYGVPSPAAQADLIAAA 287
KS0_ChiF GITVGEGVGALLLKPLALALRDGDRVYAVIKGGFVNH-GGRTHGATVPNPSAQADLIVEA 287
KS0_5919 GYLPAEGVGAVLLKPLARALADDDRILAVIKASSVNH-GGRATGYYAPNANAQVDLMEAS 276
KS0_PedH GYLPSEGVGAVLLKPLREAVADNDTILGVIKSTTINH-SGQSNGYFVPNGAAQTELMVSN 286
KS0_RhiF GYLPAEGVGAVLLKTLADAERDGDQILAVIKSTAVNH-GGHTHGFSMPSAKAEAALIDSN 286
KS0_TaiN GFVPGEGAGAIVLKRLSDAERDGDPIHAVIRGCAVNH-NGRSTSFTASDPARQADVVRDA 279
KS0_OzmJ GIVPGEGVGAVLLKPLRQARRDGDPVHAVIRGSAVNH-NGRTIGFTSPSSQRQAEVIREA 291
KS0_BryC GFTPSEGVGAVVIKRLSDAEKGNDRVLAVIRGSAVNH-SGRTNHYGVPCPRQQAAVIHEA 276
KS0_BryD GFVPGEGVGCVLLKPLSCAERDGDNIYAVILGSHTNH-SGRAGGMGPNLN-AQSDLIIEN 293
KS0_BaeR YEDTDVDPQTVTYIEAHGIASEMADALEFNAIKAGYGESANQ----EESAPCYISTVKPC 325
KS0_MmpA YRRAGIDPRHVSYIEAQGMGNPVADIAEWDALNHGLLALGREQGVQLQEGQCAISTLKPM 349
KS0_PedF YQQARVDPRQVRYIEAQGMGNVLADLVEWQAFNRALTDIARQQRVSLPPGNCLISTLKPM 347
KS0_OnnI YEQVQVDPGEIRYIEAQGMGNPLSDLGEWHAYNQALQSMAKKRGVVLPQGQCAISTLKPM 347
KS0_BaeJ LEKSGKRPEEITHIEANGSGTAVTDLLELKAIQSVYRSTD--------AGPLGIGSVKPN 325
KS0_BaeL LFKSGKKPEDISYIEANGSGSAVTDLLELKAIQSVYRSGQ--------HVPLGIGSIKPN 323
KS0_TaiK LAQSGLRADDVRHVEANGSGSRVTDLLELKSIRAVYGGQSRD------AAWCALGSVKPS 326
KS0_OzmH LAKAGVAADDVTYIETNAAGSQIPDLIELKAIAAVYRDGS--------DTPCSLGSVKPN 323
KS0_LnmI LADASVIPDEVDYVECAAAGAGIADAAELEALGSVLARCAG-------ASPVPVGTLKPN 328
KS0_RhiB LQQAGLSADEIGYVEAAAPGASLADGAEFAAISNVFGARRSD------AP-LLVGSIKAN 331
KS0_DszC_1 LADGGVAASEIDYVECAATGSGIADASEVDALKQAFEGRSPD------GPPCLLGSVKPN 339
KS0_ChiD LRRAGLGPEAVSYVEAAATGAAIADASEIAALIEVFGERQGS------APRVALGSIKPN 341
KS0_ChiC LARANVDPGSVSYIEAQGMGSTLVDAAELAAFTRVLRRG-RR------QGPCLLGSIKPN 343
KS0_DszB IERAGIDPRTISYVEAAANGAPLGDPIEVNALTRAFRRFTAD------TGFCALGTVKSN 339
KS0_DszC_2 LRRTGIDPETIGYIEVAANGSSLGDAIELAGLTQAFRRFTAR------KHFCAVGSVKSN 341
KS0_ChiF FRRAGVRPDAVSYIEVAANGSPLGDSIEIAGLKQAFRRFTVE------RGFCALGSVKSS 341
KS0_5919 FRKAGVSPESIDYIEAAANGTSLGDAVELRALARVFDGTARD------GRRVPIGTVKSN 330
KS0_PedH FTKAGIDPRTLSYVESAANGSSLGDAIEINALTAGFGRYTAD------KQFCALGSVKSN 340
KS0_RhiF FKRAGVDPRTISYVEAAANGSAMGDAIELSALNRVFGQAGVA------HQSCAIGSVKSN 340
KS0_TaiN LTRAGVDPRTIGYVEAAANGHAMGDAIEMTGLGKVFAACDGV------SGTRAIGSVKAN 333
KS0_OzmJ LRDARVDPRTIGYVEATANGSEIGDAVEMTSLTQVFEDRPDA------RGPYRIGSLKPN 345
KS0_BryC IDNANVDPRSIAYIESAANGSEMGDAIEMSALTKVFQTHRDN-----GKAQYSIGSLKSI 331
KS0_BryD LRQCGIAPDTIGYVESASNGSHLGDSIELRALDKAFSQHTKK------RDFCAIGSVKPN 347
KS0_BaeR IGHGELASGLAALIKVAMAMKHHTIPGIPRFTAANEQMAIQKSRFRFTEDNQEWTQLTD- 384
KS0_MmpA SGHMHAASAIGALFKIIRSLQTEKIHKILDFEQPNLHLHTAGQPCRLATHTVDWPRQ--- 406
KS0_PedF MGHMESASALGALFKVIRSLHTRTIHKIAHFTQYHPDMDYQGQPCAIAGETVAWPQM--- 404
KS0_OnnI MGHMESVSSLGAIMKVIRSFKTNTIHKILNVQEISPDLDPQGMPCRLLTETEPWPEQ--- 404
KS0_BaeJ IGHPLCAEGIASFIKVVLMLKEKSFIPFLSGEHEHTHFDREKANIQFTRTLADWPSPIP- 384
KS0_BaeL IGHPLCAEGIASFIKVVLMLKHKQTVPFLSGDEPMPHFDITKTDFHFHKTAGEWDAARP- 382
KS0_TaiK IGHTLCAQGIAAFIKSVLMLHHRSVPPFLSGQQPMQHSPIERSRLRFVRETIPFDVAAP- 385
KS0_OzmH IGHPQCAEGIAGVIKTVLMLRNRAIVPFLSGRQPLEHFDFAATPLRFERALTPWPDAPL- 382
KS0_LnmI IGHLEAASGLSQLIKVLLQIRHGRIAPTLVSGELSPLVDWDGLPVELVDTPRALTPRAA- 387
KS0_RhiB IGHLESASALSQITKVLMQLKHRQIAPTLGCNPLSPMICLDDNHLAIADQLSDW------ 385
KS0_DszC_1 IGHLESASALSQLTKVILQLEHGEIAPTLHTEPRNPLIQLDGTPFRINRALSPWPRAAG- 398
KS0_ChiD IGHLESASAMSQLAKVLLQIQHKTLAPHVLSGALNPMIPWDRAPFWVPEQPAAWQPR--- 398
KS0_ChiC IGHLEGAAGISQLTKVVHQLRSRQIAPSLHADPVNPEVGFDASLFRIPGALEPWPMPVV- 402
KS0_DszB IGHLEGASGVSQLAKVLLQLRHGALAPTINAEPRNPNLHLDDTPFYLQERLDDWRRPIIS 399
KS0_DszC_2 IGHPEAASGIAQLTKVLGQLHHRTLVPTLHAEPHNPNIDLRDSPFYVQRELGPWTAPTLA 401
KS0_ChiF IGHLEAASGVSQVTKVAYQLHHRTLVPTLNSEPLNPNIRLDDSPFYVQRERAPWRPAV-- 399
KS0_5919 IGHPEAASGIAQLTKVILQMQHETLVPSIKTEPVNPNLDLAHTPFRLLSRQAAWPSDP-- 388
KS0_PedH IGHGEAASGIAQLIKVLLQLKHRQLVPTIKAQPLNSNIDFTHTPFCLQRRLEPWRRPSLA 400
KS0_RhiF IGHAEAASGMSQLSKLVLQLQHQQLAPSLLLGSLNPKLDFENSPFVLQRELGHWPQPVVE 400
KS0_TaiN IGHCEAASGMSQLTKVVMAMRDGVLAPTLRDGTRNPNIAFERLPFEVQEQAAPWRRLIV- 392
KS0_OzmJ IGHGEASAGMAQLFKVILALRHRTLPPTRLPGEYNPAIDIDRLPFELSGAPVAWDQVTV- 404
KS0_BryC LGHGEAVSGMAQFMKVVLQLRNKSLCPSPDPQQKNPNIHFENLPFELQTELDEWRQLTI- 390
KS0_BryD IGHLESASGMSQLTKVLLQLRHKQLVPSIHAQPLNSNIDFEDTAFRLQKEVEEWKRLIVQ 407
KS0_BaeR ---HTGRPIPRRAAINSYGFGGMNAHVVLEQY- 413
KS0_MmpA -------ATPRLAGLHSYGAGGNNAHILVEE-- 430
KS0_PedF -------EGLRLAGIHCYGMGGVNAHLLVEESV 430
KS0_OnnI -------ARPRLAGLHSFGIGGNNVHILLEEY- 429
KS0_BaeJ -----------AAGINCFADGGTNAHVIVEAWQ 406
KS0_BaeL -----------SAAINCFADGGTNAHVILE--- 401

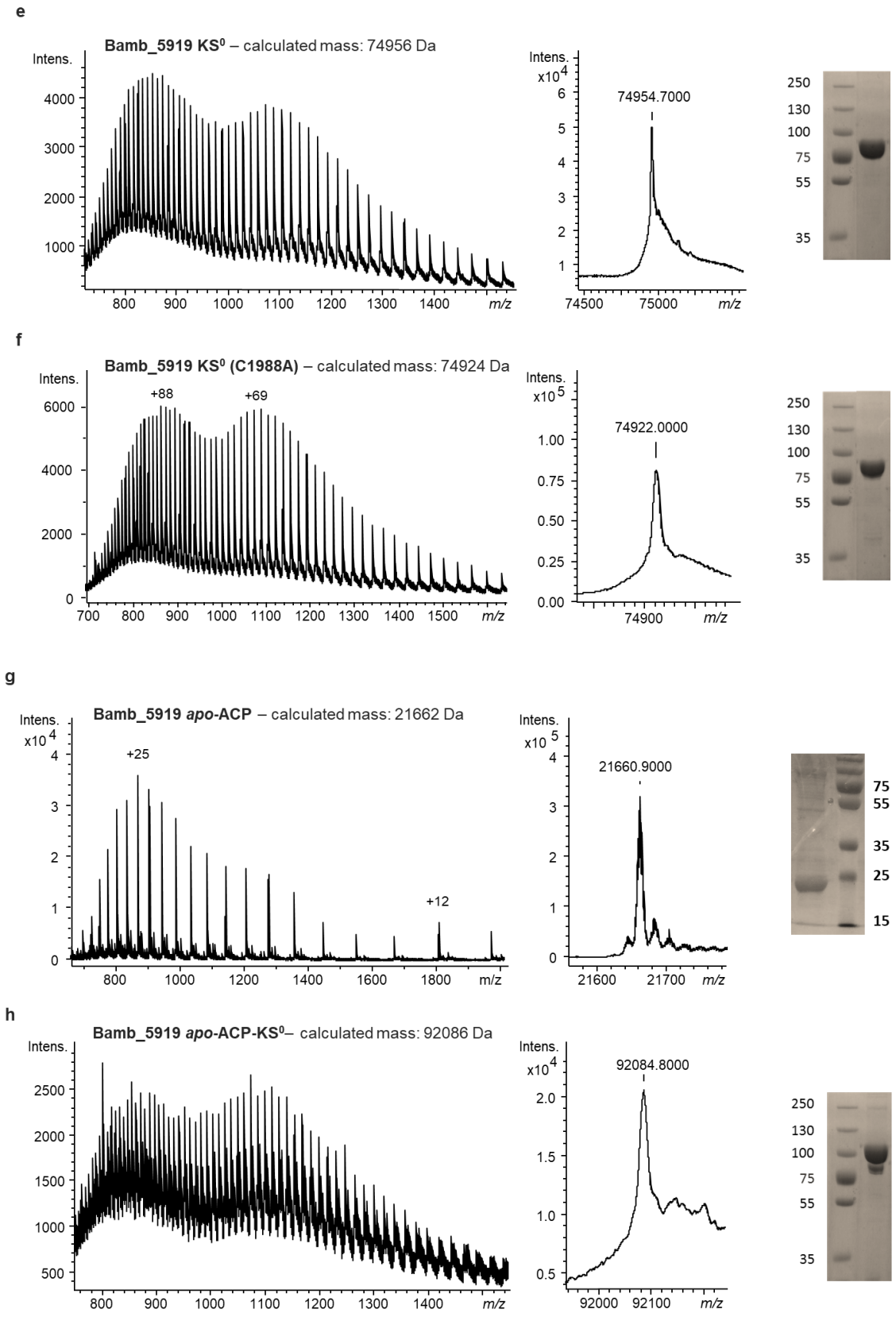
```
KS0_TaiK        -----------AVALNCFADGGTNVHAVL----  403
KS0_OzmH        -----------LAAVSSFADGGTNAHAVL----  400
KS0_LnmI        -------DGRATVLVNAVGATGSYGHVVV----  409
KS0_RhiB        -------RGPQRALINAFGASGSGGHLIVE---  408
KS0_DszC_1      -----ADAPPRRALINAFGATGSSAHAVVEEY-  425
KS0_ChiD        -------SGPRRALVNAFGATGSLGHAVIEE--  422
KS0_ChiC        --DGHAEPSTRRACISSFGAGGSGVYLIVE---  430
KS0_DszB        -----GREVPRRAMINSFGAGGGYATLVVEEH-  426
KS0_DszC_2      -GEGGTAELPRRAAISSFGAGGANTHLLVEEYS  433
KS0_ChiF        ------EGEPLRAAVASFGAGGANAYLIL----  422
KS0_5919        -------ARPRRATVSSFGASGANAHLI-----  409
KS0_PedH        LGDGPMREYPLRATVSSFGAGGSNAHLILEEF-  432
KS0_RhiF        -TDGVSRQYPRRAALSAFGAGGSNAHLVLEEY-  431
KS0_TaiN        ----DGSEVPRRAGVTSIGGGGVNAHVVLEEYV  421
KS0_OzmJ        ----DGALVPRRAGITGLGGGGTNAHVVL----  429
KS0_BryC        ----ADKKIPRRAGITALGAGGVNAHMIVEEYQ  419
KS0_BryD        -VNGENKEIPRRAAINSFGAGGVNANLIIQEY-  438
                             . :   .  *     :
```

Supplementary Fig. 2: Sequence alignment of the Bamb_5919 KS^0 domain with other KS and KS^0 domains. Universally conserved amino acid residues are highlighted in cyan. The Cys and His residues in the conserved C-H-H catalytic triad of KS domains are highlighted in green. Instances where the first of these His residues has been mutated are highlighted in yellow. Alignments were created using Clustal Omega. (**a**) Alignment of the Bamb_5919 KS^0 domain (KS^0_5919 – aa residues 1820-2238) with the Bamb_5925 (KS_5925 – aa residues 828-1252), Bamb_5924 (KS_5924_1 –aa residues 47-474, KS_5924_2 – aa residues 1455-1879), Bamb_5923 (KS_5923 – aa residues 665-1091), Bamb_5922 (KS_5922 – aa residues 44-457), Bamb_5921 (KS_5921_1 – aa residues 34-458, KS_5921_2 – aa residues 1514-1921), Bamb_5920 (KS_5920_1 – aa residues 33-455, KS_5920_2 – aa residues 1792-2208) and Bamb_5919 (KS_5929 – aa residues 33-452) KS domains from the enacyloxin PKS. (**b**) Alignment of the Bamb_5919 KS^0 domain (KS^0_5919 – aa residues 1820-2238) with the GbnD1 (KS_GbnD1_1 – aa residues 35-457, KS_GbnD1_2 – aa residues 1397-1818, KS_GbnD1_3 – aa residues 2576-3014, KS_GbnD1_4 – aa residues 4090-4491), GbnD2 (KS_GbnD2_1 – aa residues 198-632, KS_GbnD2_2 – aa residues 1720-2153, KS_GbnD2_3 – aa residues 2724-3139), GbnD3 (KS_GbnD3_1 – aa residues 557-969, KS_GbnD3_2 – aa residues 1785-2215, KS_GbnD3_3 – aa residues 3038-3453), GbnD4 (KS_GbnD4_1 – aa residues 902-1333, KS_GbnD4_2 – aa residues 2168-2587, KS_GbnD4_3 – aa residues 3266-3676, KS_GbnD4_4 – aa residues 4506-4923), GbnD5 (KS_GbnD5_1 – aa residues 492-906, KS_GbnD5_2 – aa residues 1702-2102) and GbnD6 (KS_GbnD6_1 – aa residues 424-843, KS_GbnD6_2 – aa residues 1633-2067, KS_GbnD6_3 – aa residues 3619-4053, KS_GbnD6_4 – aa residues 4907-5334) KS domains from the gladiolin PKS. (**c**) Alignment of KS^0 domains from the Bamb_5919 (KS^0_5919 – aa residues 1820-2238) subunit of the enacyloxin PKS, the BaeJ (KS^0_BaeJ – aa residues 4532-4937), BaeL (KS^0_BaeL – aa residues 4027-4427) and BaeR (KS^0_BaeR – aa residues 1488-1900) subunits of the bacillaene PKS, the PedF (KS^0_PedF – aa residues 44-473) and PedH (KS^0_PedH – aa residues 4182-4613) subunits of the pederin PKS, the LnmI (KS^0_LnmI – aa residues 1934-2342) subunit of the leinamycin PKS, The MmpA (KS^0_MmpA – aa residues 27-456) subunit of the mupirocin PKS, the OnnI (KS^0_OnnI – aa residues 76-504) subunit of the onnamide PKS, the BryC (KS^0_BryC – aa r residues 4621-5039) and BryD (KS^0_BryD – aa residues 2204-2641) subunits of the bryostatin PKS, the DszB (KS^0_DszB – aa residues 5502-5927), DszC (KS^0_DszC_1 – aa residues 1889-2313, KS^0_DszC_2 – aa residues 2671-3103) subunits of the disorazole PKS, the RhiB (KS^0_RhiB – aa residues 1041-1448) and RhiF (KS^0_RhiF – aa residues 1551-1981) subunits of the rhizoxin PKS, the ChiC (KS^0_ChiC – aa residues 7634-8063), ChiD (KS^0_ChiD – aa residues 2034-2455) and ChiF (KS^0_ChiF – aa residues 4825-5246) subunits of the chivosazol PKS, the TaiN (KS^0_TaiN – aa residues 659-1079) and TaiK (KS^0_TaiK – aa residues 1454-1856) subunits of the thailandamide PKS, and the OzmH (KS^0_OzmH – aa residues 6122-6521) and OzmJ (KS^0_OzmJ – aa residues 911-1339) subunits of the oxazolomycin PKS.

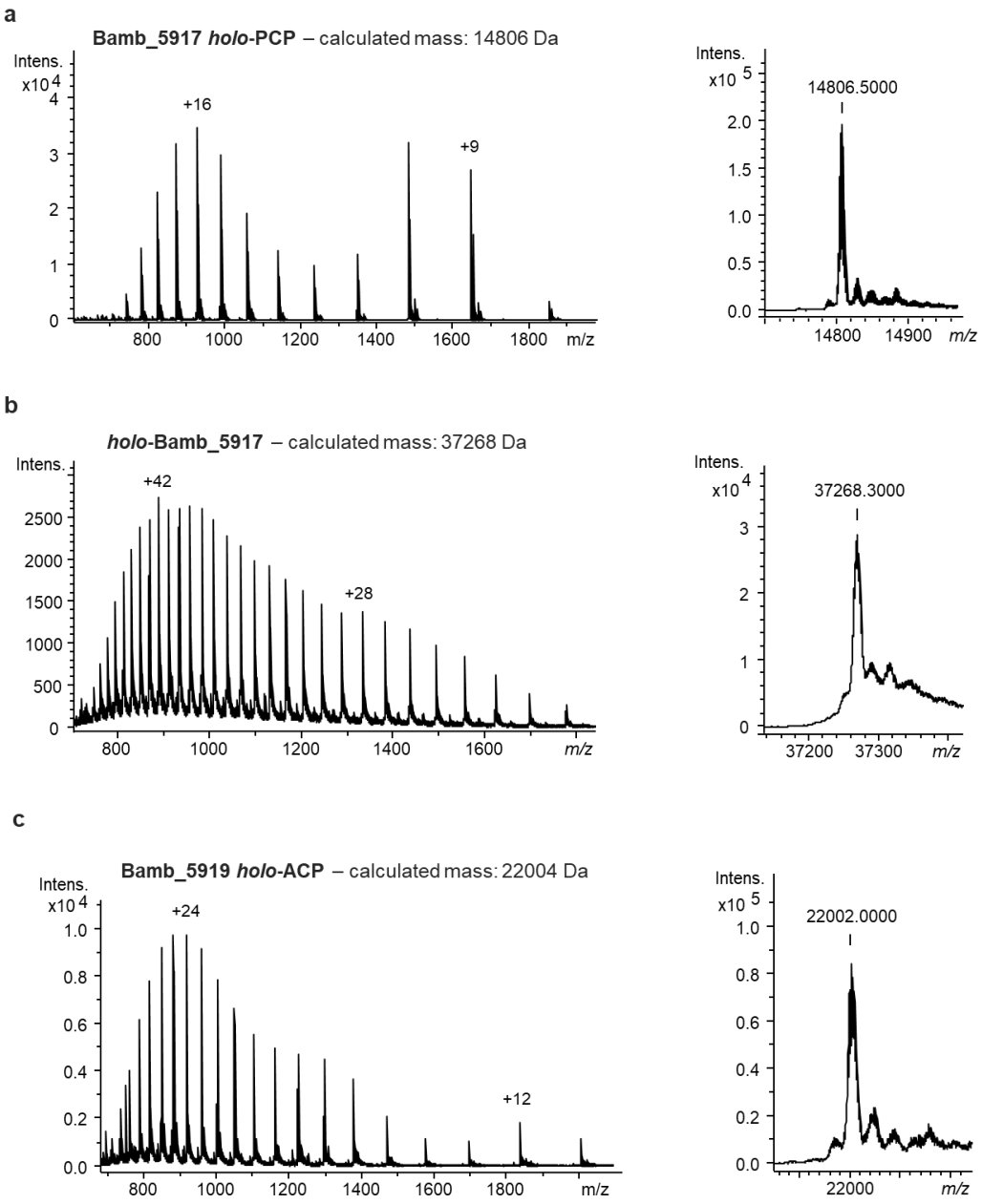


Supplementary Fig. 3: Comparison of the enacyloxin biosynthetic gene clusters from *Burkholderia ambifaria* AMMD and *B. ambifaria* BCC0203. Grey bars link homologous genes. The number on each bar indicates the percentage of amino acid sequence identity between the gene products.

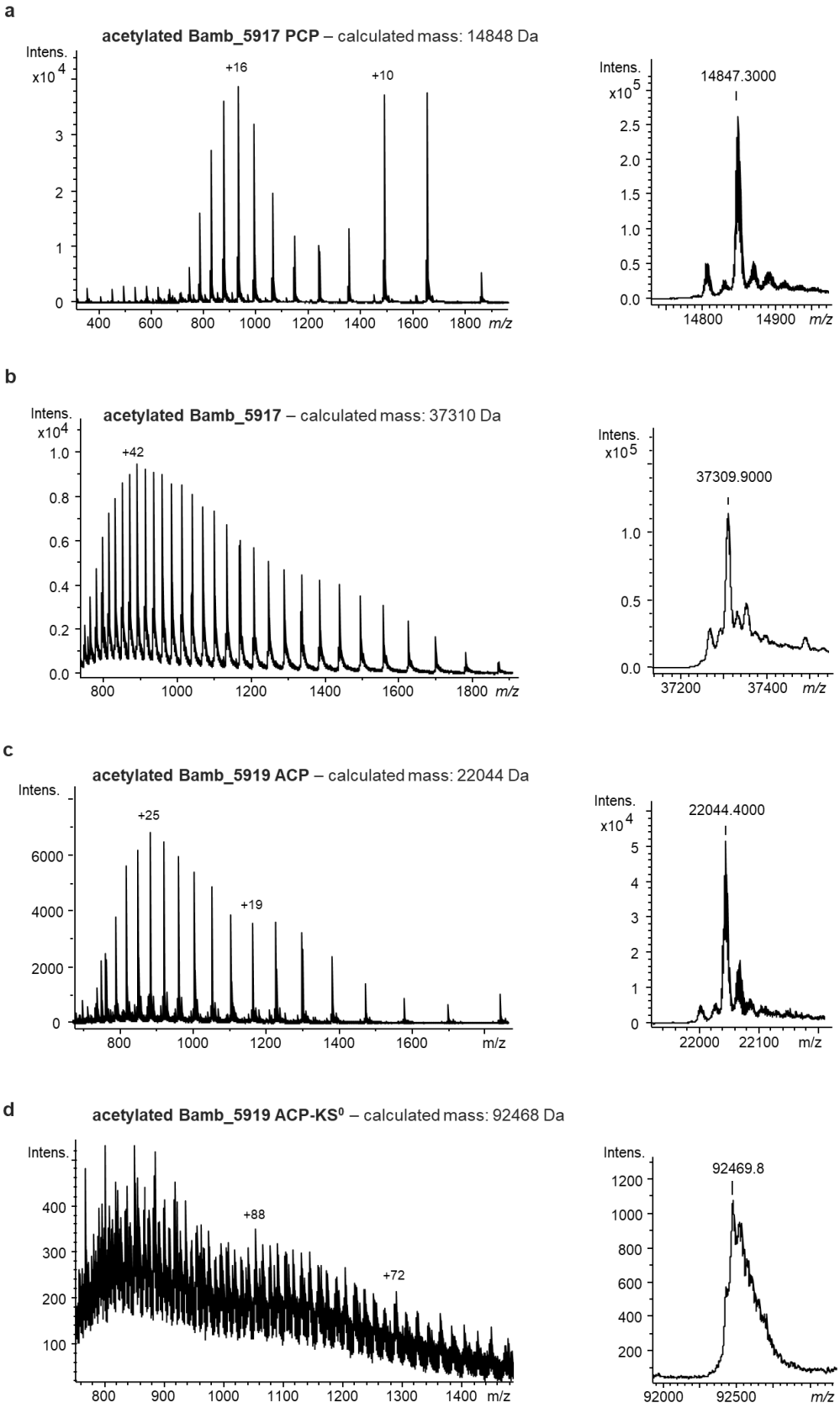

a
Bamb_5915 – calculated mass: 59692 Da
Intens.
x10 4
+72
+48
m/z
Intens.
x10 5
59691.9000
59800
m/z
250
130
100
75
55
35
b
Bamb_5915 (H205A) – calculated mass: 59625 Da
Intens.
x10 4
+72
+48
m/z
Intens.
x10 4
59623.8000
59400
59600
m/z
130
75
55
35
25
15
c
Bamb_5917 apo-PCP – calculated mass: 14466 Da
Intens.
x10 4
+16
+9
m/z
Intens.
x10 5
14465.2000
14400
14500
m/z
75
55
35
25
15
d
apo-Bamb_5917 – calculated mass: 36929 Da
Intens.
x10 4
+40
m/z
Intens.
x10 5
36927.8000
36900
37000
m/z
75
55
35
25
15



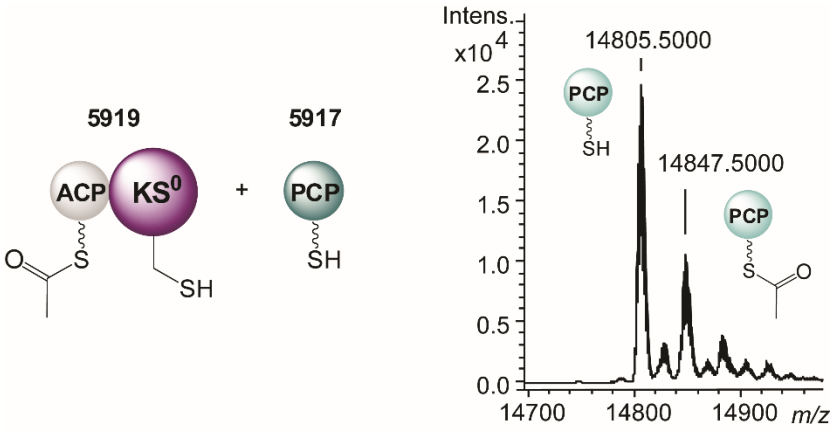
Supplementary Fig. 4: SDS-PAGE and mass spectrometry analysis of purified recombinant proteins. Measured mass spectra (left), deconvoluted mass spectra (middle) and SDS-PAGE analysis (right) of purified His_6-tagged (**a**) Bamb_5915, (**b**) Bamb_5915 (H205A), (**c**) Bamb_5917 *apo*-PCP domain, (**d**) Bamb_5917, (**e**) Bamb_5919 KS^0 domain, (**f**) Bamb_5919 KS^0 domain (C1988A), (**g**) Bamb_5919 *apo*-ACP domain and (**h**) Bamb_5919 *apo*-ACP-KS^0 di-domain.



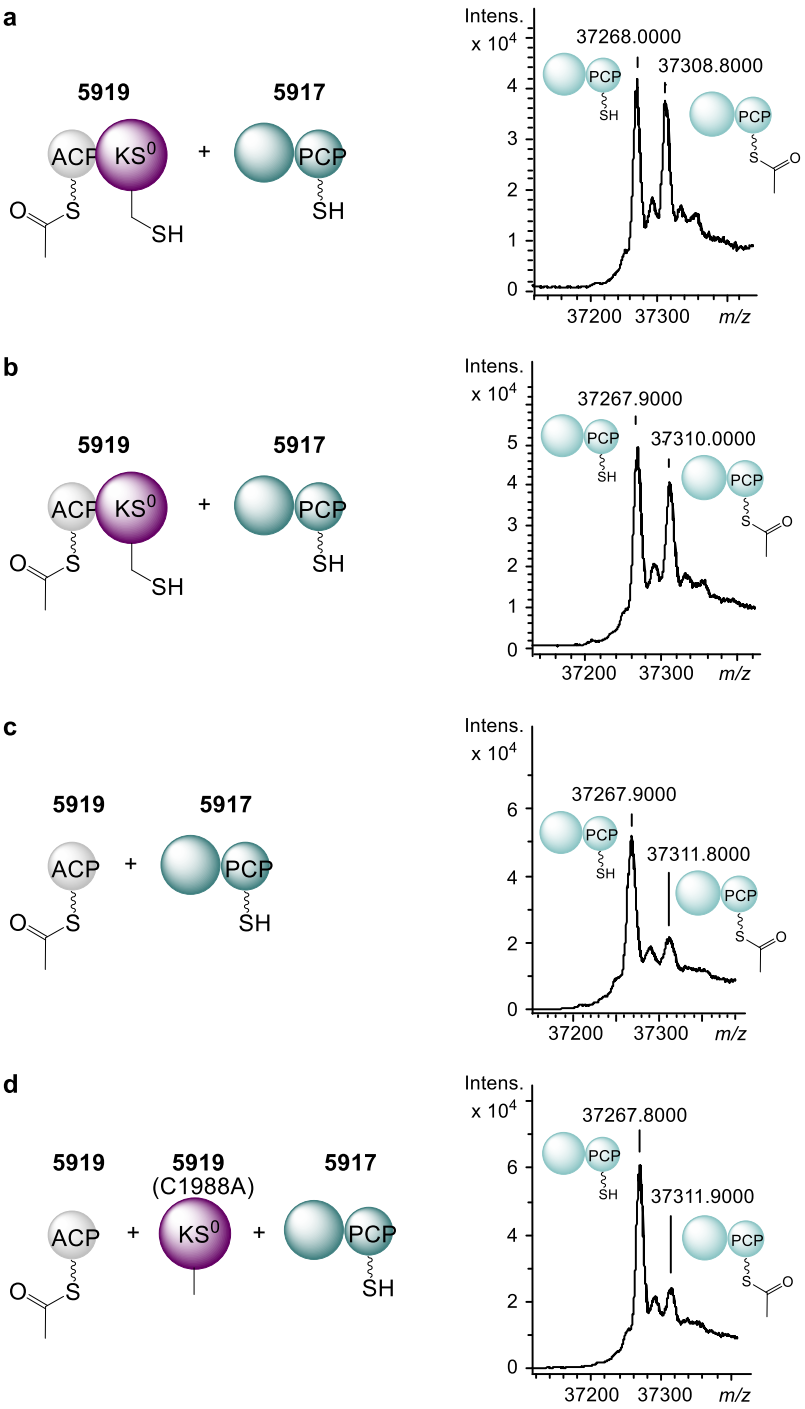
Supplementary Fig. 5: *In vitro* conversion of purified *apo*-ACP and –PCP domains to their *holo*-form. Measured (left) and deconvoluted (right) mass spectra of (**a**) the Bamb_5917 PCP domain, (**b**) Bamb_5917 and (**c**) the Bamb_5919 ACP domain following incubation with Sfp and coenzyme A.



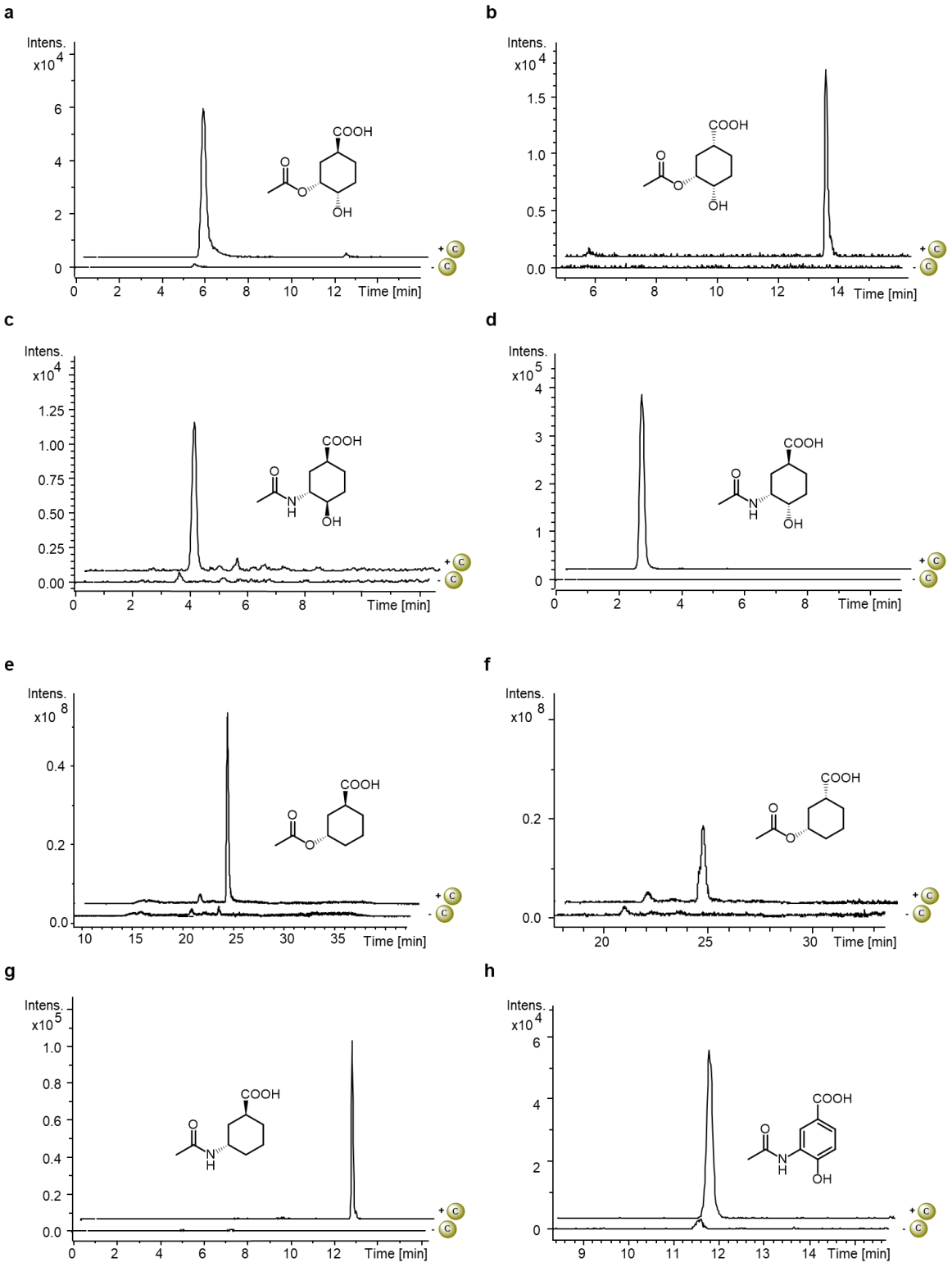
Supplementary Fig. 6: *In vitro* conversion of purified *apo*-ACP and –PCP domains to their acetylated *holo*-form. Measured (left) and deconvoluted (right) mass spectra of (**a**) the Bamb_5917 PCP domain, (**b**) Bamb_5917, (**c**) the Bamb_5919 ACP domain and (**d**) the Bamb_5919 ACP-KS^0 di-domain following treatment with Sfp and acetyl-CoA.

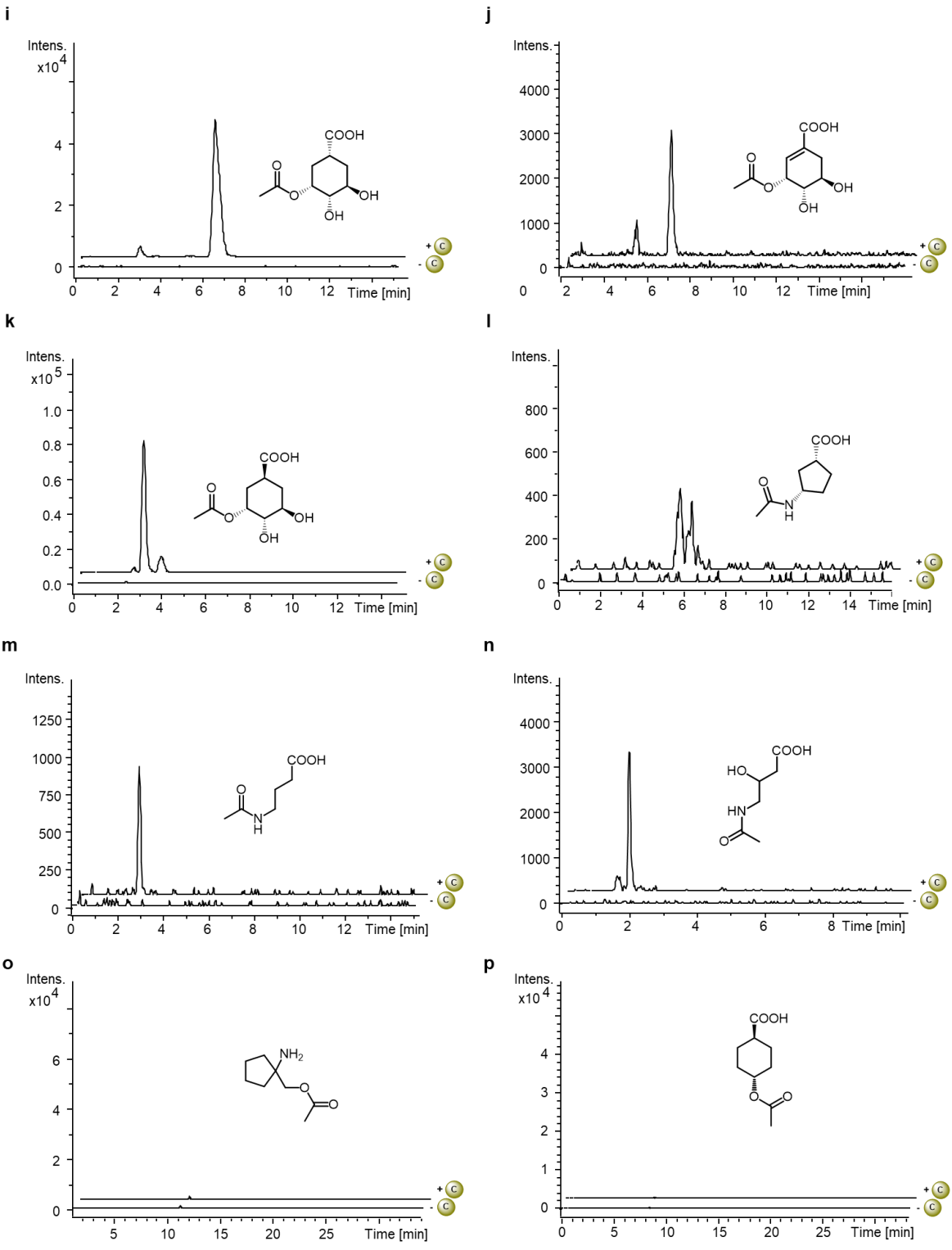


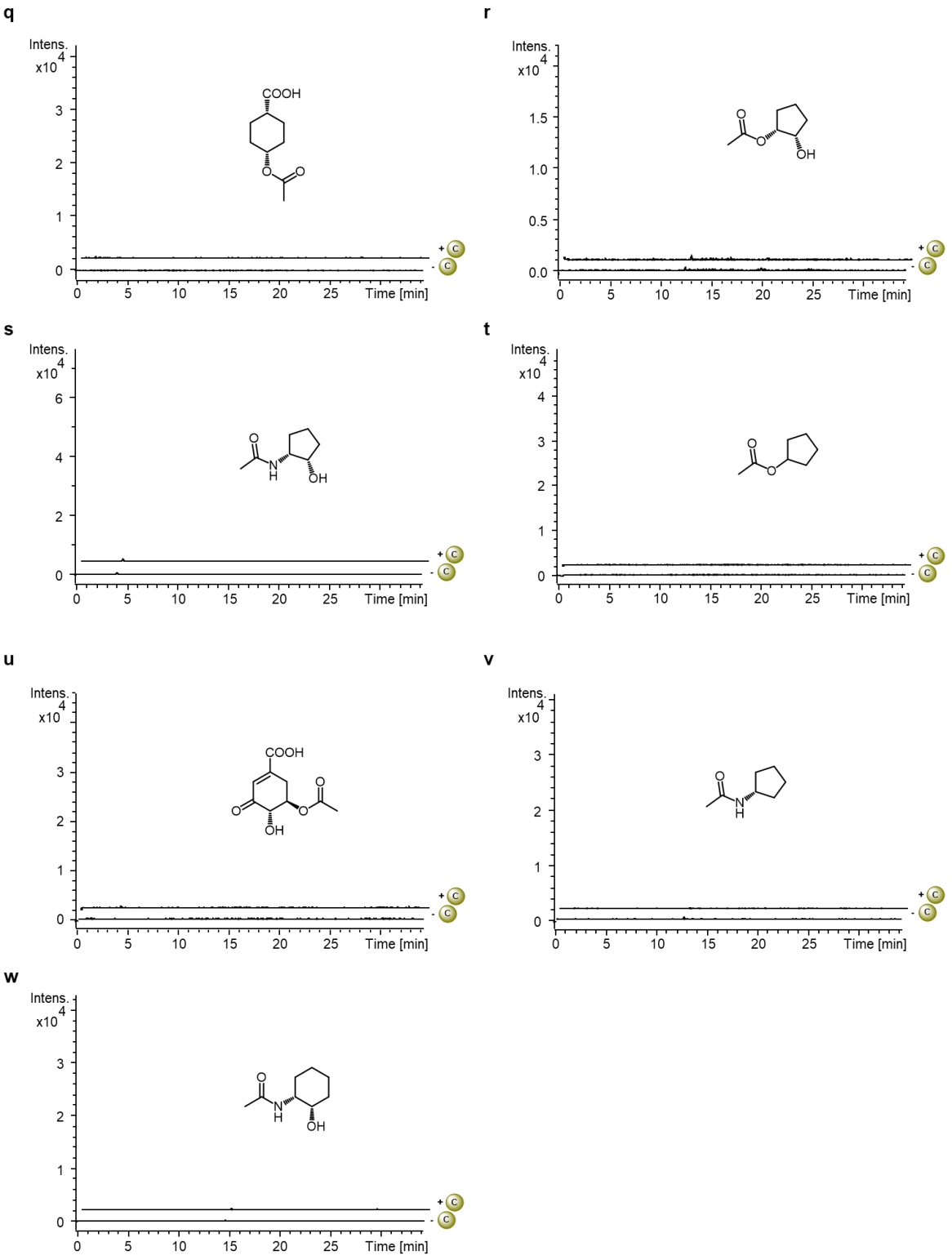
Supplementary Fig. 7: Acyl transfer assay with the acetylated Bamb_5919 ACP-KS^0 di-domain and the Bamb_5917 *holo*-PCP domain. Experimental setup (left) and deconvoluted mass spectrum of the Bamb_5917 *holo*-PCP domain (right) following incubation with the acetylated Bamb_5919 ACP-KS^0 di-domain. The ACP-bound acetyl group is transferred onto the PCP domain as indicated by a mass shift of +42 Da.



Supplementary Fig. 8: Acyl transfer assays with the full-length Bamb_5917. Experimental set up (left) and deconvoluted mass spectra of *holo*-Bamb_5917 (right) following incubation with (**a**) the acetylated Bamb_5919 ACP domain and the Bamb_5919 KS^0 domain, (**b**) the acetylated Bamb_5919 ACP-KS^0 didomain, (**c**) the acetylated ACP domain alone, and (**d**) the acetylated ACP domain and the C1988A mutant of the KS^0 domain. The data show that the KS^0 domain catalyses transfer of an acyl group from the Bamb_5919 ACP domain to Bamb_5917 (47.4 ± 2.1% and 45.3 ± 1.8% acetylation, respectively). Significantly reduced levels of acyl transfer are observed when the KS^0 domain is omitted from the reactions (12.2 ± 1.1%), or when the C1988A mutant of the KS^0 domain is employed (14.9 ± 1.1%).

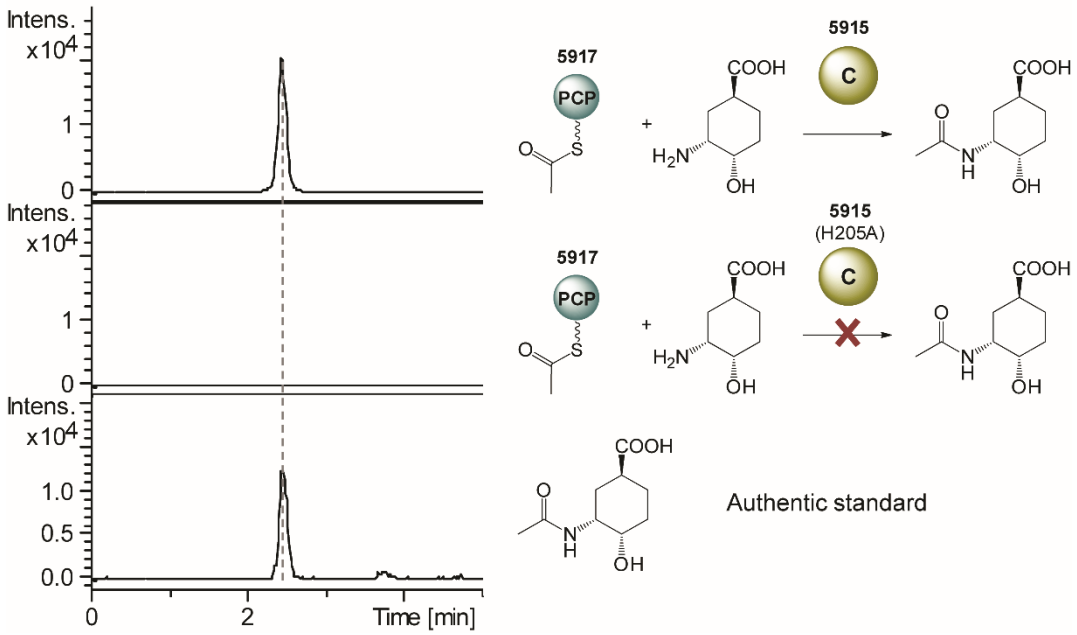

a
Intens.
x10 4
Time [min]
COOH
O
O
OH
b
Intens.
x10 4
Time [min]
COOH
O
O
OH
c
Intens.
x10 4
Time [min]
COOH
O
N
H
OH
d
Intens.
x10 5
Time [min]
COOH
O
N
H
OH
e
Intens.
x10 8
Time [min]
COOH
O
O
f
Intens.
x10 8
Time [min]
COOH
O
O
g
Intens.
x10 5
Time [min]
COOH
O
N
H
h
Intens.
x10 4
Time [min]
COOH
O
N
H
OH


i
Intens.
x10^4
COOH
OH
OH
O
O
Time [min]
j
Intens.
COOH
OH
OH
Time [min]
k
Intens.
x10^5
COOH
OH
OH
Time [min]
l
Intens.
COOH
N
H
Time [min]
m
Intens.
COOH
N
H
Time [min]
n
Intens.
COOH
HO
HN
Time [min]
o
Intens.
x10^4
NH2
Time [min]
p
Intens.
x10^4
COOH
Time [min]

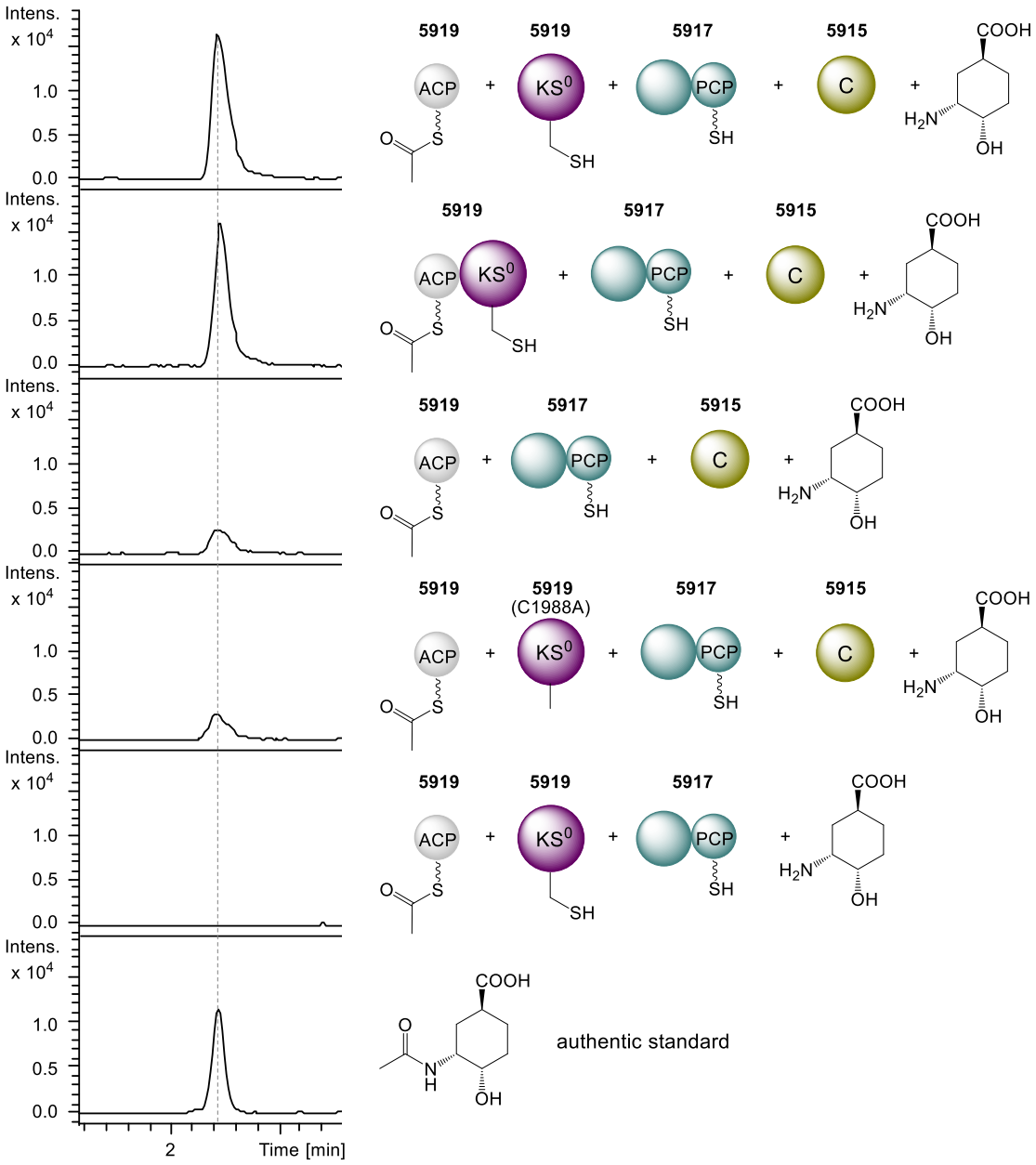


Supplementary Fig. 9: Acyl acceptor tolerance of Bamb_5915. DHCCA analogues were incubated with the acetylated Bamb_5917 PCP domain and Bamb_5915, and formation of an acetylated product was determined using UHPLC-ESI-Q-TOF-MS (chromatograms labelled +C). Bamb_5915 was omitted from negative control reactions (chromatograms labelled –C). (**a**) Extracted ion chromatograms (EICs) at *m/z* = 225.0733 ± 0.005, corresponding to $[M+Na]^+$ for the mono-acetylated derivative of (±)-(1*S*,3*R*,4*S*)-3,4-dihydroxycyclohexane-1-carboxylic acid. (**b**) EICs at *m/z* = 225.0733 ± 0.005, corresponding to $[M+Na]^+$ for the mono-acetylated derivative of (±)-(1*R*,3*R*,4*S*)-3,4-dihydroxycyclohexane-1-carboxylic acid. (**c**) EICs at *m/z* = 224.0893 ± 0.005, corresponding

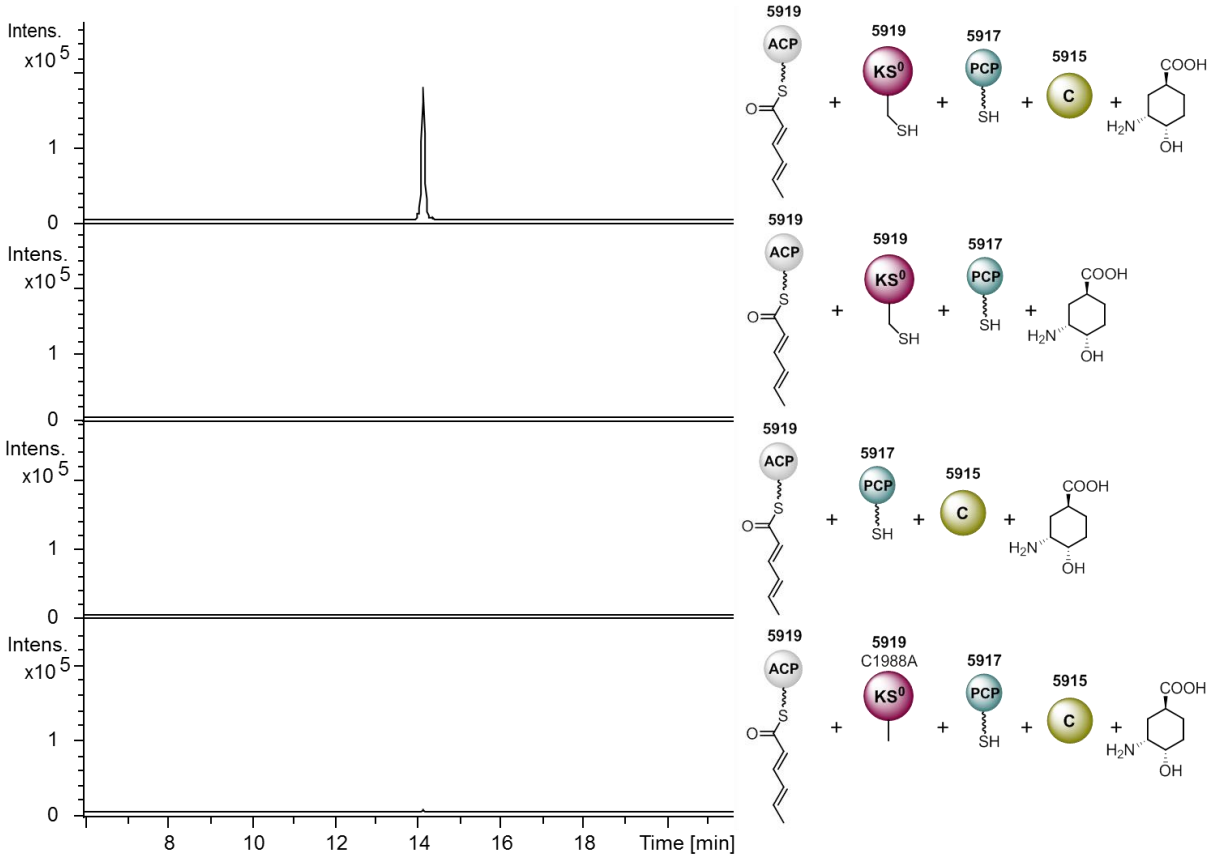
to $[M+Na]^+$ for the mono-acetylated derivative of (±)-(1*S*,3*R*,4*R*)-3-amino-4-hydroxycyclohexane-1-carboxylic acid. (**d**) EICs at *m/z* = 224.0893 ± 0.005, corresponding to $[M+Na]^+$ for the mono-acetylated derivative of (±)-(1*S*,3*R*,4*S*)-3-amino-4-hydroxycyclohexane-1-carboxylic acid. (**e**) EICs at *m/z* = 209.07845 ± 0.005, corresponding to $[M+Na]^+$ for the mono-acetylated derivative of (±)-(1*R*,3*R*)-3-hydroxycyclohexanecarboxylic acid. (**f**) EICs at *m/z* = 209.07845 ± 0.005, corresponding to $[M+Na]^+$ for the mono-acetylated derivative of (±)-(1*S*,3*R*)-3-hydroxycyclohexanecarboxylic acid. (**g**) EICs at *m/z* = 208.0944 ± 0.005, corresponding to $[M+Na]^+$ for (±)-(1*R*,3*R*)-3-acetamidocyclohexanecarboxylic acid. (**h**) EICs at *m/z* = 218.0424 ± 0.005, corresponding to $[M+Na]^+$ for 3-acetamido-4-hydroxybenzoic acid. (**i**) EICs at *m/z* = 241.0682 ± 0.005, corresponding to $[M+Na]^+$ for (±)-(1*R*,3*R*,4*S*,5*R*)-3-acetoxy-4,5-dihydroxycyclohexanecarboxylic acid. (**j**) EICs at *m/z* = 239.0526 ± 0.005, corresponding to $[M+Na]^+$ for (±)-(3*R*,4*S*,5*R*)-3-acetoxy-4,5-dihydroxy-1-cyclohexenecarboxylic acid. (**k**) EICs at *m/z* = 241.0682 ± 0.005, corresponding to $[M+Na]^+$ for (±)-(1*S*,3*R*,4*S*,5*R*)-3-acetoxy-4,5-dihydroxycyclohexanecarboxylic acid. (**l**) EICs at *m/z* = 194.0787 ± 0.005, corresponding to $[M+Na]^+$ for (±)-(1*R*,3*S*)-3-acetamidocyclopentanecarboxylic acid. (**m**) EICs at *m/z* = 168.0631 ± 0.005, corresponding to $[M+Na]^+$ for 4-acetamidobutanoic acid. (**n**) EICs at *m/z* = 184.0580 ± 0.005, corresponding to $[M+Na]^+$ for (±)-4-acetamido-3-hydroxybutanoic acid. (**o**) EICs at *m/z* = 180.0990 ± 0.005, corresponding to $[M+Na]^+$ for the mono-acetylated derivative of (1-aminocyclopentyl)methanol. (**p**) EICs at *m/z* = 209.0790 ± 0.005, corresponding to $[M+Na]^+$ for *anti*-4-acetoxycyclohexanecarboxylic acid. (**q**) EICs at *m/z* = 209.0790 ± 0.005, corresponding to $[M+Na]^+$ for *syn*-4-acetoxycyclohexanecarboxylic acid. (**r**) EICs at *m/z* = 167.0684 ± 0.005, corresponding to $[M+Na]^+$ for the mono-acetylated derivative of *syn*-1,2-cyclopentanediol. (**s**) EICs at *m/z* = 166.0844 ± 0.005, corresponding to $[M+Na]^+$ for the mono-acetylated derivative of (±)-(1*S*,2*R*)-2-aminocyclopentanol. (**t**) EICs at *m/z* = 151.0735 ± 0.005, corresponding to $[M+Na]^+$ for the mono-acetylated derivative of cyclopentanol. (**u**) EICs at *m/z* = 237.03753 ± 0.005, corresponding to $[M+Na]^+$ for the mono-acetylated derivative of 3-dehydroshikimate. (**v**) EICs at *m/z* = 150.0895 ± 0.005, corresponding to $[M+Na]^+$ for the monoacetylated derivative of cyclopentylamine. (**w**) EICs at *m/z* = 180.1007 ± 0.005, corresponding to $[M+Na]^+$ for the mono-acetylated derivative of (±)-(1S, 2R)-2-aminocyclohexanol.



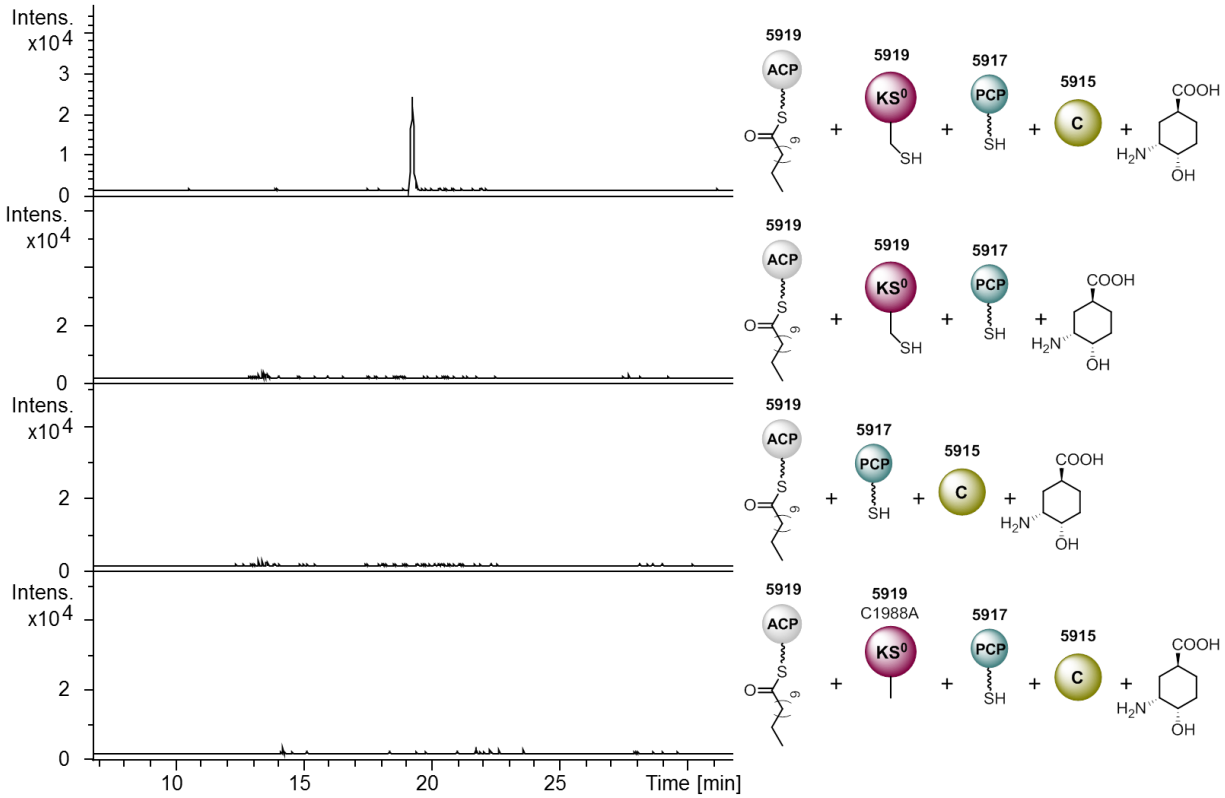
Supplementary Fig. 10: Mutation of the active site His residue (H205) in Bamb_5915 to Ala abolishes the condensation reaction. Extracted ion chromatograms at *m/z* = 224.0893 ± 0.005 (corresponding to $[M+Na]^+$ for *N*-acetyl-AHCCA) from UHPLC-ESI-Q-TOF-MS analyses of the Bamb_5915-catalysed acetylation of AHCCA with the acetylated Bamb_5917 PCP domain (top chromatogram), a control reaction in which Bamb_5915 was replaced with the H205A mutant (middle chromatogram), and the authentic standard of N-acetyl-AHCCA (bottom chromatogram).



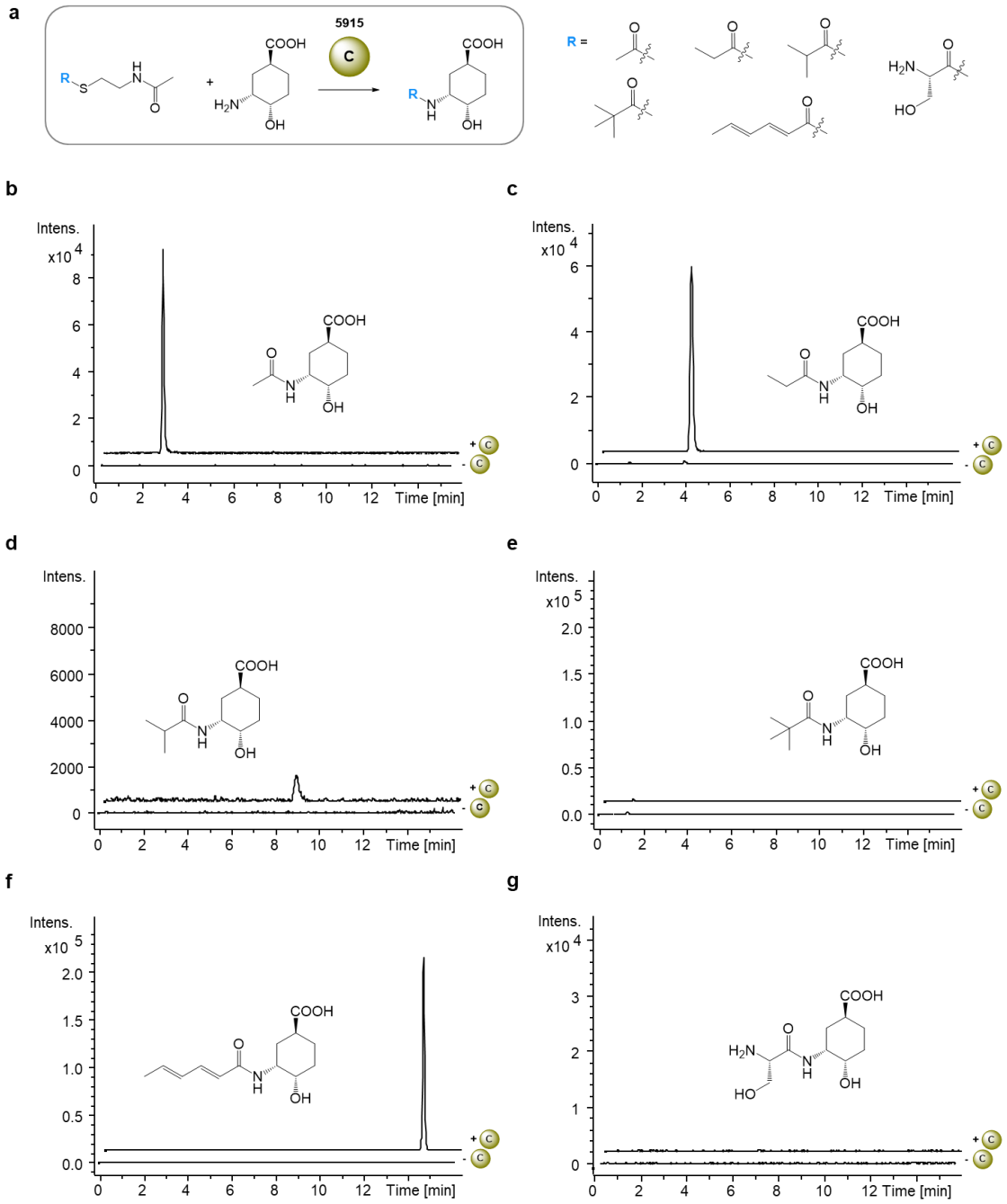
Supplementary Fig. 11: *In vitro* reconstitution of the enacyloxin chain release reaction using full-length Bamb_5917. Experimental setup (right) and extracted ion chromatograms (EICs) at *m/z* = 224.08993 (left), corresponding to the $[M+Na]^+$ ion for acetyl-AHCCA from UHPLC-ESI-Q-TOF-MS analyses of reactions containing the following. The acetylated Bamb_5919 ACP domain, the Bamb_5919 KS^0 domain, *holo*-Bamb_5917, Bamb_5915 and AHCCA (top); the acetylated Bamb_5919 ACP-KS^0 di-domain in place of the acetylated Bamb_5919 ACP domain and the Bamb_5919 KS^0 domain (second from top); lacking the Bamb_5919 KS^0 domain (third from top); with the C1988A mutant of the Bamb_5919 KS^0 domain (third from bottom) and lacking Bamb_5915 (second from bottom). The bottom chromatogram is for the authentic standard of *N*-acetyl-AHCCA. The levels of acetyl-AHCCA in reactions lacking the functional KS^0 domain were 84.5 ± 1% lower than in reactions containing it.



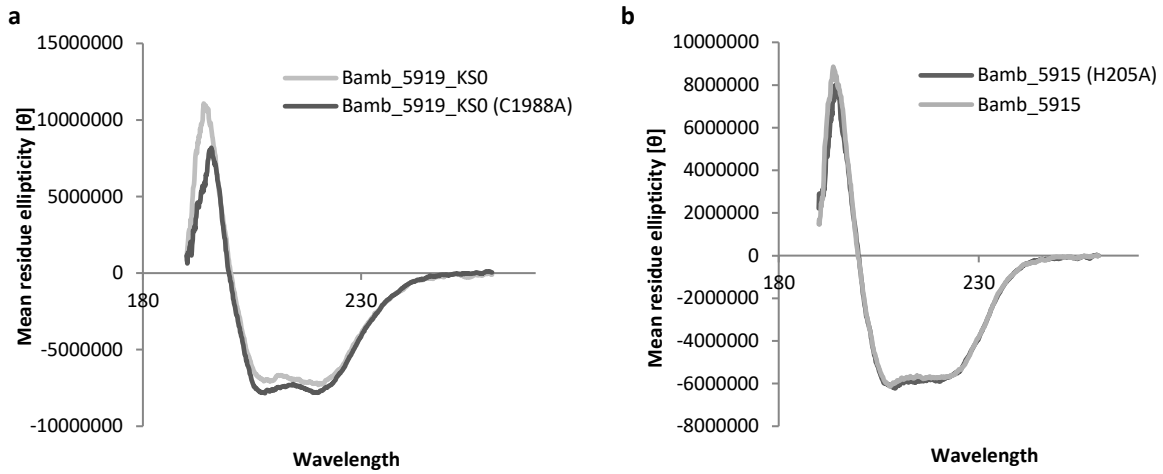
Supplementary Fig. 12: *In vitro* reconstitution of the enacyloxin chain release reaction using the 2, 4-hexadienoylated Bamb_5919 ACP domain. Experimental setup (right) and extracted ion chromatograms at *m/z* = 276.1212 ± 0.002 (left) corresponding to $[M+Na]^+$ for 2,4-hexadienoyl-AHCCA from UHPLC-ESI-Q-TOF-MS analyses of reaction mixtures containing the following. The 2,4-hexadienoylated Bamb_5919 ACP domain, the Bamb_5919 KS^0 domain, *holo*-Bamb_5917, Bamb_5915 and AHCCA (top); lacking Bamb_5915 (second from top; lacking the Bamb_5919 KS^0 domain (second from bottom); and with the C1988A mutant of the Bamb_5919 KS^0 domain (bottom).



Supplementary Fig. 13: *In vitro* reconstitution of the enacyloxin chain release mechanism using the dodecanoylated Bamb_5919 ACP domain. Experimental setup (right) and extracted ion chromatograms at *m/z* = 364.2464 ± 0.002 (left) corresponding to $[M+Na]^+$ for dodecanoyl-AHCCA from UHPLC-ESI-Q-TOF-MS analyses of the following reaction mixtures. The dodecanoylated Bamb_5919 ACP domain, the Bamb_5919 KS^0 domain, *holo*-Bamb_5917, Bamb_5915 and AHCCA (top); lacking Bamb_5915 (second from top); lacking the Bamb_5919 KS^0 domain (second from bottom); and with the C1988A mutant of the Bamb_5919 KS^0 domain (bottom).



Supplementary Fig. 14: Acyl donor tolerance of Bamb_5915. (**a**) Bamb_5915 was incubated with AHCCA and various NAC thioesters, and formation of an acylated product was determined using UHPLC-ESI-Q-TOF-MS (chromatograms labelled +C). Bamb_5915 was omitted from negative control reactions (chromatograms labelled –C). (**b**) Extracted ion chromatograms (EICs) at *m/z* = 224.0893 ± 0.005, corresponding to $[M+Na]^+$ for *N*-acetyl-AHCCA. (**c**) EICs at *m/z* =238.1050 ± 0.005, corresponding to $[M+Na]^+$ for propanoyl-AHCCA. (**d**) EICs at *m/z* = 252.1206 ± 0.005, corresponding to $[M+Na]^+$ for isobutyryl-AHCCA. (**e**) EICs at *m/z* = 266.1368 ± 0.005, corresponding to $[M+Na]^+$ for pivaloyl-AHCCA. (**f**) EICs at *m/z* = 266.1362 ± 0.005, corresponding to $[M+Na]^+$ for 2,4-hexadienoyl-AHCCA. (**g**) EICs at *m/z* = 269.1108 ± 0.005, corresponding to $[M+Na]^+$ for serinyl-AHCCA.



Supplementary Fig. 15: Secondary structure comparison of wild type and mutant proteins. Overlay of CD spectra measured for (**a**) the wild type Bamb_5919 KS^0 domain and the C1988A mutant and (**b**) Bamb_5915 and the H205A mutant.

Supplementary Table 1: Primers and templates used for PCR and site directed mutagenesis. The restriction sites and CACC sequences introduced at the 5' end of primers to allow for directional cloning are underlined and double underlined, respectively. Artificial stop codons introduced at the 5' end of reverse primers are shown in bold. Fw: forward primer, Rev: reverse primer.

Gene (region)	Size (bp)	Primer sequences (5'-3')	Construct Name	Template
Gene deletion				
bamb_5915 5'-flanking region	912	Fw: TCTAGAGAGATGAACGGCCGGATCAG Rev: AAGCTTATGAAAGCCATACTCAGCAC	pGPI/Δ*bamb_5915*	gDNA
bamb_5915 3'-flanking region	916	Fw: AAGCTTCCGCGCGCTCACAGGCTTGC Rev: GGTACCGCTACGAGGCCTGCGCCGAG	pGPI/Δ*bamb_5915*	gDNA
bamb_5917 5'-flanking region	580	Fw: CACTTAACGGCTGACATGGGAATTCTCG CCGCTCTCGAACC Rev: GGATCACGCCGGTAGCGAGGGTCATG GC	pGPI/Δ*bamb_5917*	gDNA
bamb_5917 3'-flanking region	587	Fw: CCTCGCTACCGGCGTGATCCGCTGAGGC Rev: AGCTTCCCGGGAAGATCTGGCTAGCGG GATTGCCGATCAGGTCC	pGPI/Δ*bamb_5917*	gDNA
Genetic complementation				
bamb_5915	1539	Fw: TTGGGCTAGCAGGAGGAATTCATGACG ATTCCCGCCTTG Rev: TCCGCCAAAACAGCCTCTAGATCATAAA ACCTCCGTGGTG	pMLBAD/ *bamb_5915*	gDNA
bamb_5917	948	Fw: GCGAATTCATGACCCTCGCTACCCTGCA AGCC Rev: GCTCTAGATCAGCGGATCACGCCTTCTT CGTACTC	pMLBAD/ *bamb_5917*	gDNA
Protein overproduction				
bamb_5915	1543	Fw: CACCATGACGATTCCCGCCTTG Rev: TCATAAAACCTCCGTGGT	pET151/*bamb_5915*	gDNA
bamb_5917 ACP	310	Fw: CACCGGCGCCGCCGCGGGCGTC Rev: TCAGCGGATCACGCCTTC	pET151/*bamb_5917 _ACP*	gDNA
bamb_5917	952	Fw: CACCATGGTGAGCGCACCGCGC Rev: TCAGCGGATCACGCCTTC	pET151/*bamb_5917*	3C12 fosmid
bamb_5919 ACP	568	Fw: CACCCTGGCCGAGCTGGTCGAG Rev: **TCA**CGCGAACGTGGCGCGCGA	pET151/*bamb_5919 _ACP*	3C12 fosmid
bamb_5919 KS^0	2017	Fw: CACCCCGTCGCGCGCCACGTTC Rev: **TCA**GGCCCGCCCATCGAC	pET151/*bamb_5919 _KS^0*	3C12 fosmid
bamb_5919 ACP-KS^0	2652	Fw: CACCCTGGCCGAGCTGGTCGAG Rev: TCAGGCCCGCCCATCGAC	pET151/*bamb_5919 _ACP-KS^0*	3C12 fosmid
Site directed mutagenesis				
bamb_5915 (H205A)	7299	Fw: CGCGTCGACGATGATGGCATGGAACACG CACAGC Rev: GCTGTGCGTGTTCCATGCCATCATCGTCG ACGCG	pET151/*bamb_5915_ H205A*	pET151/ *bamb_5915*
bamb_5919 KS^0 (C1988A)	8408	Fw: GCGCTCGACACAATGGCCTCGTCG TCGCTGAC Rev: GTCAGCGACGACGAGGCCATTGT GTCGAGCGC	pET151/ *bamb_5919_KS^0_ C1988A*	pET151/ *bamb_5919 _KS^0*

Supplementary Table 2: His_6-tagged proteins produced in this study, along with their molecular weights and extinction coefficients, the *E. coli* strains used to overproduce them, the molecular weight cut off (MWCO) of the filtration membranes used for protein concentration and buffer exchange, and the average yield of each protein.

Protein Name	Calculated molecular weight of His_6-tagged proteins (Da)	Extinction coefficient (M^{-1} cm^{-1})	*E. coli* strain used for protein overproduction	Amicon Ultra filtration membrane (MWCO)	Average yield (mg/L)
Bamb_5915	59692	73910	C43 (DE3)	30000	6
Bamb_5915 (H205A)	59625	73910	C43 (DE3)	30000	6
Bamb_5917 PCP*	14466	9970	BL21 Star (DE3)	10000	20
Bamb_5917	36929	40450	BL21 Star (DE3)	30000	35
Bamb_5919 ACP*	21662	1490	BL21 Star (DE3)	10000	15
Bamb_5919 KS^0	74956	77810	BL21 Star (DE3)	30000	12
Bamb_5919 KS^0 (C1988A)	74924	77810	BL21 Star (DE3)	30000	12
Bamb_5919 ACP-KS^0	92086	78185	BL21 Star (DE3)	30000	10
Sfp	28133	30370	BL21 Star (DE3)	10000	10

*The Bamb_5917 PCP domain boundaries were identified using BlastP analyses. The Bamb_5917 PCP domain sequence was then aligned with the sequence of the Bamb_5919 ACP domain to identify the boundaries of the Bamb_5919 ACP domain:

```
5919              EARAEATIGLIPPEQGAEVIARQFAHRDGDFALIPMRLAALAGQDRMPWLRALLAELVEA 1620
5917_PCP          ------------------------------------------------------------GA 2
                                                                             *
5919              EAGATGASGAPRVERRAGGTAGAALLAGLASLDAAARAARLKRHLEAAIRKLLNRADTLD 1680
5917_PCP          AAGVSAAGIEPDLT---------AIWQALFALPAVGR----------------------- 30
                   **.:.*.  * :           *:  .* :* *..*
5919              DRASMFDLGLDSLLSIDLRMQLEKDLACSLSTTVLHDHPTIEALAGFLAERVGAPPAGTV 1740
5917_PCP          -HQDFFALGGDSQLGLRMLAQLRERHGVDLPLRCLYEAPTVARLA--------------- 74
                   : .:* ** ** *.: :  **.:  . .*.   *:: **:  **
5919              RAGAAGGAGAGTGAPAGATGAAAAHAVSSASPVPAGAASAAASAASAAAAAGAPSRATFA 1800
5917-PCP          ----------------------ETIVRLAAPAPSGDQDDASEYEEGVIR---------- 101
                                        *  *:*.*:*  . *:.  ...
```

The fragment highlighted in blue was used for cloning into pET151.

References

1. Loveridge, E. J. *et al.* Reclassification of the specialized metabolite producer *Pseudomonas mesoacidophila* ATCC 31433 as a member of the *Burkholderia cepacia* complex. *J. Bacteriol.* **199,** e00125-17 (2017).
2. Flannagan, R. S., Linn, T. & Valvano, M. A. A system for the construction of targeted unmarked gene deletions in the genus *Burkholderia*. *Environ. Microbiol.* **10,** 1652–1660 (2008).
3. Agnoli, K. *et al.* Exposing the third chromosome of *Burkholderia cepacia* complex strains as a virulence plasmid. *Mol. Microbiol.* **83,** 362–378 (2012).
4. O'Sullivan, L. A. *et al.* Identifying the genetic basis of ecologically and biotechnologically useful functions of the bacterium *Burkholderia vietnamiensis*. *Environ. Microbiol.* **9,** 1017–1034 (2007).
5. Lefebre, M. D. & Valvano, M. A. Construction and evaluation of plasmid vectors optimized for constitutive and regulated gene expression in *Burkholderia cepacia* complex isolates. *Appl. Environ. Microbiol.* **68,** 5956–64 (2002).
6. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10,** 845–858 (2015).
7. Jenner, M., Jian, X., Dashti, Y., Masschelein, J., Hobson, C., Roberts, D., Jones, C., Harris, S., Parkhill, J., Raja, H., Oberlies, N., Pearce, C., Mahenthiralingam E. & Challis G.L. An unusual Burkholderia gladioli double chain-initiating nonribosomal peptide synthetase assembles 'fungal' icosalide antibiotics. *Chem. Sci.* **10**, DOI: 10.1039/C8SC04897E (2019).
8. Tosin, M., Spiteller, D. & Spencer, J.B. Malonyl carba(dethia)- and Malonyl oxa(dethia)-coenzyme A as Tools for Trapping Polyketide Intermediates. *ChemBioChem* **10**, 1714-1723 (2009).
9. Sekiyama, Y., Palaniappan, N., Reynolds, K. A. & Osada, H. Biosynthesis of phoslactomycins: cyclohexanecarboxylic acid as the starter unit. *Tetrahedron* **59,** 7465–7471 (2003).
10. Wang, X., Ma, M., Reddy, A. G. K. & Hu, W. An efficient stereoselective synthesis of six stereoisomers of 3, 4-diaminocyclohexane carboxamide as key intermediates for the synthesis of factor Xa inhibitors. *Tetrahedron* **73,** 1381–1388 (2017).
11. X. Wang, M. Ma, A. G. K. Reddy and W. Hu, *Tetrahedron*, **73**, 1381–1388 (2017).
12. Nagata, T. *et al.* Stereoselective synthesis and biological evaluation of 3,4-diaminocyclohexanecarboxylic acid derivatives as factor Xa inhibitors. *Bioorg. Med. Chem. Lett.* **18,** (2008).
13. Prasad, G., Borketey, L. S., Lin, T.-Y. & Schnarr, N. A. A mechanism-based fluorescence transfer assay for examining ketosynthase selectivity. *Org. Biomol. Chem.* **10,** 6717 (2012).
14. Mir Mohseni, M. *et al.* Discovery of a mosaic-like biosynthetic assembly line with a decarboxylative off-loading mechanism through a combination of genome mining and imaging. *Angew. Chemie Int. Ed.* **55,** 13611–13614 (2016).
15. Powell, A. *et al.* Engineered biosynthesis of nonribosomal lipopeptides with modified fatty acid side chains. *J. Am. Chem. Soc.* **129,** 15182–15191 (2007).
16. Roberts, D. M. *et al.* Substrate selectivity of an isolated enoyl reductase catalytic domain from an iterative highly reducing fungal polyketide synthase reveals key components of programming. *Chem. Sci.* **8,** 1116–1126 (2017).