Supplementary Information

Discovery and biosynthesis of gladiolin: a *Burkholderia gladioli* antibiotic with promising activity against *Mycobacterium tuberculosis*

Lijiang Song, Matthew Jenner, Joleen Masschelein, Cerith Jones, Matthew J. Bull, Simon R. Harris, Ruben C. Hartkoorn, Anthony Vocat, Isolda Romero-Canelon, Paul Coupland, Gordon Webster, Matthew Dunn, Rebecca Weiser, Christopher Paisey, Stewart T. Cole, Julian Parkhill, Eshwar Mahenthiralingam* and Gregory L. Challis*

Methods

Antimicrobial activity screen

Antimicrobial activity assays with *B. gladioli* BCC0238 were performed as described previously (17) using the following indicator strains: *S. aureus* DSM21979, *R. mannitolilytica* LMG6866T, *C. albicans* SC5314, *E. faecium* DSM25390, *B. subtilis* DSM10 and *B. multivorans* ATCC 17616. MIC values were determined using the broth microdilution method according to the official CLSI guidelines.

Gladiolin isolation, structure elucidation and stability analysis

Gladiolin-containing extracts from agar cultures were fractionated by semi-preparative HPLC which was carried out on an Agilent 1100 instrument equipped with a Zorbax SB-C18 column (100 x 21.2 mm, 5 μm), monitoring absorbance at 240 nm. Mobile phases consisted of water (A) and acetonitrile (B), each supplemented with 0.1% formic acid. A gradient of 20% B to 100% B over 20 minutes was employed at a flow rate of 5mL/min. Structure elucidation of gladiolin was achieved using a combination of UHPLC-ESI-Q-TOF-MS/MS and 1- and 2-D NMR experiments.

UHPLC-ESI-Q-TOF-MS/MS analyses were performed using a Dionex UltiMate 3000 UHPLC connected to a Zorbax Eclipse Plus C-18 column (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. A gradient of 20% B to 100% B over 30 minutes was employed at 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.

For NMR analysis, purified gladiolin was dissolved in d_4-MeOH or d_6-DMSO, and ^{1}H, ^{13}C, COSY, NOESY, TOCSY, HSQC and HMBC spectra were recorded on a Bruker Avance 700MHz spectrometer equipped with a TCI cryoprobe at 25 °C.

To assess the stability of gladiolin, it was dissolved in MeOH (1mg/ml) and left at 23-26 °C for 0, 2, 8, 24, 67, 72 and 96 h, prior to UHPLC-ESI-Q-TOF-MS analysis. After 96 hours, the majority of the material had converted to *iso*-gladiolin, the structure of which was determined using ^{1}H, HSQC and HMBC NMR experiments (**Figs. S10-11 and Table S3**).

Inhibition of the mycobacterial RNA polymerase *in vitro*

A plasmid template for real-time *in vitro* transcription was generated from a minimal pUC19-based reporter plasmid where, between transcriptional terminators, a strong mycobacterial promoter was cloned upstream of a poly(A) stretch of DNA, to allow for efficient promoter-specific transcription of poly(A) RNA by mycobacterial

RNA polymerase (RNAP) holoenzyme. Initially, a stretch of 25 adenines (poly(A) followed by a 3' *rrnABT1* transcriptional terminator (T1) was inserted into the *Sac*I site of pUC19 (as described in (26)) to give pUC-poly(A)T1. pUC-poly(A)T1 was then reduced in size by excising the DNA between two *Nsp*I sites (365 bp) and two *Aat*II sites (514 bp), maintaining the pUC19 multi-cloning site (mcs) upstream of the poly(A)T1 sequence. A 5' transcriptional terminator sequence was then inserted into the *Nsp*I site using self-annealed phosphorylated primers (CCATCCTGACGGATGGCCTTCATG and AAGGCCATCCGTCAGGATGGCATG) to generate pIVT(mcs). Finally, the promoter region of the *M. tuberculosis rrs* gene was amplified from H37Rv genomic DNA (primers: GGATATCCGTTGTTCGTGGA and TTCTCAAACAACACGCTTGC) and cloned in the correct orientation into the *Sma*I restriction site of pIVT(mcs) to give pIVT(Prrs). *M. smegmatis* RNAP holoenzyme was purified from *M. smegmatis* mc^2155 expressing a His_6-tagged RNAP β subunit (kindly provided by Dipankar Chatterjee, Indian Institute of Science) as described previously (26, 37).

To measure inhibition of promoter-specific DNA-dependent RNA transcription by gladiolin, a real-time, multiple round *in vitro* transcription assay was used that employs molecular beacons (25, 26). Briefly, 0.2 µL of 100X compound in DMSO, or DMSO alone, were added to a MicroAmp optical 384-well reaction plate, followed by addition of 17.8 µL of reaction mixture (40 mM Tris-HCl, pH 7.9, 75 mM KCl, 10 mM $MgCl_2$, 10 mM dithiothreitol (DTT), 0.1 mM EDTA, 10 U RNasin (Promega), 500 nM molecular beacon, 1 µg of pIVT(Prrs) plasmid template and 1 µg of 2.4 pmol of *M. smegmatis* RNAP holoenzyme), and allowed to pre-incubate for 10 min at room temperature. *In vitro* transcription was then initiated by adding 2 µL of ribonucleotide triphosphate (rNTP) mix (final concentration of 150 µM each rNTP). *In vitro* transcription was subsequently followed by measuring fluorescence (ex: 495 nm, em:515 nm, every 30 sec for 40 min at 25°C) using an ABI Prism 7900HT sequence detection system (Applied Biosystems). Fluorescent readouts were background-corrected (fluorescence at t=0) and the rate of linear *in vitro* transcription was determined as the change in fluorescence over time. As a control for RNAP inhibition, rifampicin was used.

Activity of gladiolin against human ovarian cancer cells

A2780 human ovarian cancer cells were obtained from the European Collection of Animal Cell Culture and grown as single monolayers in Roswell Park Memorial Institute medium (RPMI-1640) supplemented with 10% (v/v) fetal calf serum, 1% (v/v) 2 mM L-glutamine and 1% (v/v) penicillin (10 k units/mL)/streptomycin (10 mg/mL). Cells were kept at 310 K in a humidified atmosphere containing 5% CO_2 and maintenance passages were done at ca. 80% confluency.

96-well plates were used to seed 5000 A2780 cells per well. These were left to pre-incubate in drug-free media at 310 K for 48 h before adding various concentrations of gladiolin. A stock solution of gladiolin was prepared in 5% v/v DMSO and 95% v/v cell culture medium, then serial dilutions with culture medium were carried out to achieve working concentrations. The cells were exposed to various concentrations of gladiolin for a period of 24 h, the culture supernatants were removed by suction and each well was washed with PBS. A further 72 h were allowed for the cells to recover in drug-free medium at 310 K. A modified version of the SRB assay was used to determine cell viability (Skehan, P., *et al. J. Natl. Cancer Inst.* **1990**, *82*, 1107–1112; Vichai, V.; Kirtikara, K. *Nat. Protoc.* **2006**, *1*, 1112–1116.). In this assay, sulforhodamine B binds to basic amino acid residues of

proteins in fixed cells. The percentage of viable cells resulting from exposure to gladioilin is determined by measuring the absorbance due to soluble sulforohodamine relative to an untreated control. The absorbance measurements were carried out using a BioRad iMark microplate reader with a 470 nm filter. Mean percentage cell viability values +/- 1 standard deviation were calculated from duplicates of triplicates in two independent sets of experiments (**Fig. S14**).

Evaluation of gladiolin toxicity using a *Galleria mellonella* wax moth larvae model

The toxicity of gladiolin relative to rifampicin (Sigma-Aldrich) using evaluated using the *Galleria mellonella* wax moth larvae model (25). TruLarv™ research grade *G. mellonella* larvae were obtained from BioSystems Technology (Exeter, UK). Antibiotics were prepared in 10% DMSO and larvae were injected with 10 µl of either gladiolin or rifampicin at 0.5 µg ml^{-1} or gladiolin at 10 µg ml^{-1}. The doses administered (5 µg and 100 µg in larvae weighing between 0.18 - 0.35 g) are equivalent to 14.3 mg kg^{-1} and 555 mg kg^{-1}, both of which exceed the Centre for Disease Control suggested human adult daily dose of rifampicin for the treatment of active tuberculosis (10 mg kg^{-1}). Ten larvae were injected per experiment, each of which was performed in triplicate. Control larvae were injected with 10 % DMSO (n=10, in triplicate), 100 % DMSO (n = 10) and 10 larvae were left untreated. Larvae were incubated at 37 °C and the percentage surviving after 24, 48 and 72 hours was recorded (**Fig. S15**).

Structure elucidation of gladiolin by mass spectrometry and NMR spectroscopy

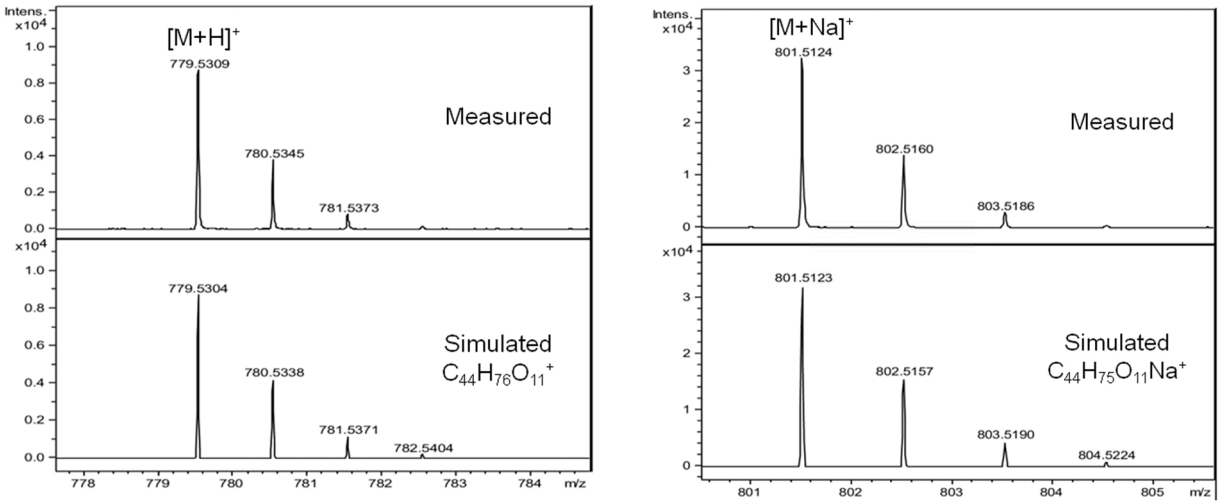

Figure S1| High-resolution mass spectrometry analysis of gladiolin. Measured spectra of gladiolin ($[M+H]^+$ and $[M+Na]^+$, top left and top right panel, respectively) and simulated spectra for $C_{44}H_{76}O_{11}^+$ and $C_{44}H_{75}O_{11}Na^+$ (bottom left and bottom right panel, respectively).

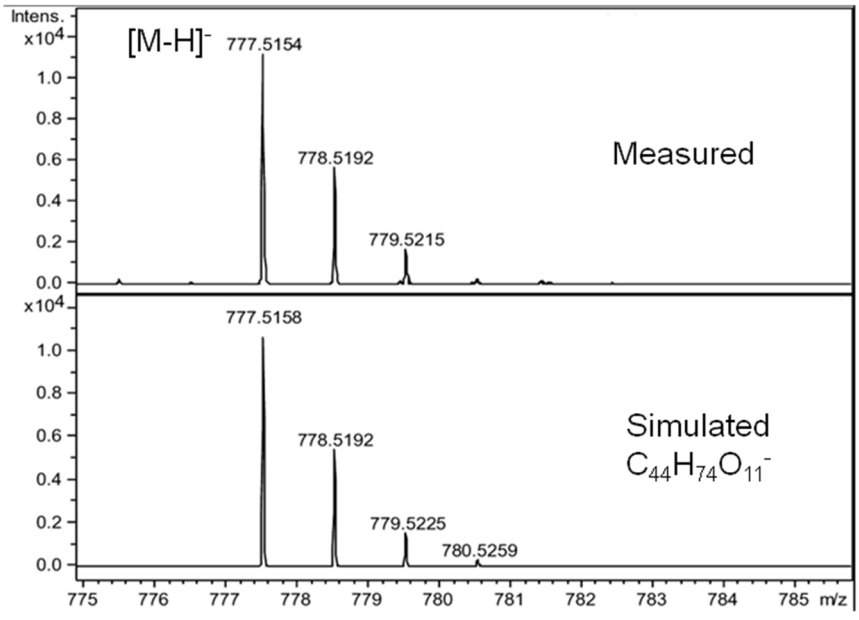

Figure S2| High-resolution mass spectrometry analysis of gladiolin in negative ion mode. Measured spectrum of gladiolin ($[M-H]^-$, top panel) and simulated spectrum for $C_{44}H_{74}O_{11}^-$ (bottom panel).

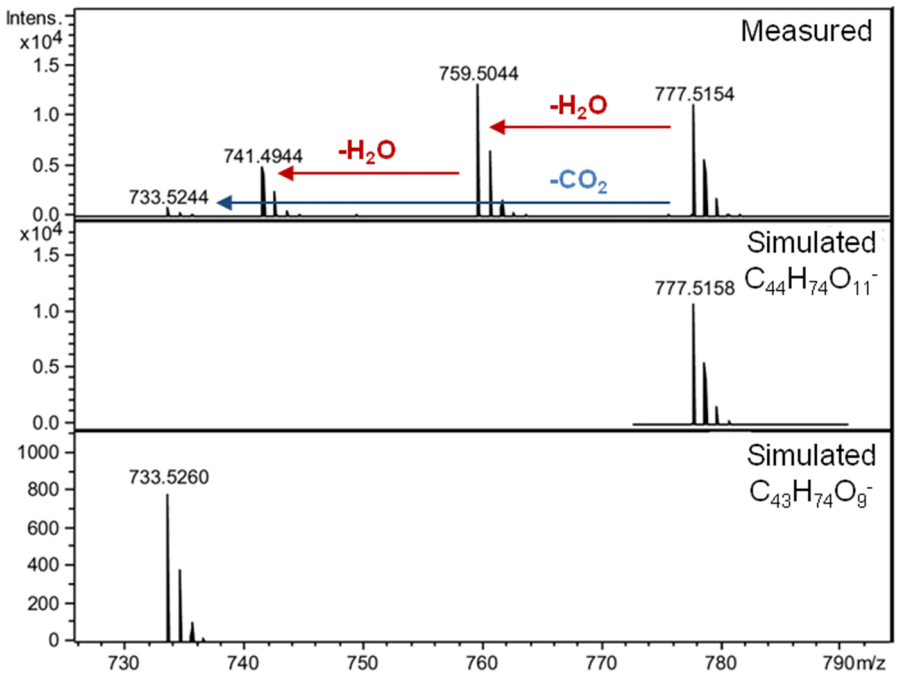

Figure S3| High-resolution MS/MS analysis of gladiolin in negative ion mode. Measured spectrum of gladiolin ([M-H]$^-$, top panel) and simulated spectrum for $C_{44}H_{74}O_{11}^-$ (middle panel) and $C_{43}H_{74}O_9^-$ (bottom panel). Fragment ions resulting from the neutral loss of H_2O and CO_2 molecules could be detected.

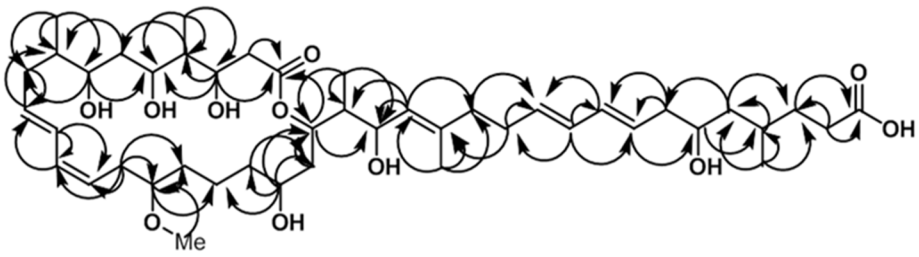

Figure S4| Correlations observed in HMBC spectra of gladiolin (d_4-MeOH, ^{1}H 700 MHz, ^{13}C 175 MHz)

Table S1| Overview of NMR signals observed for gladiolin isolated from *B. gladioli* BCC0238 (d_6-DMSO, ^{1}H 700 MHz, ^{13}C 175 MHz)

Position	^{1}H (ppm)	^{13}C (ppm)*
1		170.9
2	2.22, 2.40 (2 x 1H, 2 x m)	39.6
3	3.95(1H, m)	68.9
3-OH	4.05(ol)	
4	1.46 (1H, m)	43.8
4-Me	0.79 (3H, d, 6.5)	9.6
5	3.71 (1H, m)	68.8
5-OH	4.02 (ol)	
6	1.46 (2H, m)	39.6
7	3.54 (1H, m)	70.6
7-OH	4.02 (1H, br s)	
8	1.62 (1H, m)	37.0
8-Me	0.78 (3H, d, 6.0)	13.3
9	1.92, 2.18 (2 x 1H, 2 x m)	35.8
10	5.71 (1H, dt, 6.5, 15.0)	133.0
11	6.29 (1H, dd, 11.0, 15.0)	126.1
12	6.02 (1H, 11.0, 11.0)	129.7
13	5.31 (1H, dt, 7.0, 11.0)	125.0
14	2.26, 2.38 (2x 1H, 2 x m)	30.0
15	3.18 (1H, dd, 5.5, 5.5)	79.6
15-OMe	3.23(3H, s)	55.5
16	1.32, 1.48 (2 x 1H, 2 x m)	32.2
17	1.23, 1.36 (2 x 1H, 2 x m)	21.1
18	1.32, 1.30 (2 x 1H, 2x m)	36.2
19	3.38 (1H, m)	66.1
19-OH	3.54(1H, ol)	
20	1.49, 1.61 (2 x 1H, 2 x m)	36.3
21	5.30 (1H, m)	71.8
22	1.80 (1H, m)	42.1
22-Me	0.69 (3H, d, 6.5)	10.4
23	3.98 (1H, m)	67.8
23-OH	4.46 (1H, br s)	
24	5.06 (1H, d, 8.5)	127.6
25		134.7
25-Me	1.59 (3H, s)	16.2
26	2.02 (2H, m)	38.9
27	2.12 (2H, m)	30.1
28	5.54 (1H, dt, 7.0, 15.0)	130.2
29	5.98 (1H, dd, 11.0, 15.0)	131.1
30	5.96 (1H, dd, 11.0, 15.0)	131.9
31	5.56 (1H, dt, 7.0, 15.0)	129.2
32	2.08 (2H, m)	41.3
33	3.49 (1H, m)	67.1
33-OH	3.80 (1H, br s)	
34	1.04, 1.28 (2 x 1H, 2 x m)	43.6
35	1.61 (1H, m)	28.1
35-Me	0.80 (3H, d, 6.5)	17.9
36	1.32, 1.47 (2 x 1H, 2 x m)	32.3
37	2.14, 2.19 (2 x 1H, 2 x m)	29.9
38		174.9

*From HMBC and HSQC spectra

Table S2| NMR assignments for gladiolin (d_4-MeOH, ^{1}H 700 MHz, ^{13}C 175 MHz).

Position	^{1}H (ppm)	^{13}C (ppm)*
1		173.4
2	2.43, 2.58 (2 x 1H, 2 x d, 4.0, 15.5)	41.2
3	4.10 (1H, m)	71.6
4	1.62 (1H, m)	44.6
4-Me	0.93 (3H,d, 7.0)	10.2
5	3.97 (1H, m)	71.4
6	1.66 (2H, m)	39.8
7	3.71 (1H, m)	73.5
8	1.73 (1H, m)	38.7
8-Me	0.92 (3H, d, 7.0)	14.2
9	2.08, 2.28 (2 x 1H, 2 x m)	37.2
10	5.75 (1H, dt, 7.0, 15.0)	133.4
11	6.38 (1H, dd, 11.5, 15.0)	128.0
12	6.08 (1H, 11.5, 11.0)	131.7
13	5.34 (1H, dt, 7.5, 11.0)	125.8
14	2.36, 2.47 (2 x 1H, 2 x m)	32.0
15	3.28 (1H, dd, 5.5, 5.5)	82.4
15-OMe	3.35 (3H, s)	56.7
16	1.47, 1.62 (2 x 1H, 2 x m)	33.9
17	1.39, 1.48 (2 x 1H, 2 x m)	22.2
18	1.47, 1.38 (2 x 1H, 2 x m)	38.7
19	3.60 (1H, m)	68.4
20	1.64, 1.70 (2 x 1H, 2 x m)	38.0
21	5.39 (1H, m)	73.6
22	1.98 (1H, m)	43.5
22-Me	0.81 (3H, d, 7.0)	10.7
23	4.20 (1H, t, 9.0)	69.9
24	5.16 (1H, d, 9.5)	127.3
25		138.9
25-Me	1.68 (3H, s)	16.7
26	2.12 (2H, t, 6.5)	40.3
27	2.20 (2H, m)	31.6
28	5.58 (1H, dt, 7.0, 14.0)	132.8
29	6.04 (1H, dd, 11.5, 14.0)	131.7
30	6.04 (1H, dd, 11.5, 14.0)	133.7
31	5.59 (1H, dt, 7.0, 14.0)	129.1
32	2.19 (2H, m)	42.4
33	3.66 (1H, m)	69.9
34	1.19, 1.43 (2 x 1H, 2 x m)	44.8
35	1.70 (1H, m)	29.9
35-Me	0.91 (3H, d, 6.5)	18.5
36	1.46, 1.62 (2 x 1H, 2 x m)	34.2
37	2.28, 2.32 (2 x 1H, 2 x m)	32.8
38		178.5
*From HMBC and HSQC spectra		

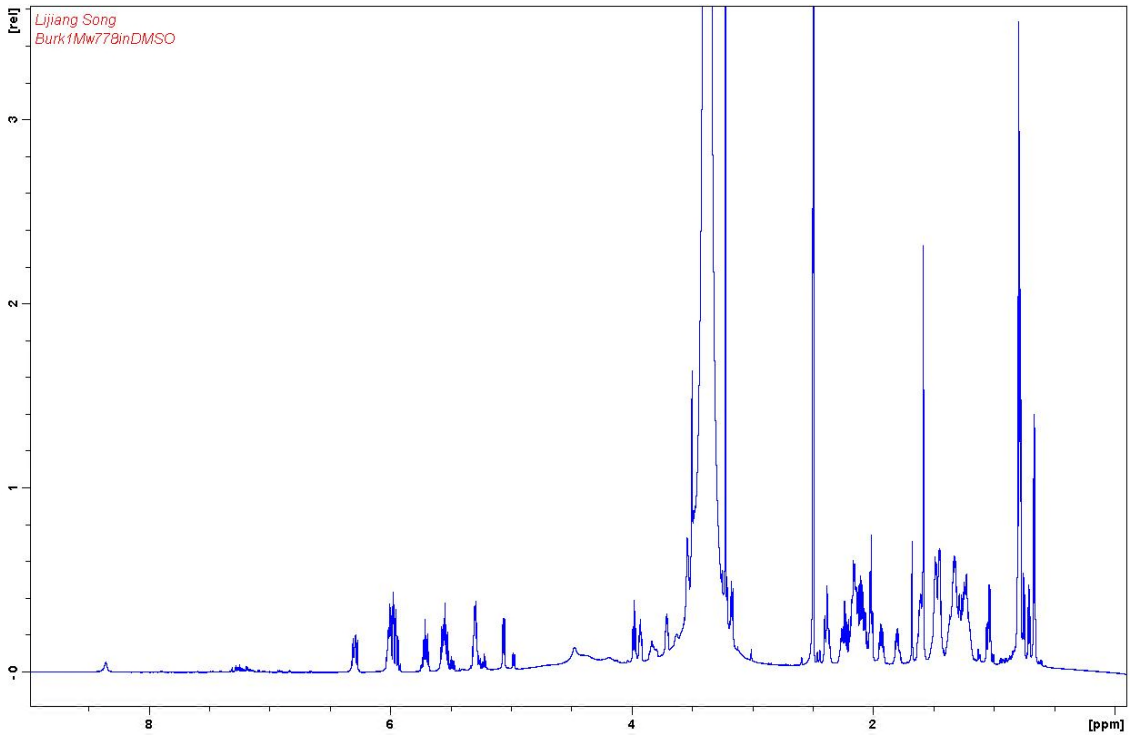

Figure S5| ^{1}H NMR spectrum of gladiolin in d_6-DMSO

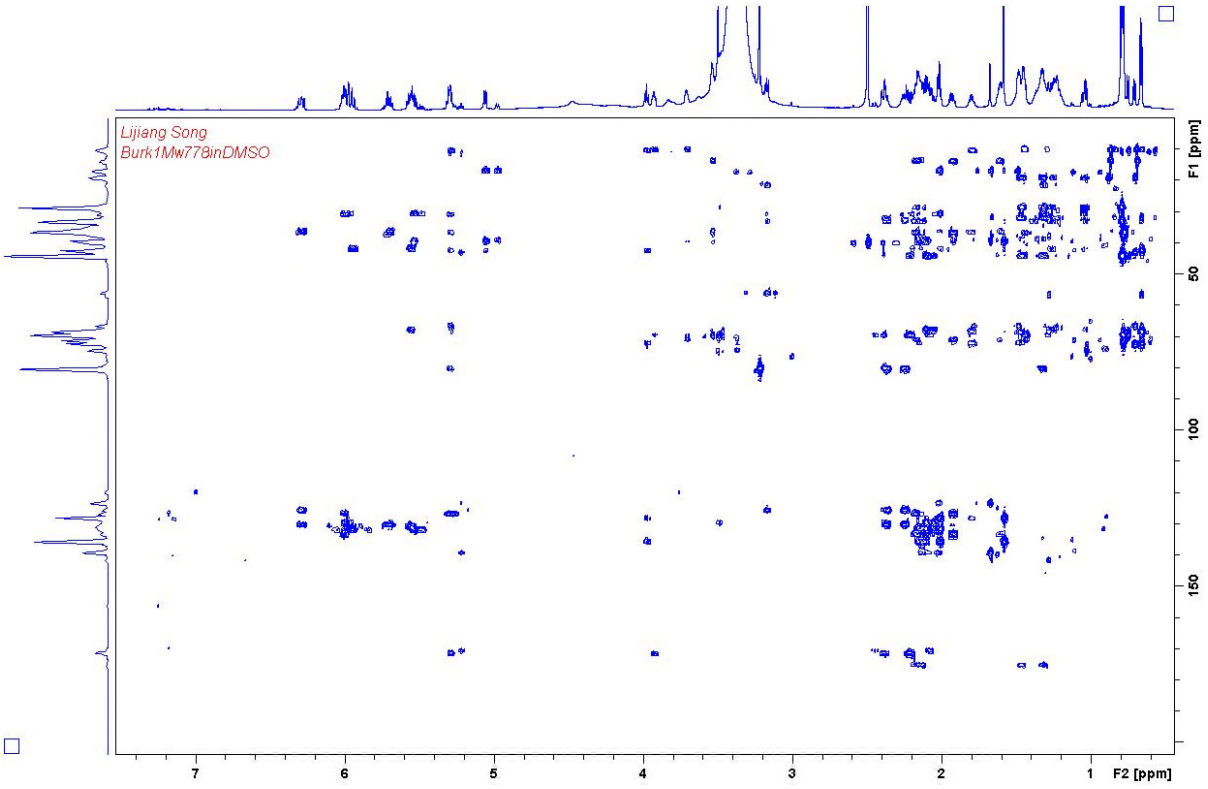

Figure S6| HMBC spectrum of gladiolin in d_6-DMSO

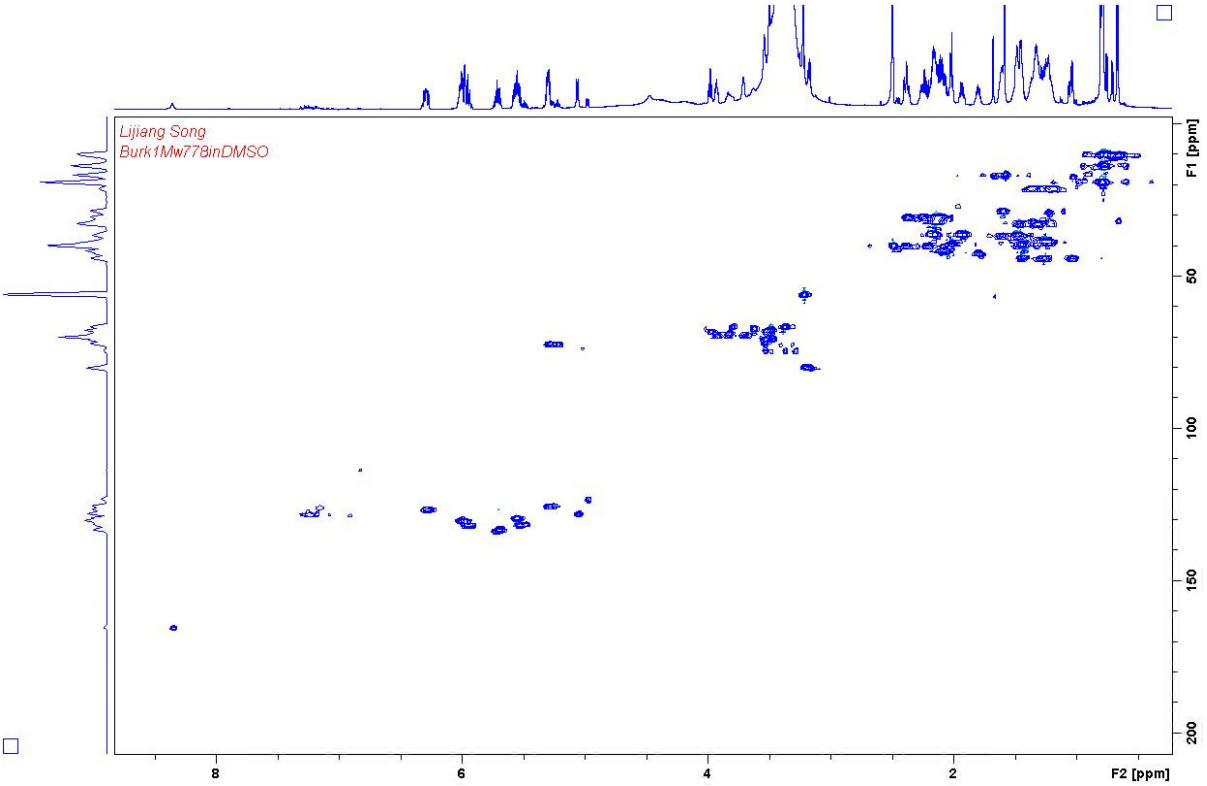

Figure S7| HSQC spectrum of gladiolin in d_6-DMSO

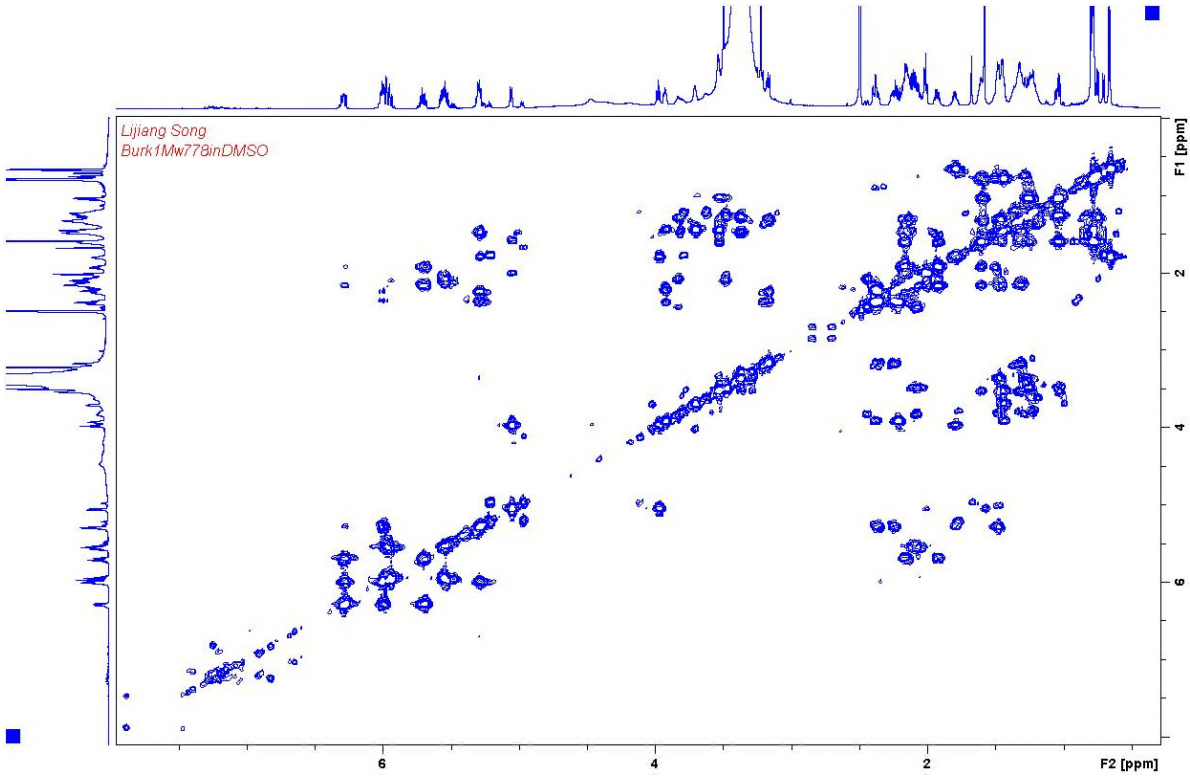

Figure S8| COSY spectrum of gladiolin in d_6-DMSO

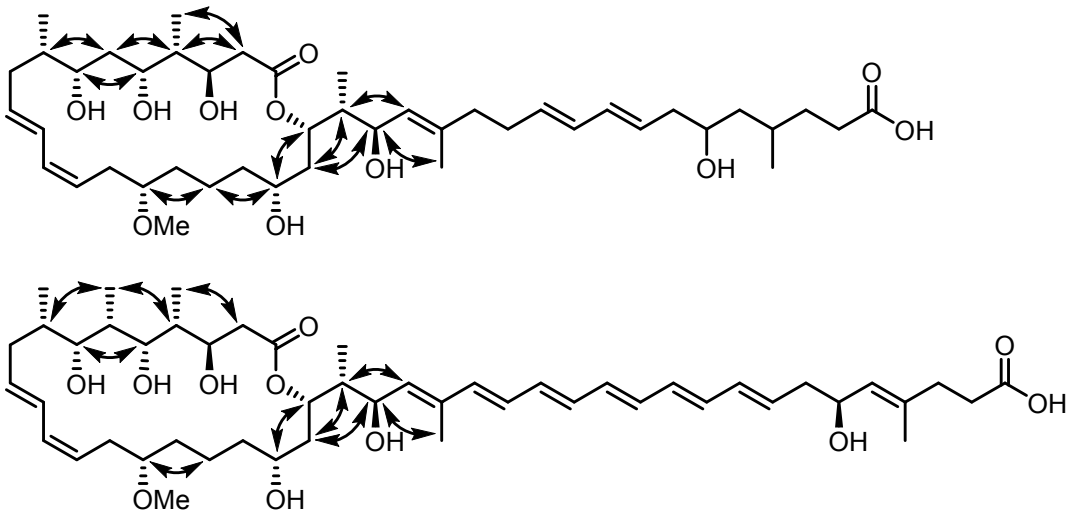

Figure S9| Comparison of key correlations in NOESY NMR spectra of gladiolin and etnangien used to assign relative stereochemistry

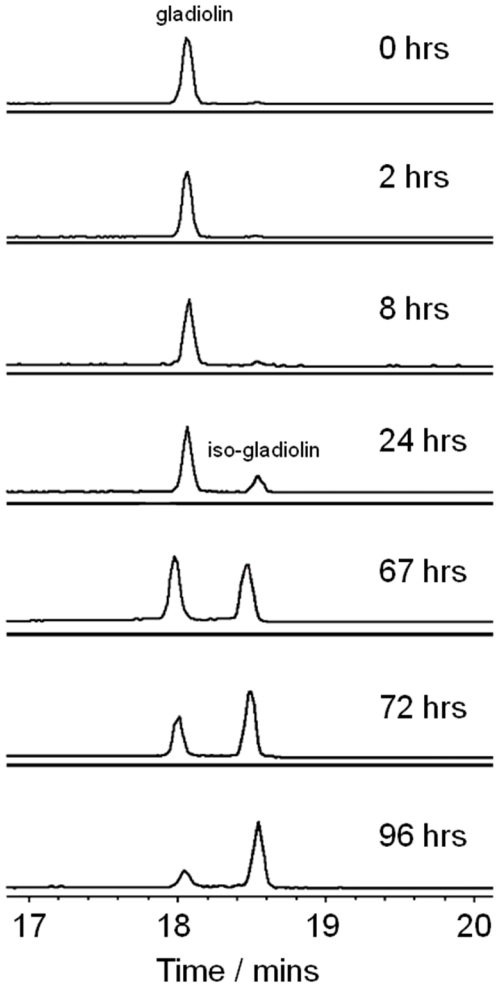

Figure S10| Time-course analysis of gladiolin stability
Chromatograms from UHPLC-ESI-Q-TOF-MS analysis, monitoring absorbance at 240 nm, of purified gladiolin left at room temperature in MeOH for 2-96 h, showing progressive conversion of gladiolin to iso-gladiolin. Iso-gladiolin has the same molecular formula as gladiolin, but a slightly longer retention time on a C_{18} column.

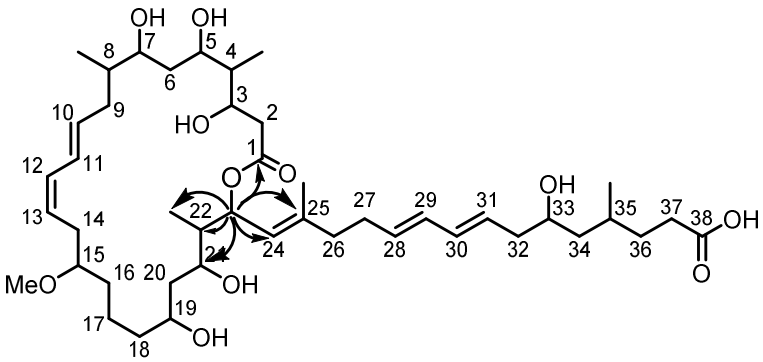

Figure S11| Key correlations observed in the HMBC spectrum of iso-gladiolin used to establish that the ester oxygen atom bridges C-1 and C-23

Table S3| Comparison of NMR signals observed for gladiolin and iso-gladiolin (d_4-MeOH, 700 MHz). Key differences are highlighted in red.

	Iso-Gladiolin		Gladiolin	
Position	**^{1}H (ppm)**	**^{13}C (ppm)***	**^{1}H (ppm)**	**^{13}C (ppm)***
1		172.7		173.4
2	2.24,2.63 (2 x 1H, 2 x d, 4.0, 15.5)	40.7	2.43, 2.58 (2 x 1H, 2 x d, 4.0, 15.5)	41.2
3	3.99 (1H, m)	70.8	4.10 (1H, m)	71.6
4	1.54 (1H, m)	45.2	1.62 (1H, m)	44.6
4-Me	0.88 (3H, d, 7.0)	10.2	0.93 (3H, d, 7.0)	10.2
5	4.05 (1H, m)	70.3	3.97 (1H, m)	71.4
6	1.53/1.70 (2H, m)	39.8	1.66 (2H, m)	39.8
7	3.75 (1H, m)	73.8	3.71 (1H, m)	73.5
8	1.71 (1H, m)	38.2	1.73 (1H, m)	38.7
8-Me	0.92 (3H, d, 7.0)	14.0	0.92 (3H, d, 7.0)	14.2
9	2.04,2.27(2 x 1H, 2 x m)	37.7	2.08, 2.28 (2 x 1H, 2 x m)	37.2
10	5.76 (1H, dt, 7.0, 14.5)	134.1	5.75 (1H, dt, 7.0, 15.0)	133.4
11	6.40 (1H, dd, 11.0, 14.5)	128.0	6.38 (1H, dd, 11.5, 15.0)	128.0
12	6.05 (1H, dd, 11.0, 11.0)	131.2	6.08 (1H, 11.5, 11.0)	131.7
13	5.30 (1H, m)	126.2	5.34 (1H, dt, 7.5, 11.0)	125.8
14	2.34, 2.45(2 x 1H, 2 x m)	32.0	2.36, 2.47 (2 x 1H, 2 x m)	32.0
15	3.28(1H, m)	82.2	3.28 (1H, dd, 5.5, 5.5)	82.4
15-OMe	3.35 (3H, s)	56.4	3.35 (3H, s)	56.7
16	1.47,1.62 (2x 1H, 2 x m)	33.9	1.47, 1.62 (2 x 1H, 2 x m)	33.9
17	1.37,1.50 (2x 1H, 2 x m)	21.9	1.39, 1.48 (2 x 1H, 2 x m)	22.2
18	1.38,1.44 (2x 1H, 2 x m)	39.0	1.47, 1.38 (2 x 1H, 2 x m)	38.7
19	3.76 (1H, m)	69.0	3.60 (1H, m)	68.4
20	1.54,1.70 (2 x 1H, 2 x m)	39.0	1.64, 1.70 (2 x 1H, 2 x m)	38.0
21	4.08 (1H, m)	68.5	5.39 (1H, m)	73.6
22	1.97 (1H, m)	44.2	1.98 (1H, m)	43.5
22-Me	0.83 (3H, d, 7.0)	11.1	0.81 (3H, d, 7.0)	10.7
23	5.29 (1H, t, 10.0)	74.5	4.20 (1H, t, 9.0)	69.9
24	5.04 (1H, d, 10.0)	124.0	5.16 (1H, d, 9.5)	127.3
25		141.6		138.9
25-Me	1.76 (3H, s)	16.9	1.68 (3H, s)	16.7
26	2.12 (2H, t, 8.0)	40.2	2.12 (2H, t, 6.5)	40.3
27	2.20 (2H, m)	31.6	2.20 (2H, m)	31.6
28	5.53 (1H, dt, 7.5, 14.0)	132.9	5.58 (1H, dt, 7.0, 14.0)	132.8
29	6.03 (1H, dd, 11.0, 14.0)	131.8	6.04 (1H, dd, 11.5, 14.0)	131.7
30	6.02 (1H, dd, 11.0, 14.0)	133.5	6.04 (1H, dd, 11.5, 14.0)	133.7
31	5.57 (1H, dt, 7.0, 14.0)	129.0	5.59 (1H, dt, 7.0, 14.0)	129.1
32	2.20(2H, m)	42.3	2.19 (2H, m)	42.4
33	3.67(1H, m)	69.6	3.66 (1H, m)	69.9
34	1.20,1.43 (2 x 1H, 2 x m)	45.1	1.19, 1.43 (2 x 1H, 2 x m)	44.8
35	1.70 (1H, m)	29.9	1.70 (1H, m)	29.9
35-Me	0.91(3H, d, 7.0)	19.1	0.91 (3H, d, 6.5)	18.5
36	1.47,1.61 (2 x 1H, 2 x m)	33.8	1.46, 1.62 (2 x 1H, 2 x m)	34.2
37	2.20,2.34 (2 x 1H, 2 x m)	31.6	2.28, 2.32 (2 x 1H, 2 x m)	32.8
38		178.6		178.5

*From HMBC and HSQC spectra

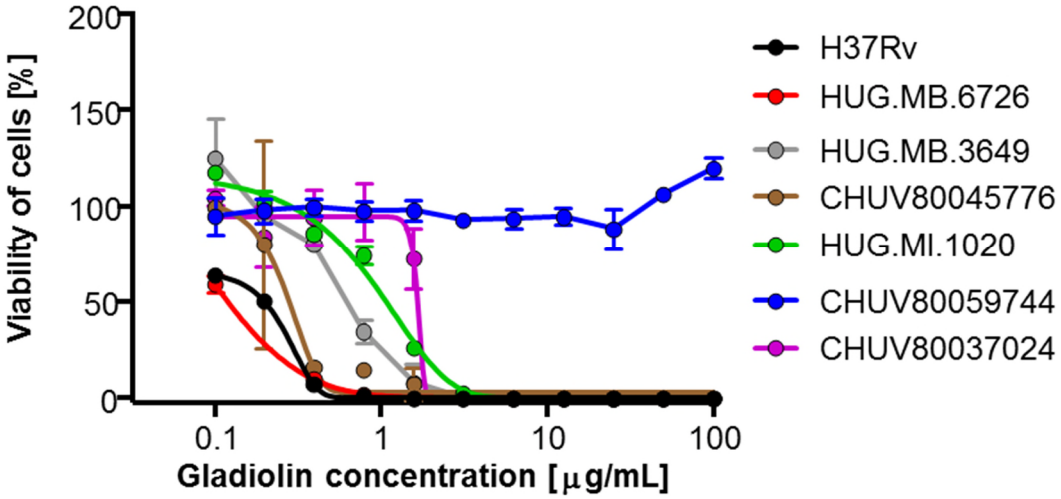

Figure S12| Activity of gladiolin against *M. tuberculosis* H37Rv and various antibiotic-resistant clinical isolates.

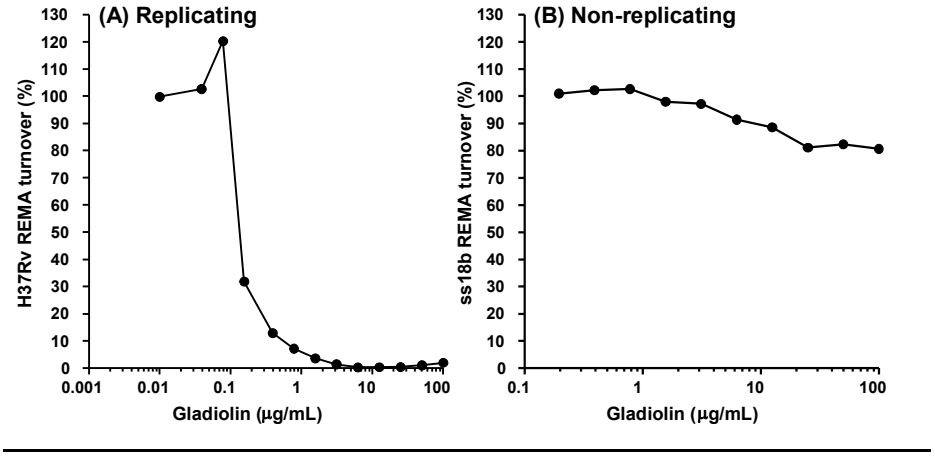

Figure S13| Comparison of the activity of gladiolin against replicating and non-replicating *M. tuberculosis*
Susceptibility testing of *M. tuberculosis* was performed using the resazurin reduction assay. Unlike replicating *M. tuberculosis* H37Rv (Panel A), non-replicating *M. tuberculosis* ss18b was found to be much less susceptible to inhibition by gladiolin (Panel B).

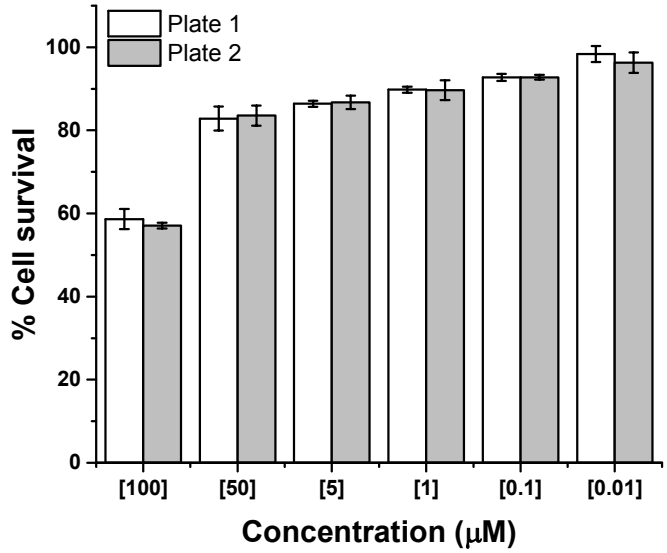

Figure S14| Percentage of A2780 ovarian cancer cells surviving after exposure to concentrations of gladiolin ranging from 0.01 to 100 μM relative to an untreated control. The gladiolin exposure and recovery times were 24 h and 72 h, respectively.

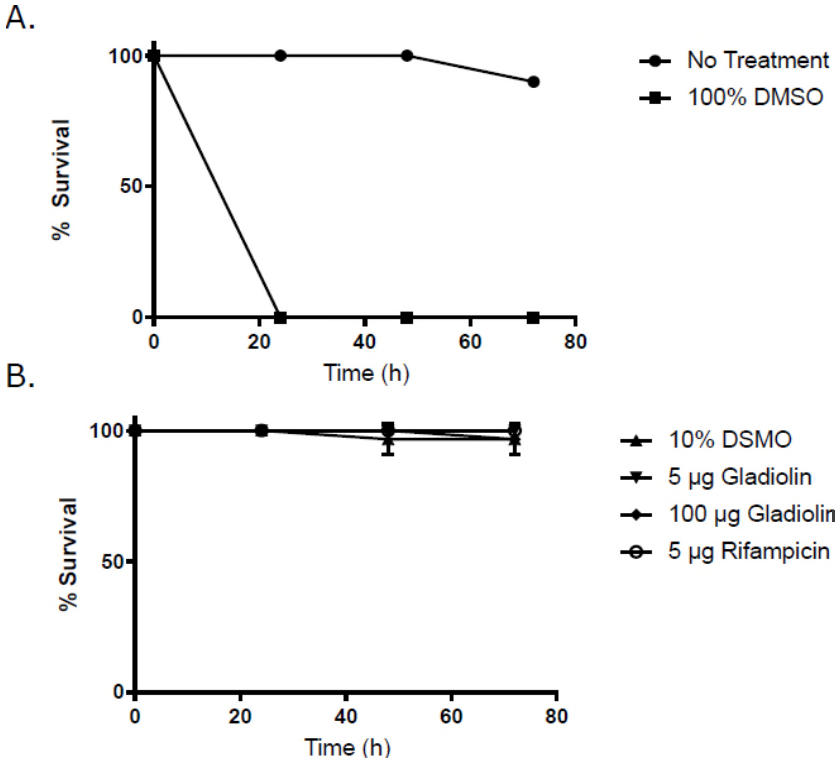

Figure S15| ***G. mellonella*** **survival following injection with gladiolin and rifampicin.** (A) Percentage survival of non-treated larvae (n=10) and 100 % DMSO (n =10) over a 72 hour period. (B) Percentage survival of larvae injected with solvent alone (10 % DMSO), 5 μg of rifampicin, 5 μg of gladiolin or 100 μg of gladiolin over a 72 hour period. Mean survival rates from three replicate experiments are shown (n = 10 per replicate). Error bars represent +/- 1 standard deviation from the mean.

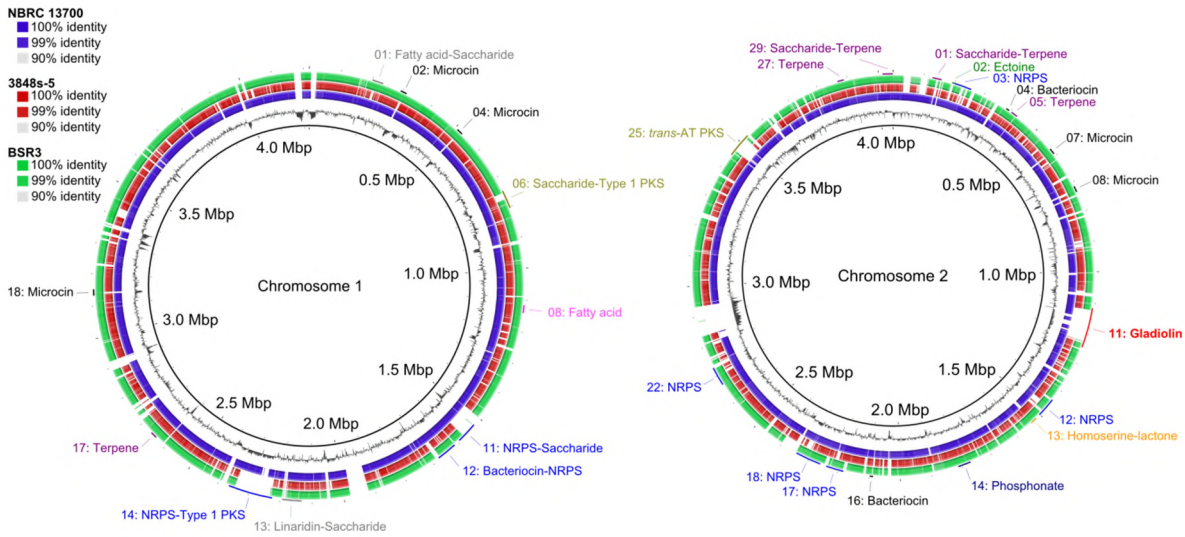

Figure S16| Genome comparisons and biosynthetic potential of *B. gladioli* BCC0238. The two largest chromosomal replicons of strain BCC0238 were compared to the complete genome sequence of *B. gladioli* BSR3 and the draft genomes of *B. gladioli* strains NBRC 13700 and 3848s-5 (percentage identity shown on the outer rings as indicated on the top left key). Positions (Mbp) and GC-content are plotted on the inner ring. The putative specialized metabolite biosynthetic gene clusters identified are shown in various colours according to the biosynthetic machinery they encode. The gladiolin biosynthetic gene cluster, located on chromosome 2, is indicated in red.

Table S4: Specialized metabolite biosynthetic gene clusters predicted to reside in the *B. gladioli* BBC0238 genome, the sequence of which has been deposited in Genbank.

Cluster number	antiSMASH functional prediction	Start position	End position	Size (bases)
Replicon 1				
R1-01	Fatty acid – Saccharide[a]	200105	235404	35300
R1-02	Microcin	294368	314462	20095
R1-03	Saccharide[a]	467828	488577	20750
R1-04	Microcin	508286	528380	20095
R1-05	Saccharide[a]	528999	551206	22208
R1-06	Saccharide[a] –Type 1 PKS	746515	803959	57445
R1-07	Saccharide[a]	844198	894237	50040
R1-08	Fatty acid[a]	1115704	1140289	24586
R1-09	Saccharide[a]	1157379	1193017	35639
R1-10	Saccharide[a]	1291051	1312982	21932
R1-11	NRPS –Saccharide[a]	1525907	1592996	67090
R1-12	Bacteriocin – NRPS	1613086	1677116	64031
R1-13	Linaridin – Saccharide[a]	2133079	2194088	61010
R1-14	NRPS – Type 1 PKS	2225763	2365560	139798
R1-15	Saccharide[a]	2546263	2567270	21008
R1-16	Saccharide[a]	2573240	2594352	21113
R1-17	Terpene	2639091	2659930	20840
R1-18	Microcin	3132596	3152690	20095
R1-19	Saccharide[a]	3187847	3208881	21035
R1-20	Saccharide[a]	3241919	3263633	21715
Replicon 2				
R2-01	Saccharide[a]– Terpene	127023	156677	29655
R2-02	Ectoine	191459	201932	10474
R2-03	NRPS	200983	256027	55045
R2-04	Bacteriocin	383579	394394	10816
R2-05	Terpene	403275	425452	22178
R2-06	Saccharide[a]	547430	569019	21590
R2-07	Microcin	573153	593247	20095
R2-08	Microcin	714475	734569	20095
R2-09	Saccharide[a]	928238	949176	20939
R2-10	Saccharide[a]	973873	995078	21206
R2-11	trans-AT PKS (gladiolin biosynthesis)	1116311	1245536	129226
R2-12	NRPS	1438232	1501658	63427
R2-13	Homoserine lactone	1516842	1537456	20615
R2-14	Phosphonate	1769564	1811339	41776
R2-15	Saccharide[a]	1991137	2013725	22589
R2-16	Bacteriocin	2086779	2097795	11017
R2-17	NRPS	2183075	2240006	56932
R2-18	NRPS	2260626	2343800	83175
R2-19	Saccharide[a]	2354860	2392727	37868
R2-20	Saccharide[a]	2416528	2437805	21278
R2-21	Saccharide[a]	2474328	2501901	27574
R2-22	NRPS	2667396	2728619	61224
R2-23	Saccharide[a]	2721864	2742820	20957
R2-24	Saccharide[a]	2987525	3013635	26111
R2-25	trans-AT PKS	3449913	3530217	80305
R2-26	Saccharide[a]	3812517	3835729	23213
R2-27	Terpene	3867072	3888142	21071
R2-28	Saccharide[a]	3940351	3961667	21317
R2-29	Saccharide[a]– Terpene	4013228	4046521	33294
Replicon 3 (no predictions)				

[a]Predicted by the ClusterFinder module within antiSMASH 3.0

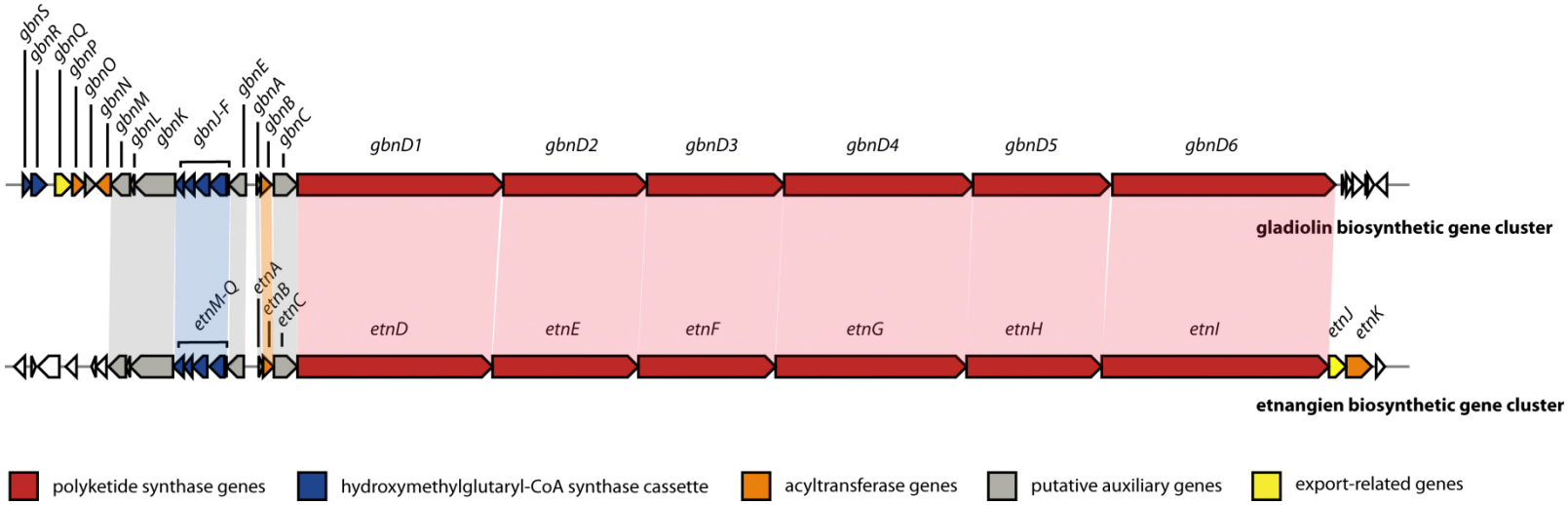

Figure S17| Comparison of the organization of the gladiolin and etnangien biosynthetic gene clusters in *B. gladioli* BCC0238 and *S. cellulosum*, respectively. Homologous genes are indicated by the colored shading. The proposed function of each protein encoded by the gladiolin gene cluster is listed in **Tables S5** and **S6**.

Table S5| Domain and module organization of the PKS-related genes in the gladiolin biosynthetic gene cluster.

PKS-related genes			
Gene/Protein	**Length bp/aa**	**Modules**	**Domains**
gbnD1/GbnD1	16,386/5,461	Load+M1	KS-ACP-ACP-ER
		M2	KS-KR-ACP
		M3	KS-DH-KR-ACP
		M4-//	KS-DH-KR-//
gbnD2/GbnD2	11,388/3,795	//-M4	//-ACP
		M5	KS-DH-KR-ACP
		M6	KS-ACP-ACP
		M7-//	KS-KR-//
gbnD3/GbnD3	10,875/3,624	//-M7	//-MT-ACP
		M8	KS-KR-ACP
		M9	KS-KR-ACP
		M10-//	KS-//
gbnD4/GbnD4	14,940/4,979	//-M10	//-DH-KR-ACP
		M11	KS-KR-ACP-KS-MT-ACP
		M12-//	KS-KR-ACP-KS-//
gbnD5/GbnD5	10,878/3,625	//-M12	//-DH-ACP
		M13	KS-KR-ACP
		M14	KS-DH-KR-MT-ACP
gbnD6/GbnD6	17,724/5,896	M15	ER-KS-KR-ACP
		M16	KS-DH-KR-MT-ACP
		M17	KS-KR-ACP-KS-ACP-TE

Table S6| Putative functions of the non-PKS-encoding genes within the gladiolin biosynthetic gene cluster

Gene/Protein	Length bp/aa	Similar Proteins	% Identity
gbnA/GbnA	264/88	Hypothetical protein (*Sorangium cellulosum*)	60
		Acyl carrier protein (*Streptomyces rimosus*)	49
		Acyl carrier protein (*Streptomyces monomycini*)	46
gbnB/GbnB	894/298	Malonyl-CoA-ACP transacylase (*Sorangium cellulosum*)	54
		Hypothetical protein (*Desulfovibrio inopinatus*)	46
		Hypothetical protein (*Sphingomonas sp.*)	47
*gbnC/*GbnC	1962/654	Asparagine Synthase (*Sorangium cellulosum*)	65
		Asparagine Synthase (*Labrenzia alba*)	51
		Asparagine Synthase (*Streptomyces sp.* DvalAA-43)	53
*gbnE/*GbnE	1389/463	Permease (*Sorangium cellulosum*)	68
		2-nitropropane dioxygenase (*Pseudomonas batumici*)	60
		BatK (*Pseudomonas fluorescens)*	60
*gbnF/*GbnF	246/82	Acyl carrier protein (*Sorangium cellulosum*)	68
		Hypothetical protein (*Paenibacillus pinihumi*)	59
		Poly(3-hydroxyalkanoate) depolymerase (*Bacillus subtilis)*	59
*gbnG/*GbnG	1212/404	Polyketide beta-ketoacyl:ACP synthase (*Sorangium cellulosum*)	65
		Polyketide beta-ketoacyl:ACP synthase (*Pelonsinus fermentans*)	57
		Polyketide beta-ketoacyl:ACP synthase (*Bacillus subtilis*)	60
*gbnH/*GbnH	1260/420	3-hydroxy-3-methylglutaryl-ACP synthase (*Sorangium cellulosum*)	99
		3-hydroxy-3-methylglutaryl-ACP synthase (*Sorangium cellulosum*)	76
		BonG (*Burkholderia gladioli*)	71
*gbnI/*GbnI	789/263	Enoyl-CoA hydratase/isomerase (*Sorangium cellulosum* So ce56)	63
		Enoyl-CoA hydratase (*Paenibacillus tyrfis*)	57
		Enoyl-CoA hydratase (*Paenibacillus sp.* 1-49)	57
*gbnJ/*GbnJ	747/249	Enoyl-CoA hydratase (*Sorangium cellulosum*)	99
		BatE (*Pseudomonas fluorescens*)	99
		Hypothetical protein (*Pseudomonas aeruginosa*)	100
*gbnK/*GbnK	3234/1078	SorM (*Sorangium cellulosum*)	99
		Hypothetical protein (*Sorangium cellulosum*)	99
		Hypothetical protein AB835 (*Candidatus endobugula sertula*)	38
*gbnL/*GbnL	360/120	alpha/beta hydrolase (*Burkholderia glumae*)	74
		alpha/beta hydrolase *(Burkholderia* plantarii)	74
		alpha/beta hydrolase (*Sorangium cellulosum*)	58
*gbnM/*GbnM	1401/467	Amidase (*Sorangium cellulosum*)	56
		SorP (*Sorangium cellulosum*)	46
		SorP (*Calothrix sp.* PCC7103)	39
*gbnN/*GbnN	1185/395	Malonyl-CoA-ACP transacylase (*Methylococcaceae bacterium* Sn10-6)	53
		Malonyl-CoA-ACP transacylase (*Scytonema sp. PCC 10023*)	

		Malonyl-CoA-ACP transacylase (*Fulvivirga imtechensis*)	48
			49
*gbnO/*GbnO	849/283	4'-phosphopantetheinyl transferase (*Collimonas arenae*)	47
		Hypothetical protein (*Janthinobacterium* sp. CG3)	44
		Hypothetical protein (*Azohydromonas lata*)	44
*gbnP/*GbnP	975/325	Malonyl-CoA-ACP transacylase (*Xanthomonas cannabis*)	50
		Malonyl-CoA-ACP transacylase (*Burkholderia thailandensis*)	53
		Malonyl-CoA-ACP transacylase (*Sorangium cellulosum So ce56*)	47
*gbnQ/*GbnQ	1353/451	Multidrug transporter MatE (*Sorangium cellulosum*)	68
		Multidrug transporter MatE (*Chromobacterium subtsugae*)	72
		Multidrug transporter MatE (*Rudaea cellulosilytica*)	59
*gbnR/*GbnR	792/264	Enoyl-CoA hydratase (*Burkholderia glumae*)	89
		Enoyl-CoA hydratase (*Burkholderia plantarii*)	86
		Enoyl-CoA hydratase (*Burkholderia ubonensis*)	79
*gbnS/*GbnS	1206/402	Enoyl-CoA hydratase (*Burkholderia glumae*)	83
		Enoyl-CoA hydratase (*Burkholderia pseudomallei*)	73
		Enoyl-CoA hydratase (*Burkholderia oklahomenesis*)	73

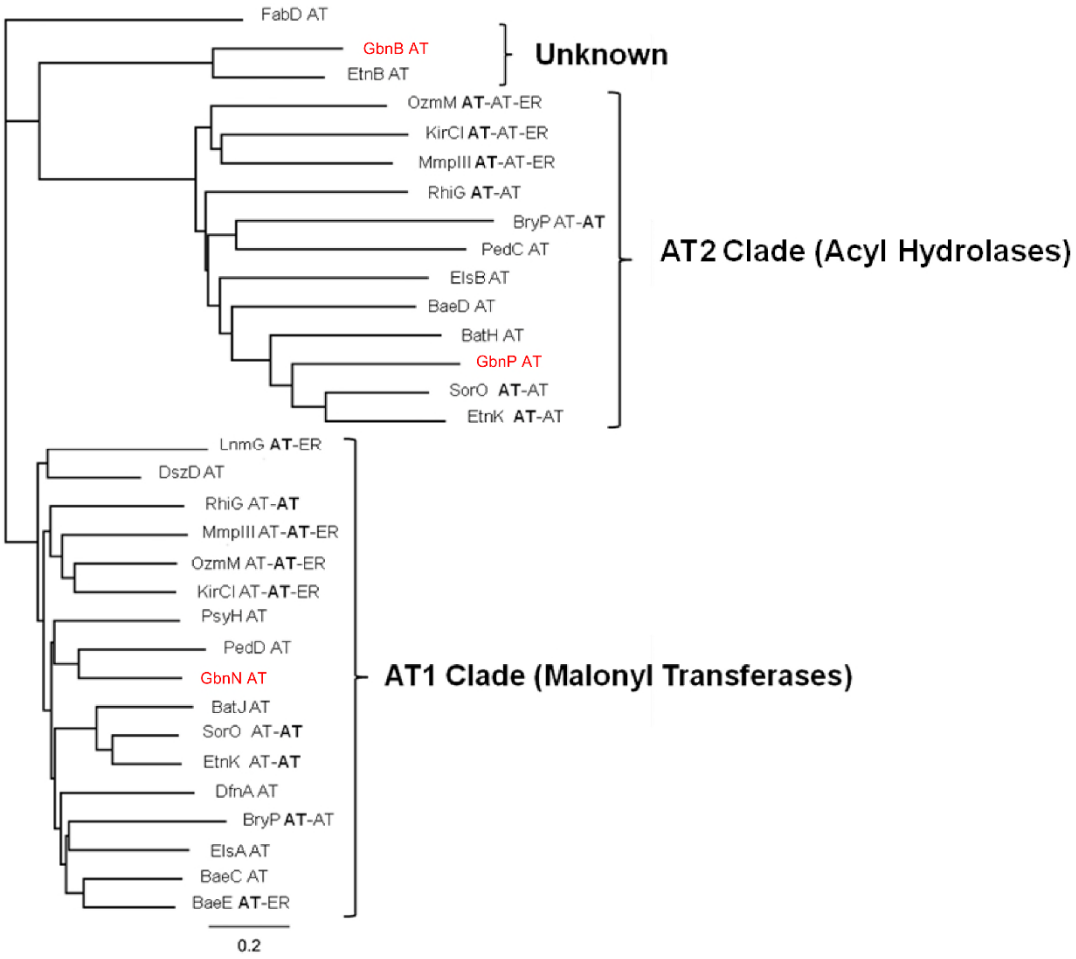

Figure S18| Phylogenetic analysis of AT-like domains in *trans*-AT PKSs.

Clade labels identify AT domain activity and tip labels consist of protein name and domain architecture. In the case of multidomain proteins, the appropriate AT domain is highlighted in bold. Sequences from the following pathways were used: Bae, bacillaene (*B. amyloliquefaciens*); Bat, batumin/kalimantacin (*Pseudomonas fluorescens*); Bry, bryostatin ("*Candidatus* Endobugula sertula"); Dfn, difficidin (*B. amyloliquefaciens*); Dsz, disorazol (*Sorangium cellulosum*); Etn, etnangien (*Sorangium cellulosum* So ce56); Els, elansolid (*Chitinophaga pinensis*); Gdn, gladiolin (*Burkholderia gladioli)*; Kir, kirromycin (*Streptomyces collinus*); Lnm, leinamycin (*Streptomyces atroolivaceus*); Mmp, mupirocin (*Pseudomonas fluorescens*); Ozm, oxazolomycin (*Streptomyces bingchenggensis* and *Streptomyces albus*); Psy, psymberin (uncultivated symbiont of sponge *Psammocinia* aff. *bulbosa*); Rhi, rhizoxin (*Burkholderia rhizoxina*); Sor, soraphen (*Sorangium cellulosum*) (3). Outgroup is the *E. coli* AT (FabD).

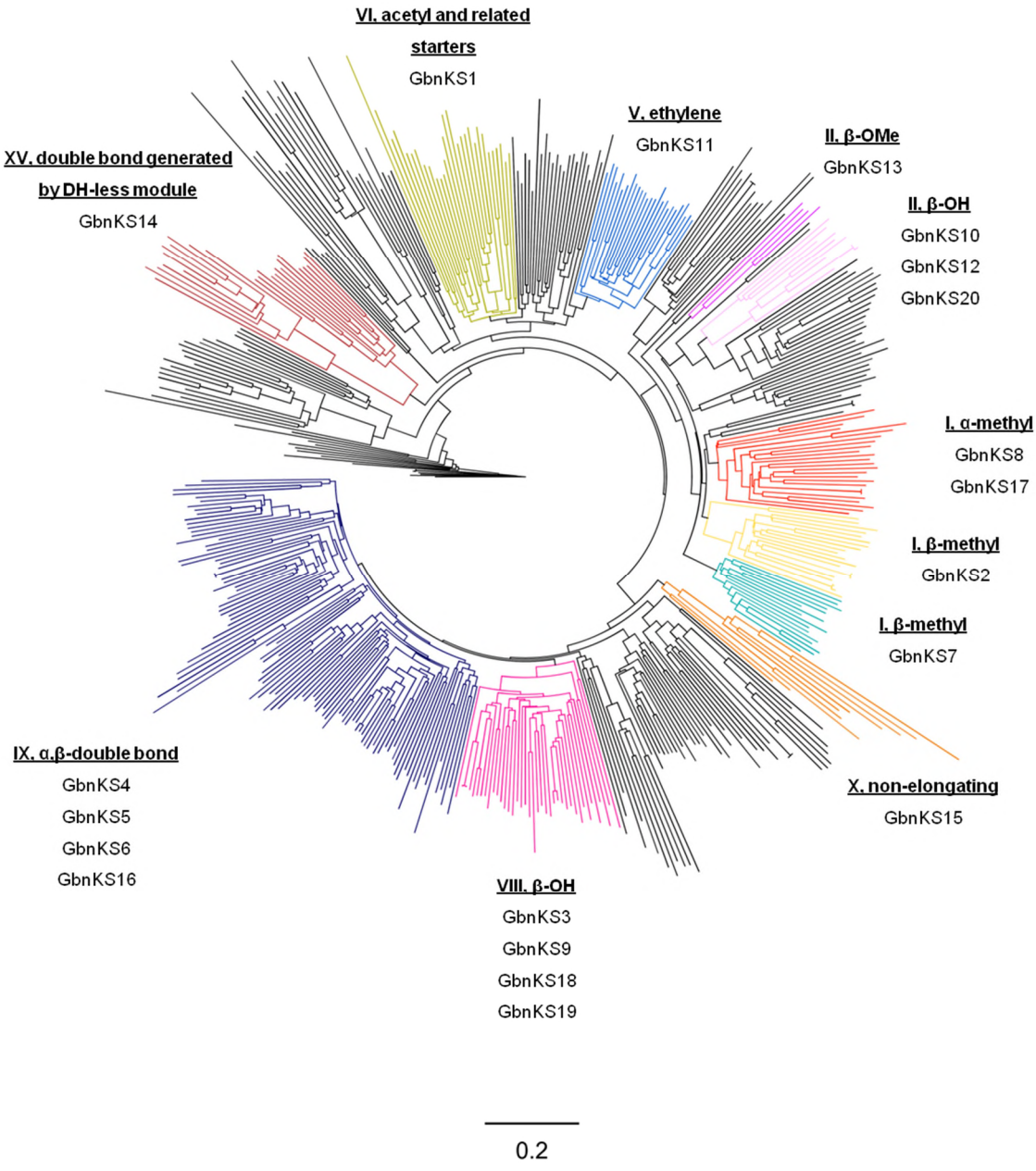

Figure S19| Phylogenetic analysis of KS domains in *trans*-AT PKSs.

Clade locations of gladiolin KS domains are highlighted, and their specificities labelled. The KS numbering denotes the position of the KS within the gene cluster starting from the upstream point. The EryKS4 was used as the outgroup.

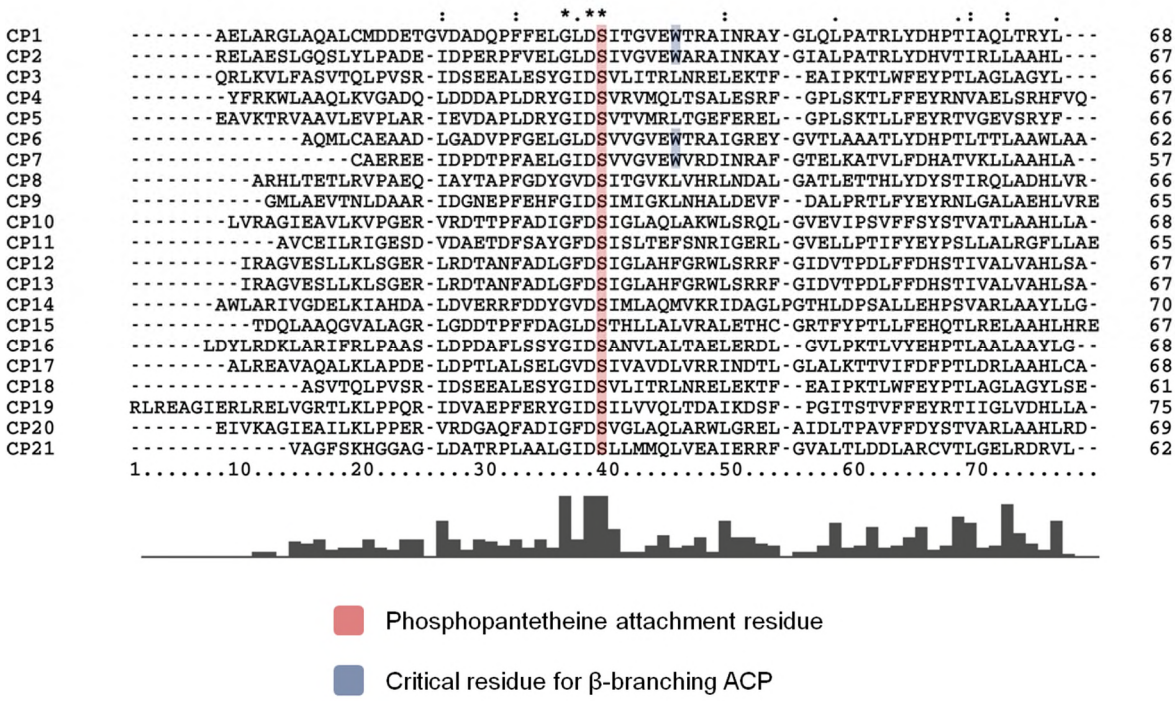

Figure S20| Sequence alignment of the ACP domains from the Gladiolin *trans*-AT PKS. The serine residue that serves as the attachment site for the 4'-PP cofactor is highlighted in red, the conserved tryptophan residue that flags ACP domains for β-branching is marked in blue.

Table S7| Predicted configurations of stereogenic carbon centers bearing hydroxyl groups in gladiolin. Sequence comparisons with KR domains of known stereospecificity were performed to determine the type of KR domain and thus predict the configuration of the stereocenters bearing the hydroxyl groups.

Ketoreductase	Diagnostic Asp Motif	KR Type	Resulting Stereochemistry
KR2	AGLI--D**D**AYA	B-type (*D*)	(*R*)
KR3	AGLL--R**D**AFL	B-type (*D*)	-
KR4	AGLL--R**D**AFL	B-type (*D*)	-
KR5	AGIV--R**D**NFV	B-type (*D*)	-
KR7	AIVL--K**D**RAL	B-type (*D*)	(*R*)
KR8	AGVL--R**D**SLI	B-type (*D*)	(*S*)
KR9	AGVA--TRGTL	A-type (*L*)	(*R*)
KR10	AGVL--R**D**GYV	B-type (*D*)	-
KR11	AGVA--TRGTL	A-type (*L*)	(*S*)
KR12	AGRVDFEAPAF	A-type (*L*)	-
KR13	AGLT--R**D**ALI	B-type (*D*)	-
KR14	ALVL--R**D**ASV	B-type (*D*)	-
KR15	AGLI--D**D**AYA	B-type (*D*)	(*S*)
KR16	AVGL--F**D**ESL	B-type (*D*)	(*R*)
KR17	AGVE--SRGSL	A-type (*L*)	(*R*)

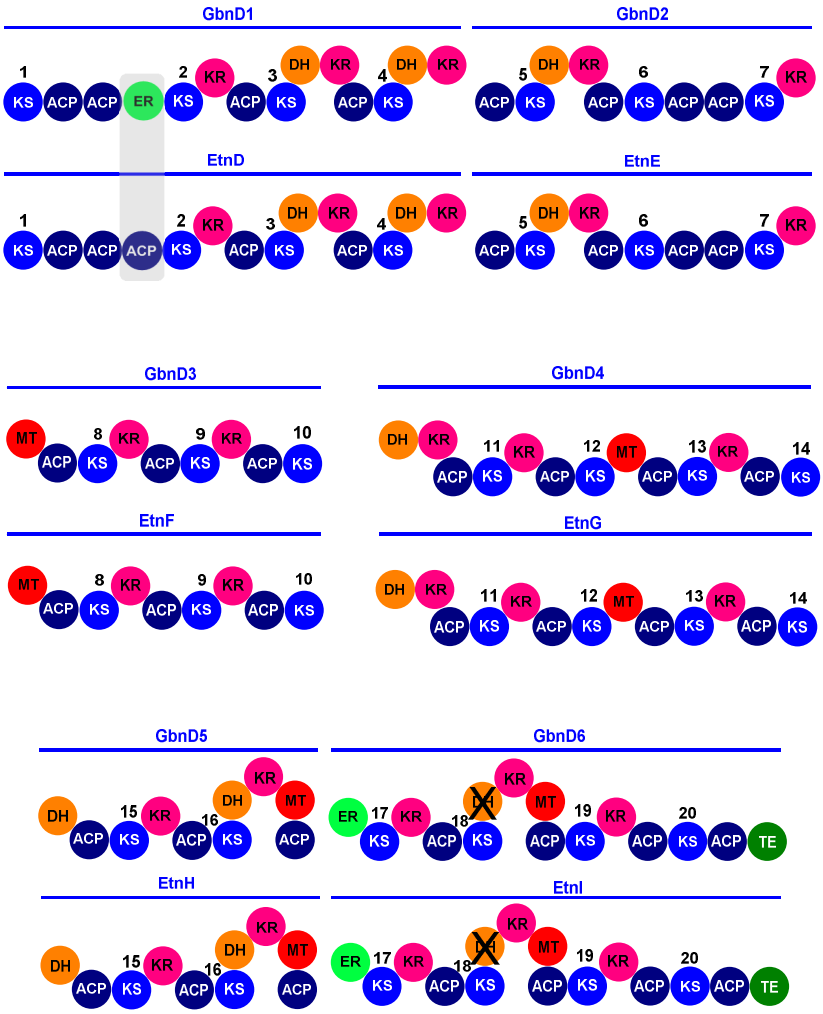

Figure S21| Comparison of gladiolin and etnangien PKSs. GbnD1-D6 is the gladiolin PKS and EtnD-I is the etnangien PKS. The only difference between the two multienzymes is the presence of an ER domain in GbnD1, which replaces an ACP domain in EtnD (highlighted in grey).