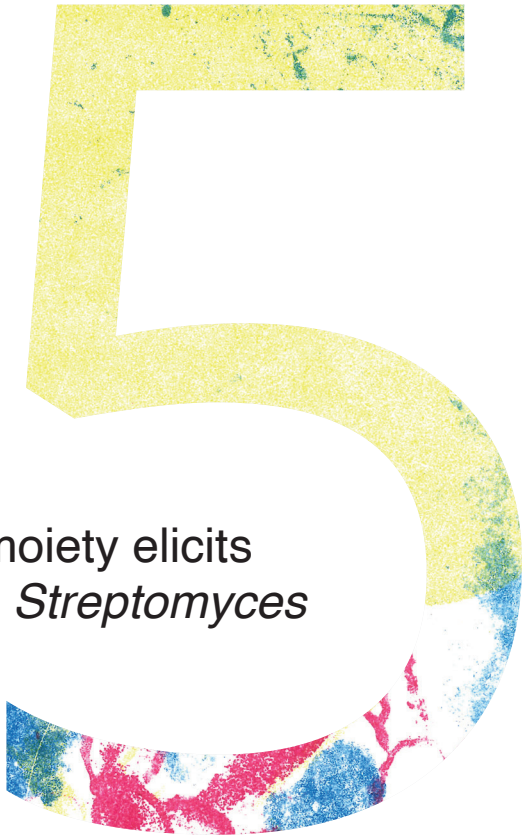
Ecology and genomics of Actinobacteria and their specialised metabolism

Bergeijk, D.A. van



The ubiquitous catechol moiety elicits angucycline production in *Streptomyces*

Doris A. van Bergeijk, Somayah S. Elsayed, Chao Du, Isabel Nuñez Santiago, Le Zhang, Victor J. Carrión, Herman P. Spaink and Gilles P. van Wezel

The work described in this chapter is part of the publication:
van Bergeijk, D.A., Elsayed, S.S., Du, C., Nuñez Santiago, I., Roseboom, A.M., Zhang, L., Carrión V.J., Spaink, H.P., van Wezel, G.P. The ubiquitous catechol moiety elicits siderophore and angucycline production in *Streptomyces*. *Commun Chem* **5**, 14 (2022).

Abstract

Actinobacteria are prolific producers of bioactive specialised metabolites. To unlock their full biosynthetic potential, better understanding of the regulatory networks and cognate signals that control antibiotic production is needed. Animal- and plant-associated molecules containing a catechol moiety, such as the human stress hormone adrenaline, enhance siderophore production in *Streptomyces* (Chapter 4). In this chapter, we analysed the response of Actinobacteria to catechol. Exploration of the catechol-responsive strain *Streptomyces* sp. MBT84 using mass spectral networking and proteomics revealed elicitation of a BGC that produces the angucycline glycosides aquayamycin, urdamycinone B, and galtamycin C. Heterologous expression of the catechol-cleaving enzymes catechol 1,2-dioxygenase or catechol 2,3 dioxygenase counteracted the eliciting effect of catechol. Taken together, these results illustrate that catechol, by itself and as part of plant- and animal-associated molecules, can act as elicitor of specialised metabolism.

Introduction

The global rise in drug resistance has urged the need for novel bioactive molecules [15,54]. Actinobacteria have been a prolific source of bioactive natural products with applications across medicine, agriculture, and biotechnology [10,11]. Over the past 20 years it has become clear that a huge amount of yet unexplored biosynthetic potential lies hidden in the genomes of Actinobacteria, which contain a wealth of uncharacterised biosynthetic gene clusters (BGCs) [20,27,50,241,266]. Many of these BGCs are silent, in other words, they are not (or hardly) expressed under routine laboratory growth conditions; this prevents discovery of their cognate natural products [29,93]. Heterologous expression and activation of these silent BGCs have shown that they encode functional biosynthetic pathways that produce novel molecules [29,30,267]. Hence, why are these BGCs silent and what activates their expression?

The metabolic versatility of Actinobacteria has been linked to their ecological diversity: Actinobacteria are widely distributed across Earth's ecosystems [2]. Within these environments, ecological forces have shaped how BGCs are regulated to ensure that the biosynthesis of their metabolically expensive products is only activated when needed [32-34]. Consequently, environmental signals have a major impact on the level and timing of specialised metabolism [37,44,138,252,268]. Identification of these eliciting signals and the regulatory pathways involved is needed to unlock the full biosynthetic potential of Actinobacteria.

In the previous chapter, we showed that the human stress hormone adrenaline enhances siderophore production by *Streptomyces* and this response was linked to the catechol moiety (Chapter 4). Interestingly, catechol is widely distributed across nature. It is common in plant metabolites [247,269] and is part of bacterial catecholate-type siderophores [5]. Additionally, catechol is present in the environment as degradation product of aromatic compounds such as polycyclic aromatic hydrocarbons originating from plant material or industrial pollution [270].

The specificity of the response of *Streptomyces* to the catechol moiety and its wide distribution in nature suggested catechol may be applied as a general elicitor of natural product biosynthesis in Actinobacteria. Here we show that catechol promotes the expression of a BGC that produces different (novel) angucycline glycosides. Expression of catechol-degrading enzymes counteracted the eliciting effect of catechol, highlighting the importance of the catechol moiety in the response. Taken together, our results provide further evidence for the potential of catechol as elicitor of specialised metabolism in *Streptomyces*.

Results

Catechol elicits the production of bioactive specialised metabolites

We have shown that plant- and animal-associated catechol compounds can impact the specialised metabolism of *Streptomyces* (Chapter 4). The specificity of this response to the catechol moiety suggested that catechol may be applied as a general elicitor of natural product biosynthesis in Actinobacteria. We therefore tested the effect of catechol on the bioactivity of some of the well-characterised Actinobacteria in our collection [138].

Similar to epinephrine, we observed both promotion and inhibition of bioactivity in the presence of catechol (Fig. S1). Catechol had a particularly strong effect on *Streptomyces* sp. MBT84, reproducibly enhancing its antibacterial activity against *B. subtilis* (Fig. S1, Fig. 1A). The bioactivity coincided with enhanced production of a yellow/brown pigment (Fig. 1A). When we investigated the specificity of this response by testing a library of structurally-related compounds including epinephrine, none of the other compounds increased antibiotic production (Fig. 1B). In Chapter 4, the effect of catechol on specialised metabolism was linked to iron availability. However, supplementation of the growth medium with iron did not affect the response of *Streptomyces* sp. MBT84 to catechol. Additionally, no siderophore production was observed in a CAS agar overlay. This strongly suggests that the change in bioactivity and pigment production was not due to altered siderophore production.

Catechol enhances angucycline production in *Streptomyces* sp. MBT84

To identify the nature of the metabolites produced by *Streptomyces* sp. MBT84 in response to catechol, the strain was streaked on MM agar plates and the metabolites were then extracted from the spent agar after five days of growth using ethyl acetate (EtOAc). The crude extracts of the catechol-grown cultures showed increased bioactivity as compared to those grown under control conditions, while media blanks showed no activity (data not shown). We wanted to rule out that catechol potentiated the bioactivity via an effect on the indicator bacteria rather than as elicitor. For this, we tested whether the presence of catechol in the growth medium affected the size of the inhibition zone of different antibiotics and the crude extracts of MBT84 against *B. subtilis*. No differences were observed (Fig. S2), which shows that addition of catechol to the growth medium does not affect the susceptibility of the indicator strains against antibiotics, nor does it potentiate the activity of the crude extracts.

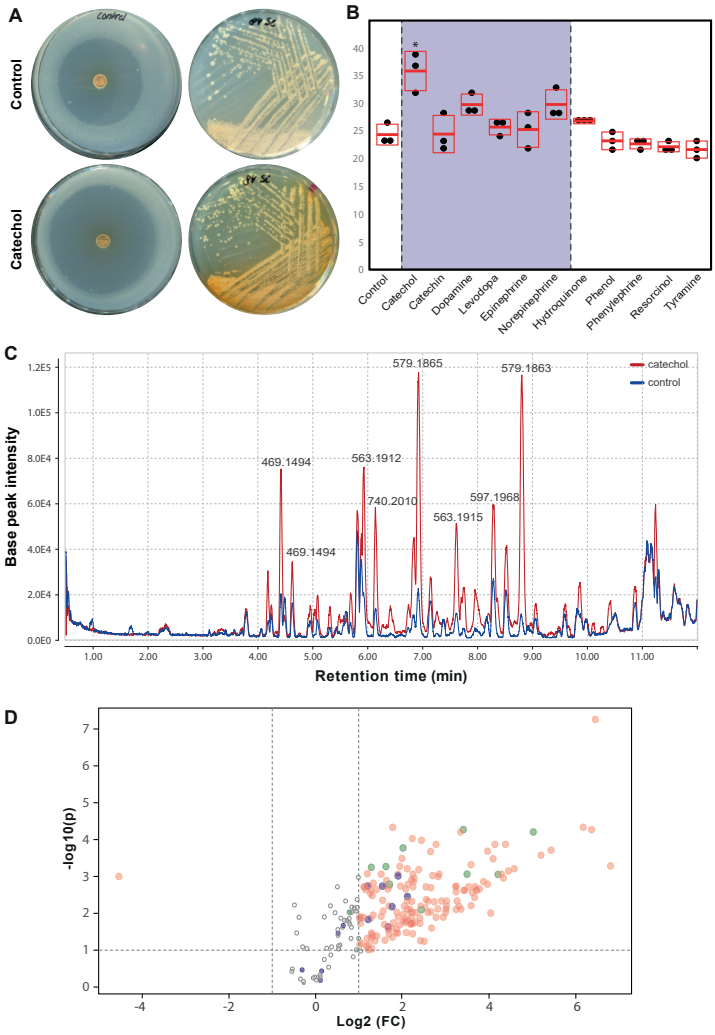


Figure 1. Catechol elicits the bioactivity and metabolite production of *Streptomyces* sp. MBT84. **A)** The antibacterial activity against *B. subtilis* 168 and yellow pigment production of *Streptomyces* sp. MBT84 are increased in the presence of catechol. **B)** None of the structurally related compounds, including epinephrine, increased the antibiotic production by *Streptomyces* sp. MBT84. One-way ANOVA, followed by a *post hoc* Tukey's HSD test, was performed to compare the difference in bioactivity between the growth conditions (* = $p < 0.001$, $n = 3$, mean and standard deviation are indicated in red, the grey box highlights the response to compounds with a catechol moiety). **C)** LC-MS chromatogram overlay of the crude extract of MBT84 grown with and without catechol. Multiple peaks were increased in intensity in the presence of catechol. **D)** Volcano plot highlighting the increased metabolite production by *Streptomyces* sp. MBT84 in response to catechol. The x and y axes represent the $\log_2$ fold changes and the corresponding $-\log_{10}$ FDR-adjusted *p*-value of all the mass features, respectively. Red circles represent the mass features in catechol-grown cultures with a significant intensity difference of more than 2-fold compared to control cultures (FDR-adjusted $p \leq 0.1$). Circles situated in the top left and top right quadrants represent the mass features which are most induced or repressed, respectively, by catechol with high statistical significance. The *m/z* values 469.149 and 579.186 are shown in purple and green, respectively ($n = 6$).

Liquid chromatography-mass spectrometry (LC-MS) analysis revealed multiple peaks in the chromatogram that were substantially overrepresented in the samples obtained from the catechol-grown cultures (Fig. 1C). The metabolomic profile of *Streptomyces* sp. MBT84 grown with and without catechol was compared through a volcano plot generated using MetaboAnalyst [271] (Fig. 1D). In the volcano plot, the majority of the mass features were significantly upregulated in catechol-grown cultures. The most notable ones were those with *m/z* values 469.1494 and 579.1864, which represented several highly intense mass features sharing the same exact masses but with different retention times. Additionally, multiple in-source fragments were among the mass features which were significantly upregulated with catechol. The mass features related to *m/z* 469.1494 and 579.1864 are likely due to yellowish-coloured metabolites, since they showed in the ultraviolet (UV) chromatogram corresponding peaks having UV absorption maxima of 420-440 nm. This corresponds well to the observed increase in yellow pigmentation when *Streptomyces* sp. MBT84 was grown in the presence of catechol.

To identify which *m/z* values correlated to the increased bioactivity of *Streptomyces* sp. MBT84, crude extracts were analysed using at-line nanofractionation coupled to LC-MS, using a method described previously [272]. In brief, the LC effluent is split post-column, after which one tenth of the flow goes to the MS, while the rest was fractionated into a 384-well plate (6 seconds/ well). Following LC solvent evaporation, the antibacterial activity of the collected nanofractions is assessed in a resazurin reduction assay and directly correlated to the mass spectral peaks. Reconstruction of the bioactivity chromatograms revealed multiple negative maxima, reflecting bioactivity. After alignment to the LC-MS chromatogram, the *m/z* values of the bioactive peaks were determined, which included those which correspond to the yellowish metabolites (*m/z* 469.149 and 579.186) (Fig. 2). In the extracted ion chromatogram of *m/z* 469.15, the multiple peaks observed were either due to quasi-molecular ions (like peaks number 7 and 9), or due to fragment ions of *m/z* 487.16 (like peaks number 1 and 2). The same pattern was observed for the peaks due to *m/z* 579.19, which were mostly fragment ions of *m/z* 597.20.

To gain more insight into which metabolites were induced by catechol and whether these were structurally related, molecular networking was employed using the Global Natural Products Social (GNPS) molecular networking web tool. This generates a network wherein molecules with related scaffolds cluster together [220]. Due to the presence of multiple isomeric metabolites, we used the classical molecular networking workflow while enabling the MSCluster algorithm [220,273]. This allows merging all mass features with similar MS and MS/MS spectra into one node, regardless of their different retention times. A network representing the ions detected in the crude extract of *Streptomyces* sp. MBT84 grown with and without catechol was constructed, revealing 258 nodes clustered in 21 spectral families (Fig. 3). Within the biggest spectral family,

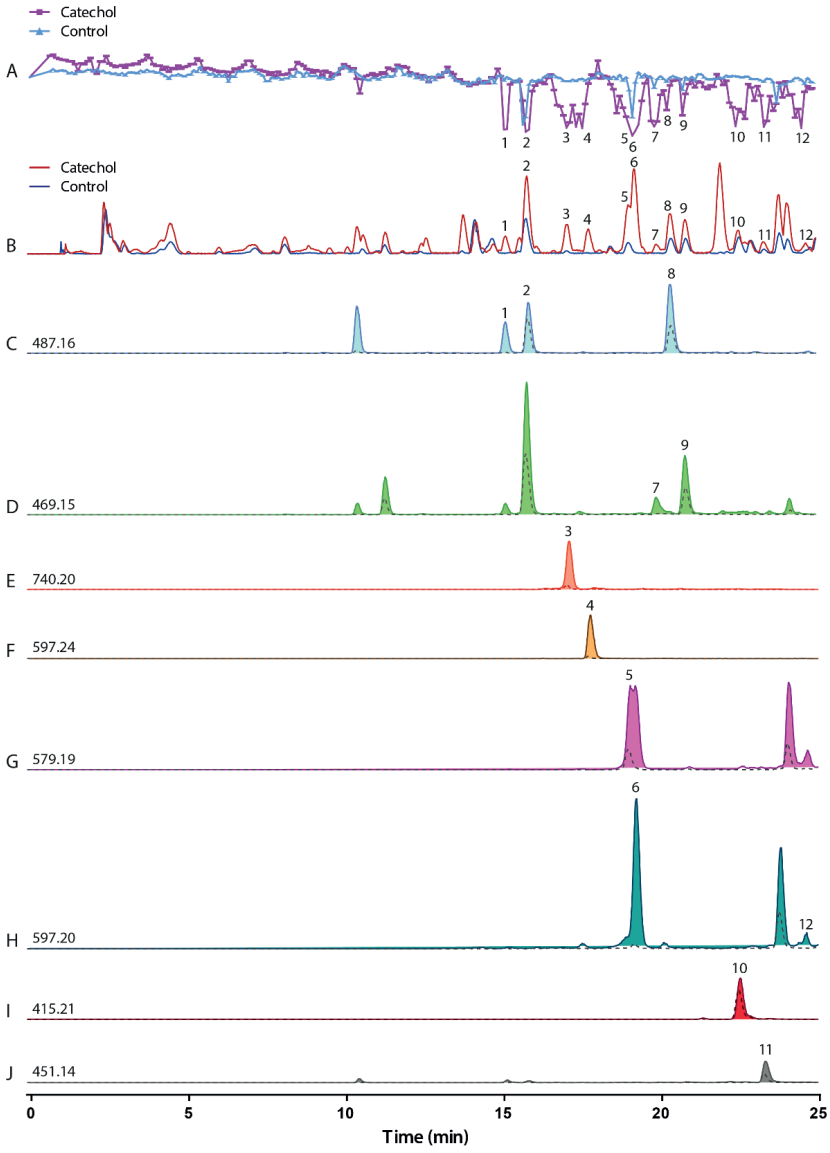


Figure 2. Nanofractionation reveals induction of several bioactive metabolites in *Streptomyces* sp. MBT84 by catechol. A) Bioactivity chromatogram overlay of the nanofractionated control and catechol-grown extracts. The negative peaks represent the bioactive compounds. The results of the resazurin reduction assay have been plotted against their corresponding LC-MS chromatograms after adjusting for the delay. **B)** Total ion current LC-MS chromatogram overlay of control and catechol-grown extracts. **C-J)** Extracted ion chromatogram overlay of m/z values 487.16 (C), 469.15 (D), 740.20 (E), 597.24 (F), 579.19 (G), 597.20 (H), 415.21 (I), and 451.14 (J) in both control (dashed line) and catechol-grown (solid line) extracts, which could be aligned to the negative maxima of the bioactivity chromatogram.

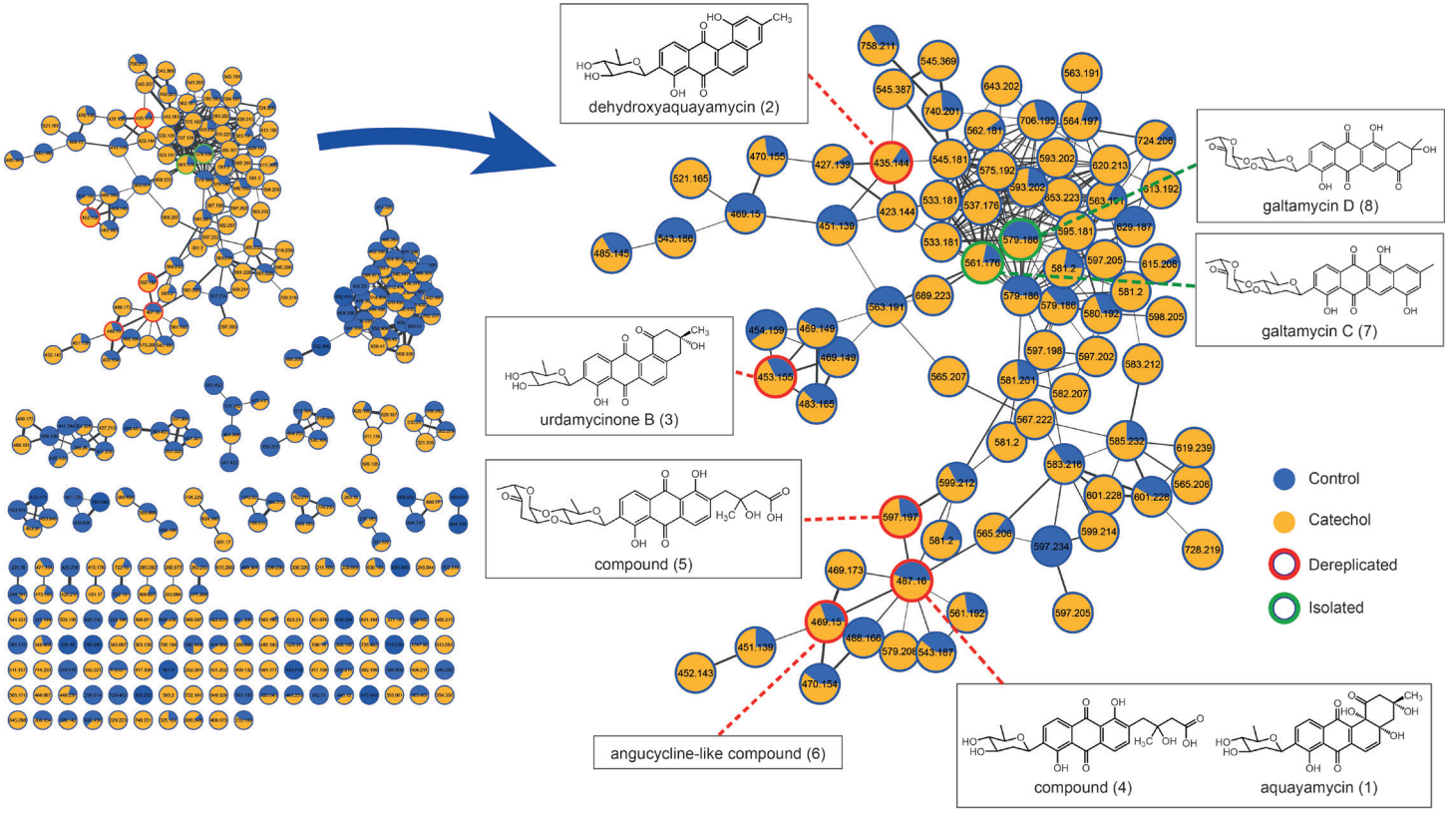


Figure 3. Molecular network of the ions detected in the crude extracts of *Streptomyces* sp. MBT84 revealing a large spectral family of angucycline compounds elicited by catechol. A pie chart was mapped to the nodes which represents the abundance of each m/z value in the control (blue) and catechol-grown crude extracts (yellow). Nodes highlighted in red represent dereplicated metabolites, while those highlighted in green represent the compounds isolated and identified in this study. *Streptomyces* sp. MBT84 was grown for five days on MM agar plates with or without 100 μM catechol (n = 6).

dereplication based on matching MS/MS spectra against the GNPS spectral library annotated several *m/z* values as being known metabolites belonging to the class of angucyclines, namely aquayamycin (**1**), dehydroxyaquayamycin (**2**), and urdamycinone B (**3**). Additionally, two structurally related compounds (**4,5**) as well as an angucycline-like compound (**6**), were annotated. This angucycline spectral family was more abundant in the presence of catechol and included the majority of the mass features that were correlated to bioactivity by nanofractionation (Fig. 2).

To ascertain that this network indeed consisted of angucycline glycoside-like compounds, we isolated the molecules with *m/z* 561.1760 (**7**) and 579.1866 (**8**). For this, *Streptomyces* sp. MBT84 was grown confluently on 12 x 12 mm MM agar plates supplemented with catechol and the metabolites were extracted with EtOAc. Following chromatographic isolation, **7** (red amorphous powder) and **8** (yellow amorphous powder) were obtained. The final structures of **7** and **8** were elucidated using nuclear magnetic resonance spectroscopy (NMR) (Supplemental data). Based on that, **7** was identified as galtamycin C, which has previously been isolated from an intertidal sediments-derived *Streptomyces* sp.[274]. As for **8**, it was found to be 1-oxo-3-hydroxy-3,4-dihydro-2*H*-galtamycin C based on the different 2D NMR correlations observed in comparison with **7**, together with the molecular formula and degrees of unsaturation obtained from its accurate mass (Supplemental data). To our knowledge **8** has not been previously described, and was thus designated as galtamycin D. Taken together, our data show that catechol enhances the biosynthesis of angucycline glycosides in *Streptomyces* sp. MBT84.

Identification of the angucycline glycoside BGC in *Streptomyces* sp. MBT84

To identify the gene cluster responsible for the biosynthesis of the angucycline glycosides in *Streptomyces* sp. MBT84, we obtained its full genome sequence using PacBio sequencing (GenBank accession number: JAHTGP000000000). A draft genome was assembled resulting in 3 contigs (Table S3). Analysis using antiSMASH 6.0 [217] revealed 21 putative BGCs (Table S4) of which BGC4 shows high similarity to the saquayamycin A BGC, and was therefore most likely responsible for the production of the angucycline glycosides (Fig. 4A, Table S5).

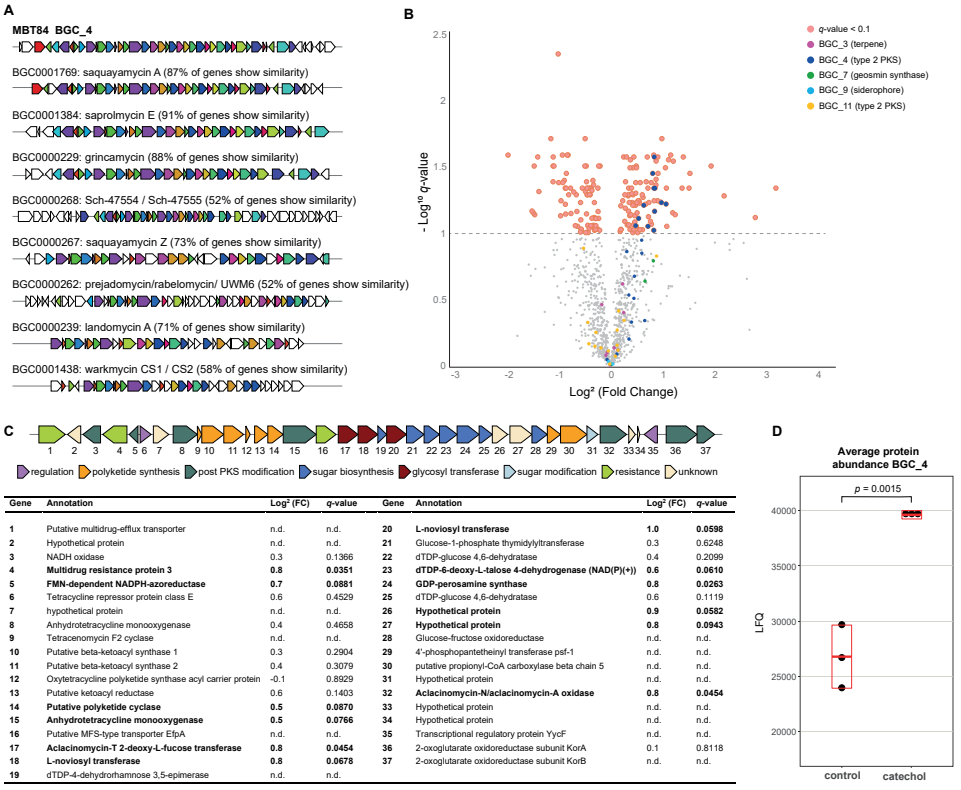
We have previously shown the applicability of quantitative proteomics combined with metabolomics to connect BGCs to the natural product(s) they specify [249]. Therefore, MS-based quantitative proteomic analysis was performed on total protein samples obtained from *Streptomyces* sp. MBT84 grown on MM agar plates with or without 100 μM catechol covered with a cellophane disc. After five days of growth, the biomass was harvested and snap-frozen in liquid nitrogen. Subsequent quantitative proteomics analysis was performed on three replicate samples per growth condition, yielding 1322 quantifiable proteins, of which 187 were significantly differentially expressed between catechol-grown and control cultures (FDR-

adjusted *p*-value < 0.1; Fig. 4B). Proteins belonging to five of the 21 putative BGCs annotated by the antiSMASH algorithm were expressed under the chosen conditions. However, only the biosynthetic proteins of BGC4 were mostly significantly upregulated in catechol-grown cultures, which provided validation of the involvement of this BGC in the observed change in bioactivity induced by catechol (Fig. 4B and C). The minimal polyketide synthase (PKS) enzymes responsible for the generation of the initial angucycline or angucyclinone backbone [275], namely two ketosynthase units (KSα and KSβ) and an acyl carrier protein, were not significantly differentially expressed (Fig. 4C).

To confirm that BGC4 was indeed responsible for the production of the angucycline glycosides, we knocked down transcription of the gene for the KSβ of the minimal PKS using CRISPR-dCas9 interference [276,277]. The knock-down was enforced from the CRISPRi construct pGWS1517 that expresses a spacer targeting the non-template strand of the gene for KSβ. Pigment production was strongly inhibited in the strain with reduced expression of KSβ, while no inhibition was seen in the control strain harbouring a construct (pGWS1516) that targets the template strand (Fig. S3). Metabolomics analysis revealed that the production of the angucycline glycosides was significantly reduced when the expression of KSβ was inhibited (Fig. S3). All the evidence together proves that indeed BGC4 specifies the angucyclines.

Expression of catechol-degrading enzymes counteracts antibiotic elicitation

Various microorganisms (e.g., *Pseudomonades*) enzymatically degrade aromatic compounds [270]. To establish whether catechol itself is essential for the eliciting effect, we expressed two enzymes in *Streptomyces* sp. MBT84 that degrade catechol. Catechol can be cleaved into *cis, cis*-muconic acid by catechol 1,2-dioxygenase (C12O) (*ortho*-pathway) or into 2-hydroxymuconic semialdehyde by catechol 2,3-dioxygenase (C23O) (*meta*-pathway) (Fig. S4) [270]. We wondered whether these degradation products would have a similar eliciting effect on angucycline production or whether the effect of catechol would disappear once degraded. To this end, we separately cloned the *catA1* gene encoding C12O from *Pseudomonas putida* KT2440 and the *xylE* gene encoding C23O from the promoter-probe vector pIJ4083 [278] onto the multi-copy vector pWHM3. The strong constitutive promoter SF14 [279] was used to drive expression. The resulting constructs pGWS1519 and pGWS1520 were introduced in *Streptomyces* sp. MBT84 via protoplast transformation, whereby the empty vector was used as a control. Of each recombinant strain (named MBT84-pWHM3, MBT84-C12O, MBT84-C23O) three independent transformants were tested for their ability to respond to catechol.



Gene	Annotation	Log^2 (FC)	*q*-value	Gene	Annotation	Log^2 (FC)	*q*-value
1	Putative multidrug-efflux transporter	n.d.	n.d.	**20**	**L-noviosyl transferase**	**1.0**	**0.0598**
2	Hypothetical protein	n.d.	n.d.	**21**	Glucose-1-phosphate thymidylyltransferase	0.3	0.6248
3	NADH oxidase	0.3	0.1366	**22**	dTDP-glucose 4,6-dehydratase	0.4	0.2099
4	**Multidrug resistance protein 3**	**0.8**	**0.0351**	**23**	**dTDP-6-deoxy-L-talose 4-dehydrogenase (NAD(P)(+))**	**0.6**	**0.0610**
5	**FMN-dependent NADPH-azoreductase**	**0.7**	**0.0881**	**24**	**GDP-perosamine synthase**	**0.8**	**0.0263**
6	Tetracycline repressor protein class E	0.6	0.4529	**25**	dTDP-glucose 4,6-dehydratase	0.6	0.1119
7	hypothetical protein	n.d.	n.d.	**26**	**Hypothetical protein**	**0.9**	**0.0582**
8	Anhydrotetracycline monooxygenase	0.4	0.4658	**27**	**Hypothetical protein**	**0.8**	**0.0943**
9	Tetracenomycin F2 cyclase	n.d.	n.d.	**28**	Glucose-fructose oxidoreductase	n.d.	n.d.
10	Putative beta-ketoacyl synthase 1	0.3	0.2904	**29**	4'-phosphopantetheinyl transferase psf-1	n.d.	n.d.
11	Putative beta-ketoacyl synthase 2	0.4	0.3079	**30**	putative propionyl-CoA carboxylase beta chain 5	n.d.	n.d.
12	Oxytetracycline polyketide synthase acyl carrier protein	-0.1	0.8929	**31**	Hypothetical protein	n.d.	n.d.
13	Putative ketoacyl reductase	0.6	0.1403	**32**	**Aclacinomycin-N/aclacinomycin-A oxidase**	**0.8**	**0.0454**
14	**Putative polyketide cyclase**	**0.5**	**0.0870**	**33**	Hypothetical protein	n.d.	n.d.
15	**Anhydrotetracycline monooxygenase**	**0.5**	**0.0766**	**34**	Hypothetical protein	n.d.	n.d.
16	Putative MFS-type transporter EfpA	n.d.	n.d.	**35**	Transcriptional regulatory protein YycF	n.d.	n.d.
17	**Aclacinomycin-T 2-deoxy-L-fucose transferase**	**0.8**	**0.0454**	**36**	2-oxoglutarate oxidoreductase subunit KorA	0.1	0.8118
18	**L-noviosyl transferase**	**0.8**	**0.0678**	**37**	2-oxoglutarate oxidoreductase subunit KorB	n.d.	n.d.
19	dTDP-4-dehydrorhamnose 3,5-epimerase	n.d.	n.d.				

Figure 4. Identification of the BGC responsible for the production of angucycline glycosides in *Streptomyces* sp. MBT84 using antiSMASH and MS-based quantitative proteomics. A) KnownClusterBlast output from antiSMASH which shows similar clusters from the MIBiG database. Genes marked with the same colour are interrelated; white genes have no relationship. **B)** Volcano plot of MS-based quantitative proteomics for cultures grown with and without catechol ($n = 3$). Proteins with an FDR-adjusted p-value ≥ 0.1 are greyed out. Proteins with a positive $\log_2$ fold change are higher expressed in catechol-grown cultures. **C)** BGC4 coding for the biosynthesis of angucycline glycosides. Annotations are based on BLAST homology searches and genes are colour-coded based on putative function. Significantly differentially expressed proteins are depicted in bold (FDR-adjusted $p < 0.1$, two-sample t-test, $n = 3$) **D)** Average protein level of BGC4 of cultures grown with and without catechol ($n = 3$). Mean and standard deviation are indicated in red. Two sample t-test was done showing BGC4 is significantly differentially expressed in presence of catechol.

The transformants were grown for five days on MM agar plates with and without 100 μM catechol, and the metabolites were extracted as described above. While catechol induced a clear increase in yellow pigmentation in MBT84 harbouring the control plasmid, such induction was no longer visible when MBT84-C12O was grown in the presence of catechol (Fig. 5A). In strain MBT84-C23O, the effect of catechol on the yellow pigmentation was strongly reduced, but in some transformants still some induction was observed. The visual

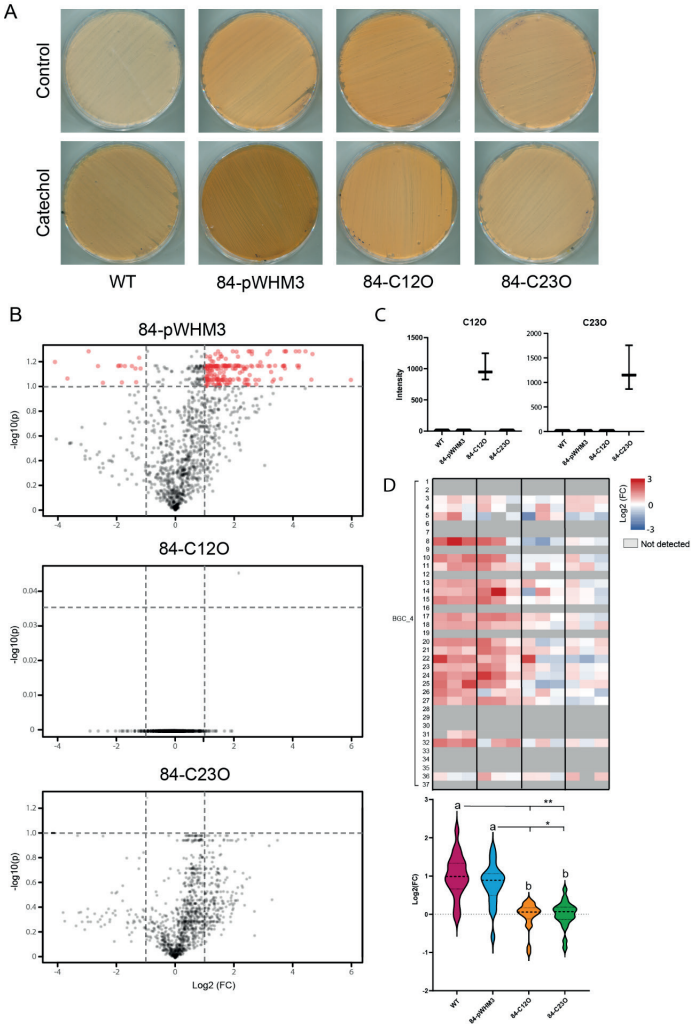


Figure 5. Heterologous expression of the catechol cleaving enzymes C12O and C23O resulted in a strong reduction of the eliciting effect of catechol. A) After five days of growth on MM with and without 100 μM catechol a clear increase in yellow pigmentation is visible in the WT and MBT84-pWHM3 (empty plasmid control) strain. This increase is not visible when C12O and C23O are expressed. **B)** Volcano plot highlighting the changes in metabolite production of the different recombinant strains in response to catechol. While no difference in metabolite profile was found when C12O is expressed, the metabolite profile of MBT84-C23O follows a similar distribution as the control but none of the mass features are significantly upregulated in the catechol-grown cultures (fold change >2 and FDR-adjusted *p*-value < 0.1) (n = 3). **C)** Protein abundance of the catechol dioxygenases in the WT and recombinant strains (n = 3) **D)** Expression profile of BGC4 coding for the biosynthesis of angucycline glycosides in the WT and recombinant strains. The heatmap shows the $\log_2$ fold change of protein level of each transformant (or WT replicate) comparing samples from catechol-grown cultures and control cultures (n = 3). The violin plots show the distribution of the $\log_2$ fold change of BGC4 expression in response to catechol (median = dashed line, quartiles = dotted line). Note that the empty vector control follows the same pattern as the WT strain, while this pattern is not visible when C12O and C23O are expressed. The average $\log_2$ fold changes in protein level of BGC4 in response to catechol were compared by one-way ANOVA, followed by a *post hoc* Tukey's HSD test. Similar letters indicate no significant difference (* = $p < 0.05$, ** $p < 0.01$).

assessment of antibiotic production was confirmed by detailed metabolomic analysis (Fig. 5B). Indeed, while catechol again elicited the production of metabolites in the control strain, catechol did not alter the metabolic profile of transformants expressing the catechol-degrading enzyme C12O or C23O. However, for the latter, the volcano plot showed a similar skewing to the upregulated side as the control, which suggests a reduced response.

To analyse the expression of the catechol dioxygenases and of BGC1.4 in the recombinant strains, we used quantitative proteomic analysis as described previously [216]. C12O and C23O were expressed individually in the respective recombinant strains (Fig. 5C). To study the protein levels of the gene products of BGC4 of the transformants or their parent with control plasmid, a heatmap of the $\log_2$ fold changes in expression between catechol and control was generated. Additionally, violin plots were generated to show the distribution of the expression differences of BGC1.4 proteins (Fig. 5D). The small sample size and variability between the different transformants, did not allow for a thorough statistical analysis. However, while catechol induced a similar upregulation of the expression of BGC1.4 in MBT84-pWHM3 compared to WT, this pattern was no longer visible in MBT84-C12O and MBT84-C23O. These results indicate that catechol is the elicitor and not one of its degradation products.

Discussion

Actinobacteria harbour many BGCs that are not or poorly expressed under laboratory growth conditions. A better understanding of the environmental signals that influence BGC expression as well as the regulatory pathways involved, will allow us to further explore actinobacterial specialised metabolism for therapeutic uses. Our work reveals catechol as elicitor of angucycline production in *Streptomyces* sp. MBT84. Angucyclines constitute the largest group of aromatic polyketides and are well known for their structural diversity and their therapeutic potential as anticancer and antibiotic compounds [275]. Our recent discovery of lugdunomycin and a range of novel angucyclines revealed that this extensively explored family of polyketides still has potential for drug discovery [60]. Indeed, the new compound described in this study, designated as galtamycin D, is another example that many angucyclines and derivatives are still to be discovered. Environmental cues may not only elicit the expression of (cryptic) biosynthetic pathways, but also of previously unseen branches of known pathways. Indeed, co-cultivation of the model streptomycete *Streptomyces coelicolor* A3(2) with *Aspergillus niger* elicited the production of GTRI-02, which is produced by the actinorhodin biosynthetic pathway [89]. The type II polyketide actinorhodin has been studied extensively by many laboratories for over 60 years, and the surprising discovery of an entirely new branch of its biosynthetic pathway underlines the importance of microbial interactions for drug discovery. Similarly, supplementation of

catechol to growth conditions might offer a new strategy to further explore the immense chemical space of angucyclines and other polyketides.

To coordinate appropriate production of specialised metabolites, Actinobacteria have evolved a vast array of complex, multi-level regulatory pathways of which many remain to be elucidated. The identification of a signal that influences specialised metabolism can provide a first step towards the elucidation of a pathway, as was illustrated by the discovery that GlcNAc had a stimulating effect on antibiotic production of many *Streptomyces* sp.[36]. Elucidating the signal transduction pathway identified DasR as pleiotropic transcriptional repressor of several pathway-specific activators, including ActII-ORF4 and RedD, whereby GlcN-6P acts as ligand for DasR [36]. However, the mechanism via which catechol impacts angucycline production in *Streptomyces*, remains yet unexplained. Our experiments show that degradation of catechol will annihilate its effect. Heterologous expression of the catechol-cleaving enzymes C12O or C23O in *Streptomyces* sp. MBT84 strongly reduced its response to catechol. This strongly suggests that catechol is the elicitor and not one of its degradation products. Of the two enzymes, C23O had slightly less impact than C12O, whereby the latter fully counteracted the effect of catechol. Our proteomics experiments show that both enzymes are expressed well, and we anticipate that the difference may lie in the nature of their degradation products. C23O is responsible for the extradiol ring cleavage of catechol, producing 2-hydroxymuconic semialdehyde in which the 1,2 di-hydroxy architecture of catechol is preserved, while C12O does not maintain this scaffold. This could indicate that specific protein interactions are involved in the response to catechol and might also explain why *Streptomyces* sp. MBT84 only responds to catechol. Catecholamines are protonated at neutral pH, which could interfere with protein binding. The next step will be to elucidate the signal transduction pathway for catechol and identify the key regulatory network it targets inside the cell.

Catechol impacts different types of specialised metabolites, as exemplified by its effect on the biosynthesis of angucyclines and siderophores (Chapter 4). Preliminary experiments in our laboratory revealed many elicited mass features in the metabolome of other *Streptomyces* species grown in presence of catechol (not shown). Our data indicate that these mass features belong to other classes of specialised metabolites, such as non-ribosomal peptides, which we are currently investigating. Interestingly, high throughput elicitor screening (HiTES) technology identified the plant-derived piceatannol as elicitor of the expression of a BGC for a cryptic NRPS[280]. Considering the data presented in our work, we hypothesise that this response might also be mediated through the catechol moiety that is present in piceatannol.

Our work again provides a new example of the power of using chemical elicitors to activate cryptic BGCs and identify the corresponding metabolites. Crude extracts are complex mixtures of metabolites and identifying the bioactive metabolites is often time consuming. Additionally, many 'cryptic' BGCs have not yet been linked to a natural product. We propose catechol as a new addition to the catalogue of elicitors that could be used in screens for novel bioactive molecules, which may allow for prioritization of bioactive metabolites for isolation from complex crude extracts and assist in 'decrypting' cryptic BGCs. A major issue to solve in the future is to identify the signal transduction pathway by which streptomycetes sense catechol, from internalization through metabolism to the activation of BGCs. Understanding of the mechanisms involved is of great interest as they may offer new tools for natural product-based drug discovery.

Methods

Bacterial strains, growth conditions and antimicrobial activity assay

All media and routine *Streptomyces* techniques have been described previously[227]. The bacterial strain collection used in this study were obtained from the Leiden MBT strain collection. *Streptomyces sp.* MBT84 has been isolated from soil samples collected from the QinLing Mountains (Shanxi province, China) [138]. *Bacillus subtilis* 168 and *Escherichia coli* ASD19 [237] were used as indicator strains for antimicrobial activity and were cultured in LB media at 37 °C.

Antimicrobial activity assays were conducted using the double-layer agar method. Strains were spotted on minimal medium agar plates (MM) supplemented with 0.5% mannitol and 1% glycerol (w/v) as non-repressing carbon sources, and nutrient agar (NA) (Difco) plates, using a pin replicator. For individual testing of strains, 2 μL spore stock was manually spotted. Growth media were supplemented with 25 mM TES buffer and 100 μM of either (+)-catechin hydrate (Sigma-Aldrich, CAS# 225937-10-0) , catechol (Sigma-Aldrich, CAS# 120-80-9), dopamine hydrochloride (Sigma-Aldrich, CAS# 62-31-7), (−)-epinephrine (+)-bitartrate salt (Sigma-Aldrich, CAS# 51-42-3), hydroquinone (Sigma-Aldrich, CAS# 123-31-9), levodopa (Sigma-Aldrich, CAS# 59-92-7), norepinephrine bitartrate monohydrate (MCE, CAS# 108341-18-0), phenol (VWR, CAS# 108-95-2), (R)-(−)-phenylephrine hydrochloride (Sigma-Aldrich, CAS# 61-76-7), tartaric acid (Sigma-Aldrich, CAS# 87-69-4), or tyramine hydrochloride (Sigma-Aldrich, CAS# 60-19-5).

After four days of incubation at 30 °C, plates were overlaid with soft LB agar (1.8% w/v agar) containing one of the indicator strains (2% v/v) pre-grown in liquid LB to exponential phase (OD600 = 0.4 – 0.6) and incubated overnight at 37 °C (± 18 hours). The following day, antibacterial activity was determined as the ratio between the inhibition zone diameter and the spot diameter.

Bioactivity of crude extracts and metabolite profiling

Streptomyces sp. MBT84 was grown confluently on MM agar plates with and without 100 μM catechol (n = 6). After five days of growth, the agar plates were cut into small pieces, soaked overnight in ethyl acetate (EtOAc) to extract the metabolites, evaporated at room temperature, and dissolved in MeOH to a final concentration of 20 mg/mL. 10 μL was spotted onto a sterile filter disc placed onto a soft agar layer inoculated with *B. subtilis* 168. As controls, 10 μL MeOH and 6 μL of 1 mg/mL ampicillin solutions were used. Following incubation overnight at 37 °C (± 18 hours), the plates were examined visually for halos.

For liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis, the dry extracts were dissolved in MeOH to a final concentration of 1 mg/mL. LC-MS/MS acquisition was performed using Shimadzu Nexera X2 ultra high-performance liquid chromatography (UPLC) system, with attached photodiode array detector (PDA), coupled to Shimadzu 9030 QTOF mass spectrometer, equipped with a standard electrospray ionisation (ESI) source unit, in which a calibrant delivery system (CDS) is installed. A total of 2 μL was injected into a Waters Acquity HSS C_{18} column (1.8 µm, 100 Å, 2.1 × 100 mm). The column was maintained at 30 °C, and run at a flow rate of 0.5 mL/min, using 0.1% formic acid in H_2O, and 0.1% formic acid in acetonitrile (ACN) as solvents A and B, respectively. The gradient used was 5% B for 1 min, 5–85% B for 9 min, 85–100% B for 1 min, and 100% B for 4 min. The column was re-equilibrated to 5% B for 3 min before the next run was started. The PDA acquisition was performed in the range of 200-600 nm, at 4.2 Hz, with 1.2 nm slit width. The flow cell was maintained at 40 °C.

All the samples were analysed in positive polarity, using data dependent acquisition mode. In this regard, full scan MS spectra (m/z 100–1700, scan rate 10 Hz, ID enabled) were followed by two data dependent MS/MS spectra (m/z 100–1700, scan rate 10 Hz, ID disabled) for the two most intense ions per scan. The ions were fragmented using collision induced dissociation (CID) with fixed collision energy (CE 20 eV), and excluded for 1 s before being re-selected for fragmentation. The parameters used for the ESI source were: interface voltage 4 kV, interface temperature 300 °C, nebulizing gas flow 3 L/min, and drying gas flow 10 L/min.

LC-MS based comparative metabolomics

Raw data obtained from LC-MS analysis were converted to mzXML centroid files using Shimadzu LabSolutions Postrun Analysis. The files were imported into Mzmine 2.53 for data processing [281]. Unless stated otherwise, *m/z* tolerance was set to 0.002 *m/z* or 10.0 ppm, RT tolerance was set to 0.05 min, noise level was set to 2.0E2 and the minimum absolute intensity was set to 5.0E2. Raw data were cropped to RT 0.5–12 min. Mass ion peaks were detected (positive polarity, mass detector: centroid) and their chromatograms were built using

ADAP chromatogram builder[282] (minimum group size in number of scans: 10; group intensity threshold: 2.0E2). The detected peaks were smoothed (filter width: 9), and the chromatograms were deconvoluted (algorithm: local minimum search; chromatographic threshold: 85%; search minimum in RT range: 0.05; minimum relative height: 1%; minimum ratio of peak top/edge: 2; peak duration: 0.03–2.00 min). The detected peaks were deisotoped (monotonic shape; maximum charge: 2; representative isotope: most intense). Peak lists from different extracts were aligned (weight for RT = weight for *m/z* = 20; compare isotopic pattern with a minimum score of 50%). Missing peaks detected in at least one of the samples were filled with the gap filling algorithm (RT tolerance: 0.1 min). Among the peaks, we identified fragments (maximum fragment peak height: 50%), adducts ($[M+Na]^+$, $[M+K]^+$, $[M+NH_4]^+$, maximum relative adduct peak height: 3000%) and complexes (Ionization method: $[M+H]^+$, maximum complex height: 50%). Duplicate peaks were filtered. Artifacts caused by detector ringing were removed (*m/z* tolerance: 1.0 *m/z* or 1000.0 ppm). The aligned peaks were exported to a MetaboAnalyst file.

In Excel, features that were not consistently present with an intensity higher than 8000 in all samples were removed from the file. Additionally, all features that originate from the culture medium were removed by retaining only features with an average peak intensity of at least 50 times greater in the bacterial extracts than in the culture medium extracts. The resulting peak list was uploaded to MetaboAnalyst [271] for statistical analysis. Log transformation with pareto scaling was applied to the data. Differences with a twofold change and an FDR-adjusted *p*-value < 0.1 were considered statistically significant. Based on these criteria, volcano plots and heat maps were generated.

MS/MS-based molecular networking and dereplication

MS/MS raw data (obtained from Shimadzu 9030 QTOF MS) were converted to a 32-bit mzML file using MSConvert (ProteoWizard) and a molecular network was assembled using the online Global Natural Product Social Molecular Networking (GNPS) tool [220]. Both the precursor ion and the MS/MS fragment ion mass tolerance were set to 0.02 Da. The minimum cosine score was set to 0.7 and the minimum matched peaks set to 5. The MSCluster algorithm was run with a minimum cluster size of 2 spectra. The spectra in the network were searched against the GNPS spectral libraries. For this, the precursor ion and the MS/MS fragment ion mass tolerance were set to 0.5 Da. Matches between network spectra and library spectra required a minimum score of 0.7 and at least 6 matched peaks. Cytoscape 3.5.1 was used for visualization of the generated molecular networks. The edge thickness was set to represent the cosine score, with thicker lines indicating higher similarity between nodes. LC-MS/MS data were deposited in the MassIVE Public GNPS data set (MSV000087784). The molecular networking job in GNPS can be found at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=78cfa392f9c94ac0a110dc682a2d8e6f.

The annotated MS/MS spectra were deposited in the GNPS spectral library for galtamycin C (CCMSLIB00006675753) and the new compound **8** (CCMSLIB00006675754).

At-line nanofractionation and bioactivity assessment

Liquid chromatography separation, subsequent at-line nanofractionation and parallel mass spectrometry analysis were performed as previously described [272]. Each sample was measured in duplicate. The extracts were dissolved in MeOH to a final concentration of 4 mg/mL, and 10 µL was injected into Waters Acquity UPLC system with attached PDA, equipped with XBridge Peptide BEH C_{18} column (5 μm, 300 Å, 4.6 × 100 mm). The flow rate used was 0.6 mL/min. Solvent A was 0.1% formic acid in H_2O and solvent B was 0.1% formic acid in ACN. The gradient used was 0–50% B for 20 min, 50–90% for 5 min, 90% for 5min, 90–100% for 1 min and 100% for 10 min. After the flow was split in a 1:9 ratio, the smaller fraction was fed to a Thermo Instruments MS system (LTQ Orbitrap XL, Bremen, Germany) equipped with an ESI source. The following ESI parameters were used: capillary voltage 5 V, spray voltage 3.5 kV, capillary temperature 300 °C, auxiliary gas flow rate 10 arbitrary units, and sheath gas flow rate 50 arbitrary units. MS spectra were acquired in the Orbitrap in positive mode at a mass range of 100-2000 *m/z*, and FT resolution of 30,000. The larger fraction eluted from the UPLC was sent to the chip-based nano-electrospray ionization source/fractionation robot (NanoMate Triversa, Advion BioSciences). The delay was determined as 0.3 min by injection 1 μL of 0.5 mg/mL ampicillin dissolved in water. Dried plates with the nanofractionated extract were tested for antibacterial activity in a resazurin reduction assay against *B. subtilis* 168 following the method described by Mladic et al. (2018) [272]. The bioactivity chromatograms were plotted in GraphPad Prism 9 software.

Large-scale fermentation and isolation of metabolites 7 and 8

Streptomyces sp. MBT84 was grown on MM agar plates supplemented with 50 μM catechol, 1% glycerol and 0.5% mannitol at 30 °C for five days. Agar plates were cut into small pieces and soaked in EtOAc to extract metabolites as described earlier. The solvent was subsequently evaporated under reduced pressure at 40 °C to obtain 1.6 g crude extract. This extract was adsorbed onto 1.6 g silica gel (pore size 60 Å, 70-230 mesh, Sigma Aldrich), and loaded on a silica column, followed by gradient elution using mixtures of n-hexane, EtOAc, and MeOH. One of the fractions that eluted with 50% EtOAc : 50% n-hexane was combined with the fraction that eluted with 75% EtOAc : 25% n-hexane, and reconstituted in acetonitrile. This fraction was further purified using a Waters preparative HPLC system comprised of 1525 pump, 2707 autosampler, and 2998 PDA detector. The pooled fraction was injected into a SunFire C_{18} column (10 μm, 100 Å, 19 × 150 mm). The column was run at a flow rate of 12.0 mL/min, using solvent A (dH_2O) and solvent B (acetonitrile), and a gradient of 70-100% B over 20 min to yield **7** (3.4 mg) and **8** (1.9 mg).

Galtamycin D (8): yellow amorphous powder; UV (LC-MS) λmax 222, 266, and 441 nm; HRESIMS *m/z* 579.1866 $[M+H]^+$ (calcd for $C_{31}H_{31}O_{11}$, 579.1861); ^1H and ^{13}C NMR data (Table S5).

NMR measurements

NMR measurements for the purified compounds were recorded on Bruker Ascend 850 NMR spectrometer (Bruker BioSpin GmbH), equipped with a 5 mm cryoprobe. The samples were measured in DMSO-d_6 in a 3 mm NMR tube through the use of an adapter. The spectra were referenced using the solvent residual peak and processed in MestReNova software.

Genome sequencing, assembly, and annotation

MBT84 was grown in 25 mL of YEME supplemented with 0.5% glycine and 5 mM $MgCl_2$ and cultivated at 30 °*C* with 200 rpm shaking speed. Genomic DNA was isolated by phenol-chloroform extraction as described previously[227]. Genome sequencing was performed using Pacbio Sequel RSII at DNA link Sequencing Lab – South Korea. Raw sequences were demultiplexed with Lima 1.9.0 to produce CCS reads which were converted to fastq with bam2fastq 1.3.0. The assembly was performed with Flye 2.5 [261]. The genome has been deposited at GenBank under accession number JAHTGP000000000. Biosynthetic gene clusters were identified using the genome mining tool antiSMASH 6.0 [217].

Proteomics sample preparation

MM agar plates with and without 100 μM catechol were covered with cellophane. MBT84 spores were spread on the plates using glass beads and incubated for five days at 30 °*C*. Biomass was scraped off and snap-frozen in liquid nitrogen, lysed in a precooled TissueLyser adaptor (Qiagen, The Netherlands) and proteins extracted using lysis buffer [4% SDS, 100 mM tris-HCl (pH 7.6), 50 mM EDTA]. Protein extracts were prepared as described in Chapter 4

Proteomics of catechol-grown cultures (proteomining)

For the analysis of the expression of biosynthetic enzymes in response to catechol, we applied natural product proteomining [249]. Samples were prepared from *Streptomyces* surface-grown cultures grown with or without catechol. Desalted peptide solutions were injected into Waters nanoAcquity UPLC system equipped with a Waters HSS T3 C_{18} (1.8 µm, 100 Å, 75 µm X 250 mm). A gradient from 1% to 40% acetonitrile in water (with added 0.1% FA) over 110 min was applied. Online MS/MS analysis was done using a Waters Synapt G2-Si HDMS mass spectrometer with a UDMS^E method set up as described previously [283]. [Glu1]-fibrinopeptide B was used as a lock mass compound and sampled every 30 s. Raw data from all samples were first analysed using the vender software ProteinLynx

Global SERVER (PLGS, version 3.0.3, waters, USA). The resulting dataset was imported in ISOQuant version 1.8 [283] for label-free quantification. TOP3 quantification was filtered to remove identifications that had an average value of less than 4000, in all replicates. This led to the removal of 152 protein quantification results. Log_2 fold changes were calculated, and proteins were considered significantly altered in expression when FDR-adjusted $p <$ 0.1 were obtained. Volcano plots were made from filtered data, with the biosynthetic gene clusters color-coded. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [265] partner repository with the dataset identifier PXD030319.

Proteomics of *Streptomyces* sp. MBT84 expressing catechol dioxygenases

Heterologous expression of catechol dioxygenases

Constructs for the heterologous expression of catechol 1,2-dioxygenase (C12O) and catechol 2,3-dioxygenase (C23O) were constructed as follows: The *catA1* gene encoding C12O was amplified from the genomic DNA of *Pseudomonas putida* KT2440 using the primer pair SF14_catA1_F and catA1_T0_R. The *xylE* gene encoding C23O was amplified from promoter-probe vector pIJ4083 using the primer pair SF14_xylE_F and xylE_T0_R. The forward primers contain the sequence of the strong constitutive promoter SF14[279] including RBS and the reverse primers contain a t0 terminator sequence. The PCR products SF14-*catA1*-T0 and SF14-*xylE*-T0 were placed into the BamHI / XbaI site of the multi-copy vector pWHM3 to create vectors pGWS1519 and pGWS1520 respectively. These constructs were introduced into *Streptomyces sp.* MBT84 via protoplast transformation, whereby the empty vector was used as a control. An overview of the constructs and oligonucleotides is presented in Table S1 and Table S2. For all experiments, the growth medium was supplemented with 10 μg/mL thiostrepton.

Analysis of the recombinant strains by quantitative proteomics

Quantitative proteomics was used to analyse the expression of catechol dioxygenases and the expression profile of BGC1.4 in recombinant strains of *Streptomyces* sp. MBT84 as described previously [216]. Briefly, the desalted peptide solution was separated on an UltiMate 3000 RSLCnano system (Thermo Scientific) set in a trap-elute configuration, coupled to QExactive HF (Thermo Scientific) mass spectrometer. The LC system used a Waters nanoEase M/Z Symmetry C_{18} trap column (5 μm, 100 Å, 180 μm × 20 mm) for peptide loading/retention, and Waters nanoEase M/Z HSS T3 C_{18} analytical column (1.8 μm, 100 Å, 75 μm × 250 mm) for peptide separation. The MS was operated in positive mode with data dependent acquisition. Raw LC-MS/MS files were analysed using MaxQuant software (v1.6.17.0)[264] with label

free quantification (LFQ) method applied. Proteins were considered significantly altered in expression when FDR-adjusted $p < 0.1$ were obtained. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [265] partner repository with the dataset identifier PXD030484.

CRISPRi technology

To knock down the expression of genes of the angucycline BGC, we applied CRISPRi RNA interference technology. The CRISPRi system was modified from pCRISPR-dCas9 [284] by expressing Cas9 from the constitutive *gapdh* promoter, using vector pSET152 that integrates at the øC31 attachment site on the *S. coelicolor* chromosome [285]. For this, a 20 nt spacer sequence was introduced into the sgRNA scaffold by PCR using forward primers KSβ_TF or KSβ_NT2F together with reverse primer SgTermi_R_B. The PCR products were cloned into pGWS1370 [286] via NcoI / BamHI restriction sites. The resulting constructs pGWS1516 (targeting template strand of KSβ, control) and pGWS1517 (targeting non-template strand of KSβ) were then introduced into *Streptomyces sp.* MBT84 via protoplast transformation[227]. An overview of the constructs and oligonucleotides is presented in Table S1 and S2.

Supplementary information for Chapter 5

Table S1. Overview of the primers used in this study

Primer name	sequence
SgTermi_R_B	CTAGGGATCCCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTAGTTACGCCTACGTAAAAAAAGCACCGACTCGGTGCC
KSβ_TF	CATGCCATGGGTTCGAGATGGGTGTGATCAGTTTTAGAGCTAGAAATAGC
KSβ_NT2F	CATGCCATGGGATGTCGCGGGGCAGTGTCCGTTTTAGAGCTAGAAATAGC
SF14_catA1_F	CATGGGATCCTAATGAGTTACGTAGACCTACGCCTTGACCTTGATGAGGCGGCGTGAGCTACAATCAATACTCGATTAGAATTCAAGGGAGAGAACATGACCGTGAAAATTTCCCAC
catA1_T0_R	CATGTCTAGATGGACTCACAAAGAAAAAACGCCCGGTGTGCAAGACCGAGCGTTCTGAACAATCAGCCCTCCTGCAACGCCCG
SF14_xylE_F	CATGGGATCCTAATGAGTTACGTAGACCTACGCCTTGACCTTGATGAGGCGGCGTGAGCTACAATCAATACTCGATTAGAATTCAAGGGAGAGAACATGAACAAAGGTGTAATGCGA
xylE_T0_R	CATGTCTAGATGGACTCACAAAGAAAAAACGCCCGGTGTGCAAGACCGAGCGTTCTGAACAATCAGGTCAGCACGGTCATGAA

Table S2. Overview of the plasmids used in this study

Plasmid	Description	Reference
pWHM3	Unstable *E. coli* / *Streptomyces* shuttle vector with high copy number	[287]
GWS1370	pGWS1369 (pSET152 lacking its NcoI site) containing sgRNA scaffold (no spacer) and Pgapdh-dCas9	[286]
GWS1516	GWS1370 containing a spacer targeting the template strand of KSβ	This work
GWS1517	GWS1370 containing a spacer targeting the non-template strand of KSβ	This work
GWS1519	pWHM3 containing *catA1* behind SF14 promoter and a t0 terminator	This work
GWS1520	pWHM3 containing *xylE* behind SF14 promoter and a t0 terminator	This work

Table S3. Genomic features of *Streptomyces* sp. MBT84

	***Streptomyces* sp. MBT84**
Number of contigs	3
Largest contig	9,591,886
Total length	10,344,466
N50	9,591,886
CDS	9,713
rRNAs	18
tRNAs	90

Table S4. Biosynthetic gene clusters identified by antiSMASH 6.0 [217]

Protocluster	Type	Most similar known cluster
Contig_1		
1	RiPP-like	Informatipeptin (57%)
2	NAPAA	-
3 *	terpene	Hopene (92%)
4 *	T2PKS, oligosaccharide	Saquayamycin A (87%)
5	T3PKS	-
6	siderophore	Grincamycin (8%)
7 *	terpene	Geosmin (100%)
8	RiPP-like	-
9 *	siderophore	-
10	terpene	Albaflavenone (100%)
11 *	T2PKS	Spore pigment (83%)
12	siderophore	Desferrioxamine B (66%)
13	RiPP-like	-
14	ectoine	Ectoine (100%)
15	NAPAA	Chalcomycin (9%)
16	T3PKS	Herboxidiene (8%)
17	RRE-containing	
18	NRPS, betalactone	Cyclomarin D (17%)
19	Melanin	Melanin (42%)
20	NRPS, NAPAA, RiPP-like	Stenothricin (18%)
Contig_2		
21	Bacteriocin	-
Contig_3		
-		

* Expressed BGCs: ≥ 1 core biosynthetic protein was identified

Table S5. Comparison of BGC4 to saquayamycin cluster

	Protein locus tag	Size (aa)	Putative function	Homologue in saquayamycin BGC0001769	%identity*
1	_08725	534	Multidrug-efflux transporter	sqnZ	98.50
2	_08730	246	Hypothetical protein	sqnAA	61.02
3	_08735	372	NADH oxidase	Not found	-
4	_08740	526	Multidrug resistance protein	sqnB	99.24
5	_08745	199	FMN-dependent NADPH-azoreductase	sqnC	100.00
6	_08750	227	Tetracycline repressor protein class E	sqnD	96.38
7	_08755	283	Hypothetical protein	sqnE	97.88
8	_08760	492	Anhydrotetracycline monooxygenase	sqnF	98.98
9	_08765	602	Tetracenomycin F2 cyclase	sqnBB	99.07
10	_08770	426	polyketide putative beta-ketoacyl synthase 1	sqnH	99.28
11	_08775	408	polyketide putative beta-ketoacyl synthase 2	sqnI	98.77
12	_08780	89	polyketide synthase acyl carrier protein	sqnJ	98.88
13	_08785	261	Putative ketoacyl reductase	sqnK	100.00
14	_08790	311	Putative polyketide cyclase	sqnL	99.04
15	_08795	665	Anhydrotetracycline monooxygenase	sqnM	98.95
16	_08800	404	putative MFS-type transporter EfpA	sqnN	99.26
17	_08805	430	Aclacinomycin-T 2-deoxy-L-fucose transferase	sqnG1	99.53
18	_08810	404	L-noviosyl transferase	sqnG2	99.26
19	_08815	193	dTDP-4-dehydrorhamnose 3,5-epimerase	sqnS1	100.00
20	_08820	376	L-noviosyl transferase	sqnG3	99.47
21	_08825	355	Glucose-1-phosphate thymidylyltransferase	sqnS2	99.72
22	_08830	327	dTDP-glucose 4,6-dehydratase	sqnS3	99.08
23	_08835	353	dTDP-6-deoxy-L-talose 4-dehydrogenase (NAD(P)(+))	sqnS4	97.79
24	_08840	434	GDP-perosamine synthase	sqnS5	99.54
25	_08845	254	dTDP-glucose 4,6-dehydratase	sqnS6	98.82
26	_08850	323	Hypothetical protein	sqnO	98.45
27	_08855	466	Hypothetical protein	sqnS7	99.57
28	_08860	318	Glucose--fructose oxidoreductase	sqnS8	98.11
29	_08865	241	4'-phosphopantetheinyl transferase psf-1	sqnCC	97.88
30	_08870	530	putative propionyl-CoA carboxylase beta chain 5	sqnP	99.62
31	_08875	221	Hypothetical protein	sqnDD	96.82
32	_08880	528	Aclacinomycin-N/aclacinomycin-A oxidase	sqnQ	99.81
33	_08885	145	Hypothetical protein	sqnEE	99.31
34	_08890	76	Hypothetical protein	Not found	-
35	_08895	254	Transcriptional regulatory protein YycF	sqnR	100.00
36	_08900	639	2-oxoglutarate oxidoreductase subunit KorA	sqnT	99.51
37	_08905	364	2-oxoglutarate oxidoreductase subunit KorB	sqnU	99.16

* Sequences were aligned using protein blast

Table S6. ^{1}H and ^{13}C NMR data of 8 in DMSO-*d*6 at 298 K

Position	δC, type	δH, mult. (*J* in Hz)
1	195.8, C	
2	50.4, CH_2	3.01, d (15.7) 2.82, dd (15.7, 2.6)
3	75.8, C	
4	33.4, CH_2	3.57, dd (18.1, 2.4) 2.94, dd (18.1, 2.9)
4a	138.3, C	
5	136.7, C	
6	114.8, CH	8.16, s
6a	130.3, C	
7	187.3, C	
7a	115.6, C	
8	158.1, C	
9	137.6, C	
10	133.8, CH	7.95, d (8.0)
11	119.2, CH	7.84, d (8.0)
11a	131.7, C	
12	187.4, C	
12a	117.1, C	
12b	160.5, C	
13	24.9, CH_3	1.45, s
8-OH		12.92, s
12b-OH		12.89, s
Sugar A, β-D-olivose		
1A	70.5, CH	5.01, dd (11.2, 2.2)
2A	35.9, CH_2	2.26, m 1.62, m
3A	75.7, CH	3.88, m
4A	73.6, CH	3.52, dd (9.1, 9.1)
5A	73.5, CH	3.62, m
6A	17.4, CH_3	1.27, d (6.1)
Sugar B, α-l-cinerulose B		
1B	90.5, CH	5.23, d (2.7)
2B	70.9, CH	4.35, m
3B	39.8, CH_2	2.90, dd (17.4, 2.8) 2.48, dd (17.4, 3.7)
4B	208.7, C	
5B	77.0, CH	4.71, q (6.8)
6B	16.1, CH_3	1.24, d (6.8)

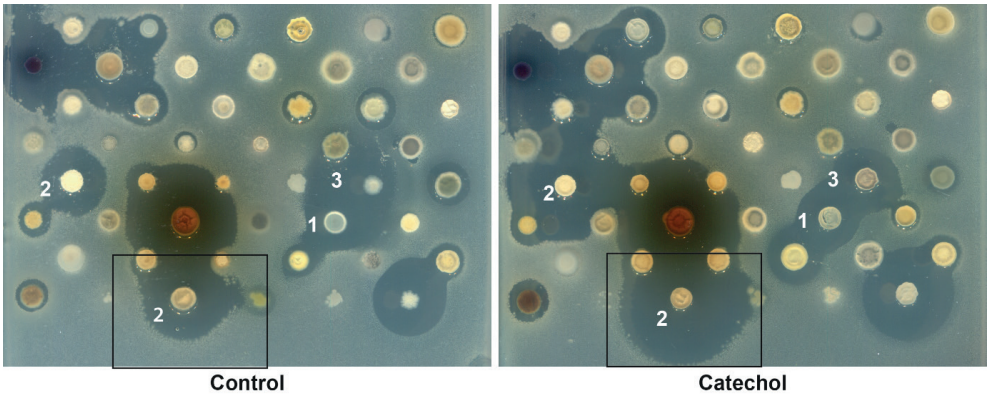


Figure S1. Catechol changes the bioactivity profile of different *Streptomyces*. *Streptomyces* sp. were spotted on MM with and without 100 μM catechol. Both inhibition (1) and promotion (2) of antibiotic production were observed. Additionally, we observed that catechol influenced interactions between different strains illustrated by the change in halo shape (3). Catechol significantly enhanced the bioactivity of *Streptomyces* sp. MBT84 (black square) (n = 3).

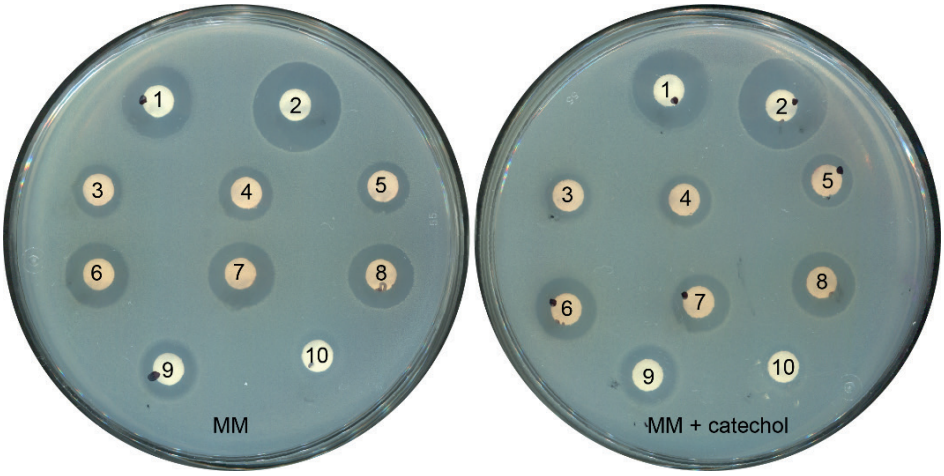


Figure S2. Addition of catechol to growth medium does not affect the susceptibility of *B. subtilis* 168 to different antibiotics and crude extracts of MBT84. MM glycerol mannitol agar plates with and without 100 μM catechol were overlaid with LB soft agar containing *B. subtilis* 168. Diffusion discs contain: 10 μL of (1) 0.5 mg/mL apramycin, (2) 0.5 mg/mL ampicillin, (3-5) crude extracts of MBT84 grown without catechol (20 mg/mL), (6-8) crude extracts of MBT84 grown in presence of catechol (20 mg/mL), (9) 1.5 mg/mL kanamycin, and (10) 5 mg/mL chloramphenicol. Experiment was conducted in duplicate with 2 replicates. No differences between the growth conditions were observed.

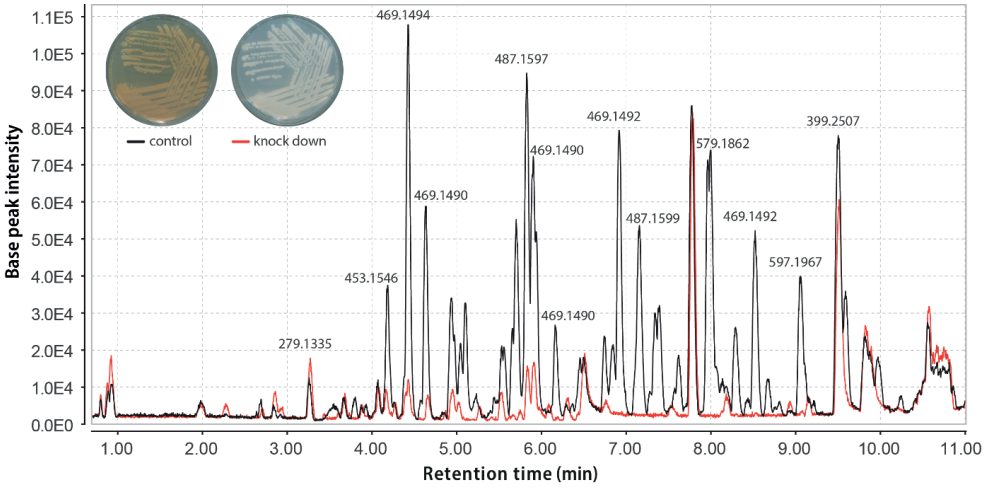


Figure S3. BGC4 specifies the angucyclines glycosides produced by *Streptomyces* sp. MBT84. Knock-down of the gene encoding KSβ of BGC4 in *Streptomyces* sp. MBT84 using CRISPRi resulted in almost complete inhibition of pigment production compared to the control strain harboring a construct that targets the template strand. The LC-MS chromatogram overlay of the crude extracts of the knock-down and control strains shows that inhibition of the expression of the KSβ resulted in decreased production of the majority of the metabolites. Three independent transformants were tested.

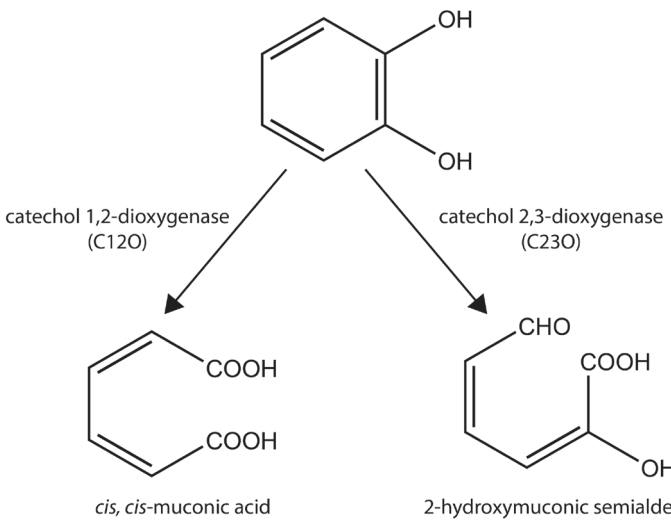


Figure S4. Degradation of catechol by catechol 1,2-dioxygenase and catechol 2,3-dioxygenase

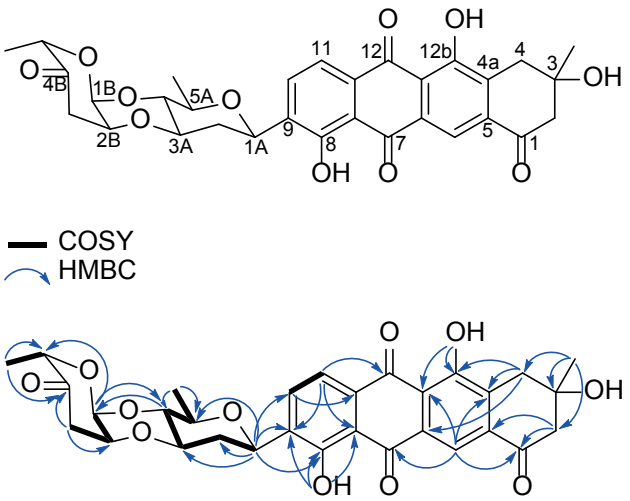


Figure S5. 2D structure (Top) and key COSY and HMBC correlations (bottom) of compound 8

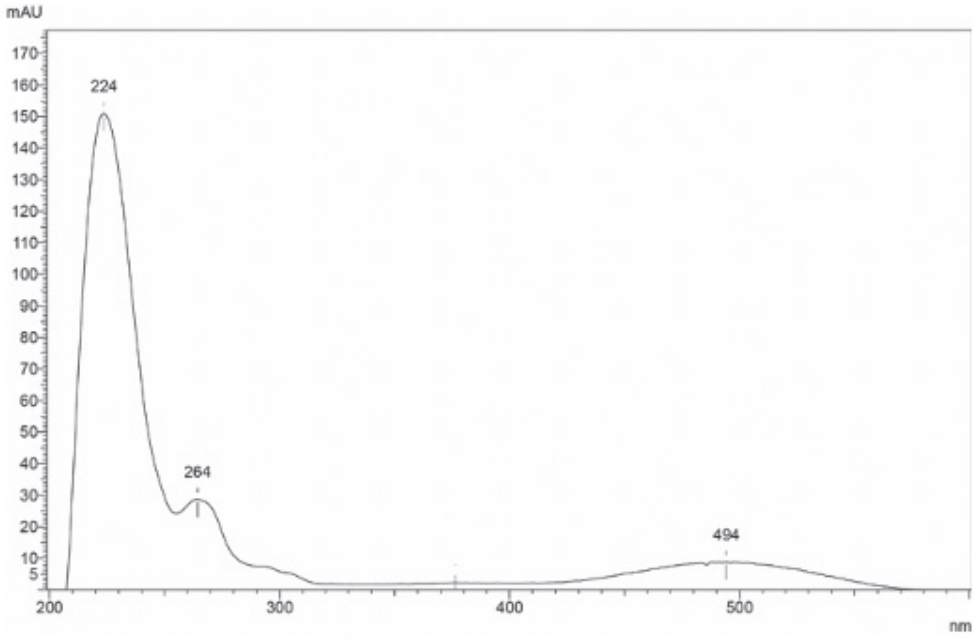


Figure S6. UV spectrum of 7

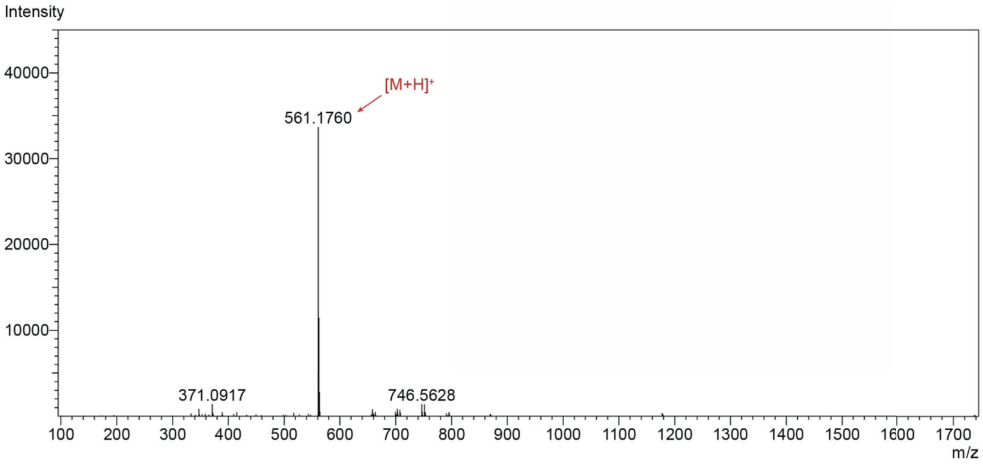


Figure S7. (+)-HR-ESI-MS spectrum of 7

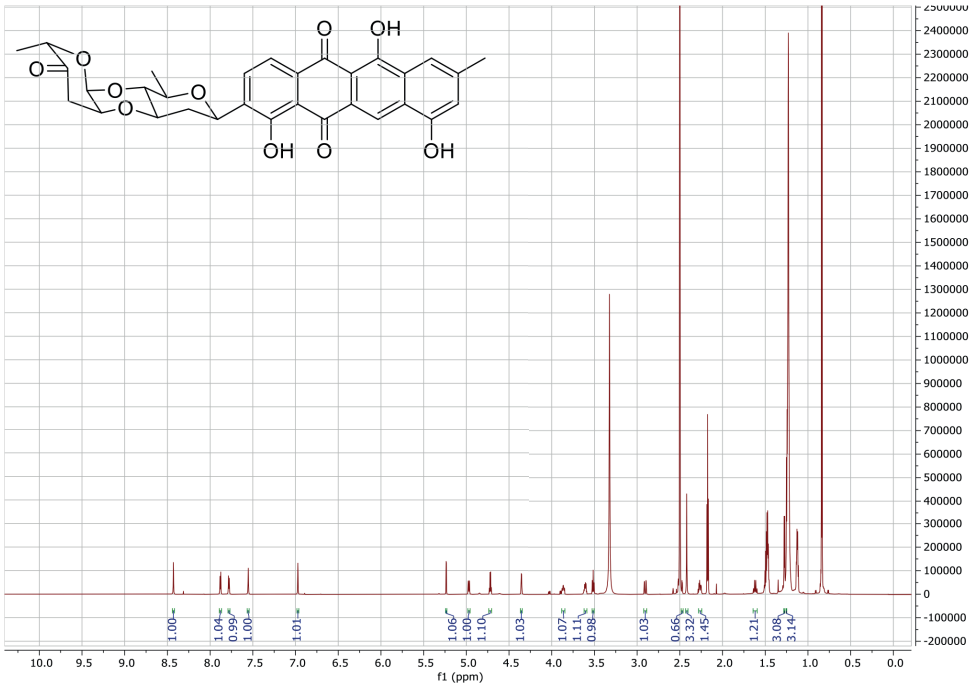


Figure S8. ^{1}H NMR spectrum of 7 (850 MHz, in DMSO-d_6)

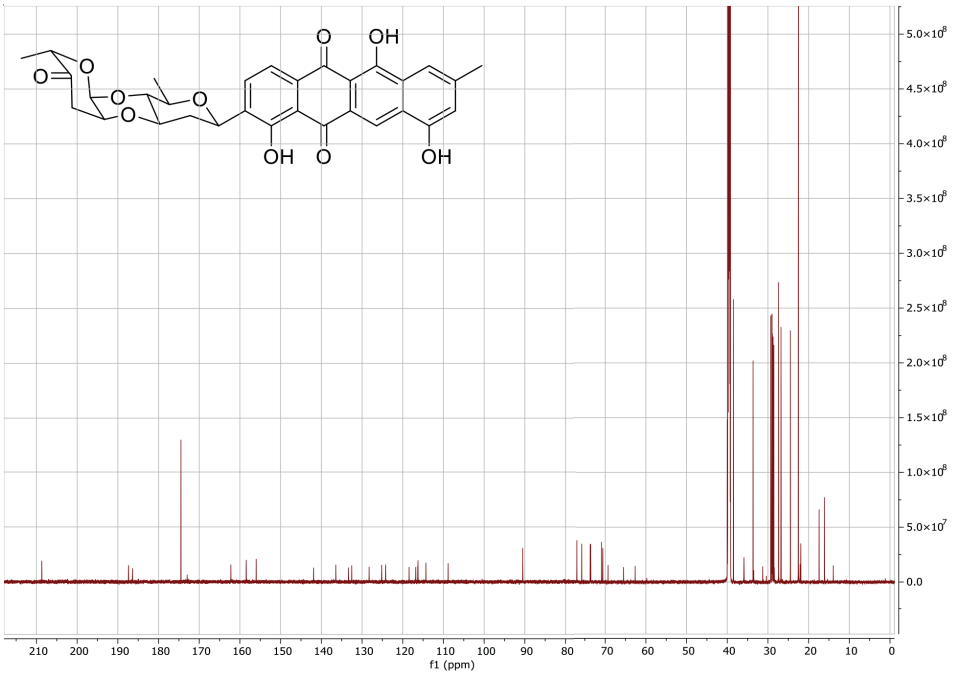


Figure S9. ^{13}C NMR spectrum of 7 (213 MHz, in DMSO-d_6)

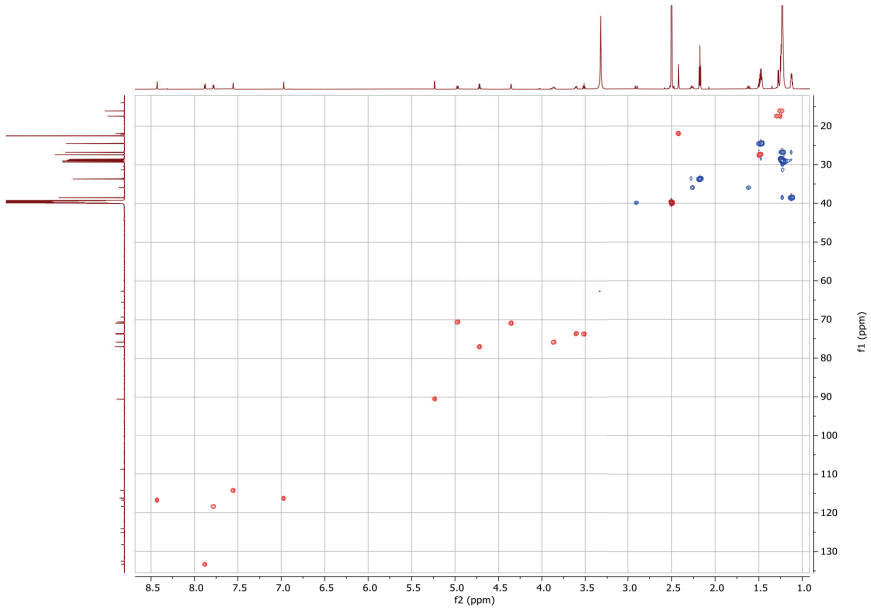


Figure S10. Multiplicity-edited HSQC spectrum of 7 (850 MHz, in DMSO-d_6)

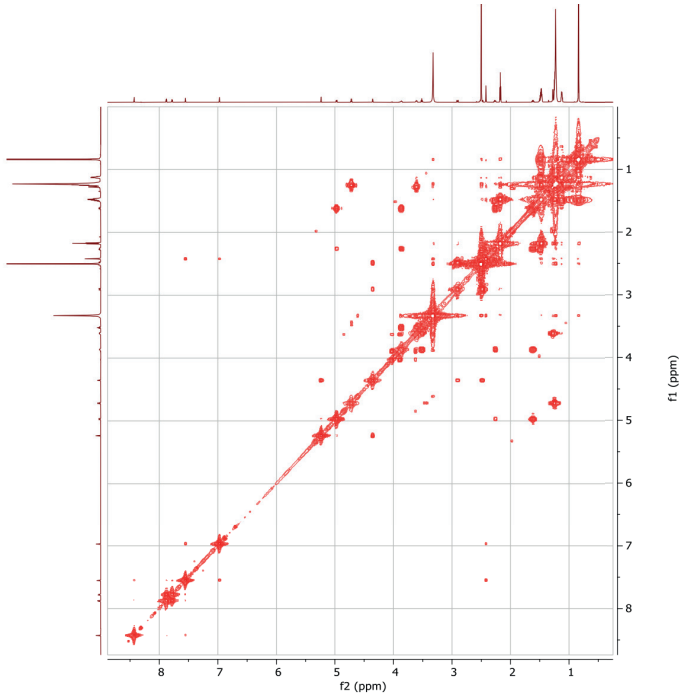


Figure S11. ^{1}H–^{1}H COSY spectrum of 7 (850 MHz, in DMSO-d_6)

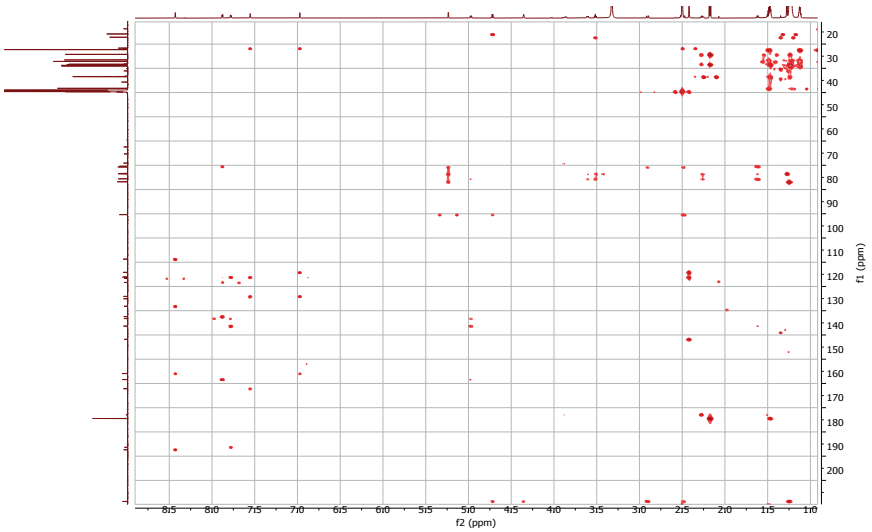


Figure S12. HMBC spectrum of 7 (850 MHz, in DMSO-d_6)

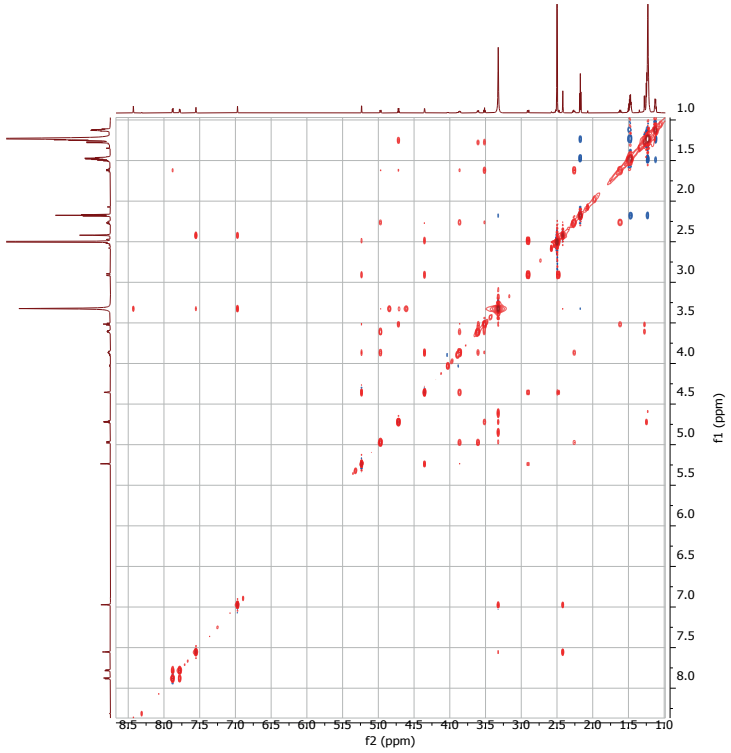


Figure S13. NOESY spectrum of 7 (850 MHz, in DMSO-d_6)

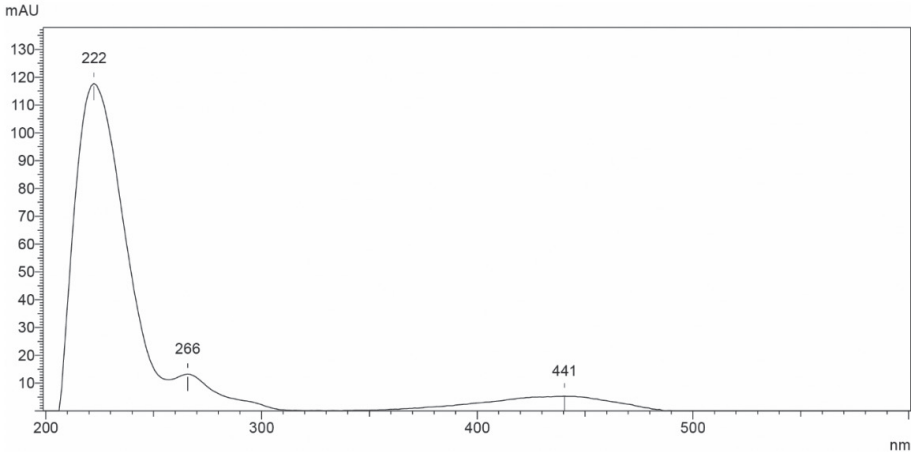


Figure S14. UV spectrum of 8

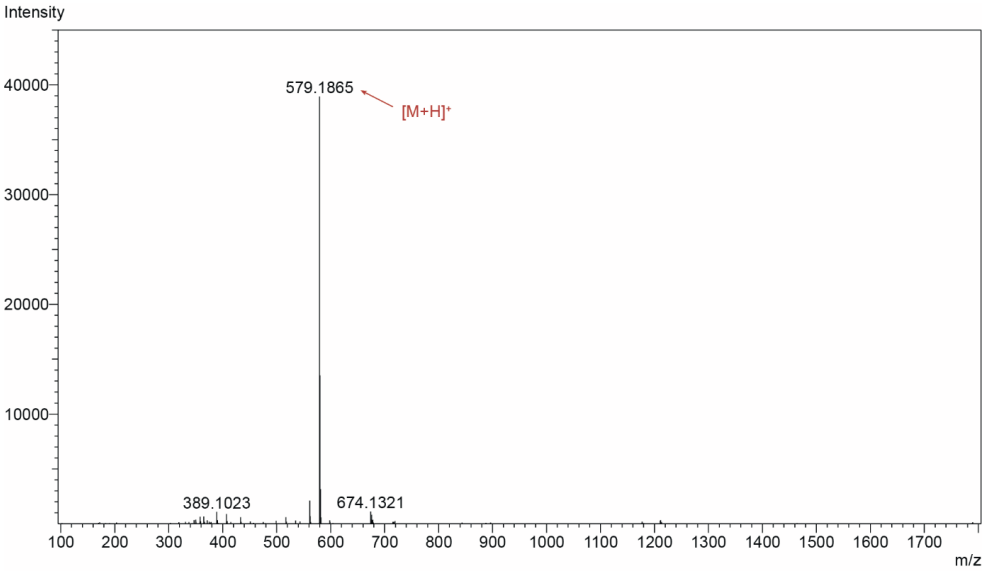


Figure S15. (+)-HR-ESI-MS spectrum of 8

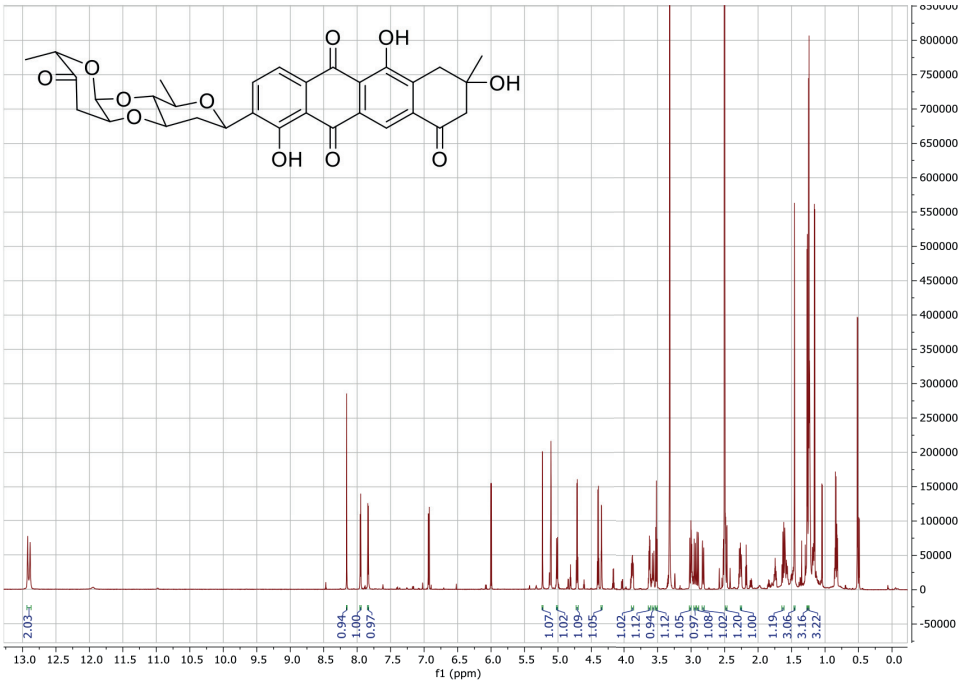


Figure S16. ^{1}H NMR spectrum of 8 (850 MHz, in DMSO-d_6)

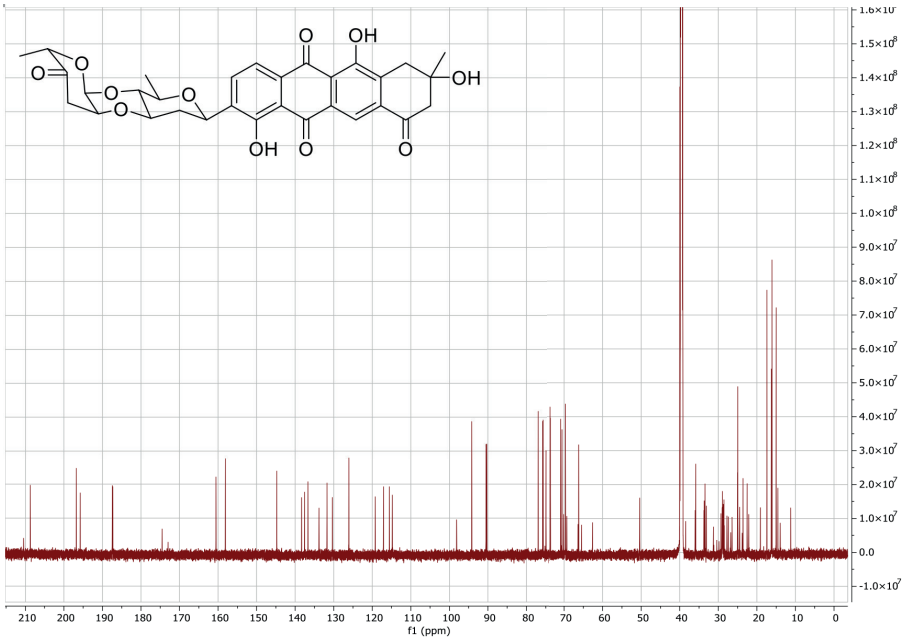


Figure S17. ^{13}C NMR spectrum of 8 (213 MHz, in DMSO-d_6)

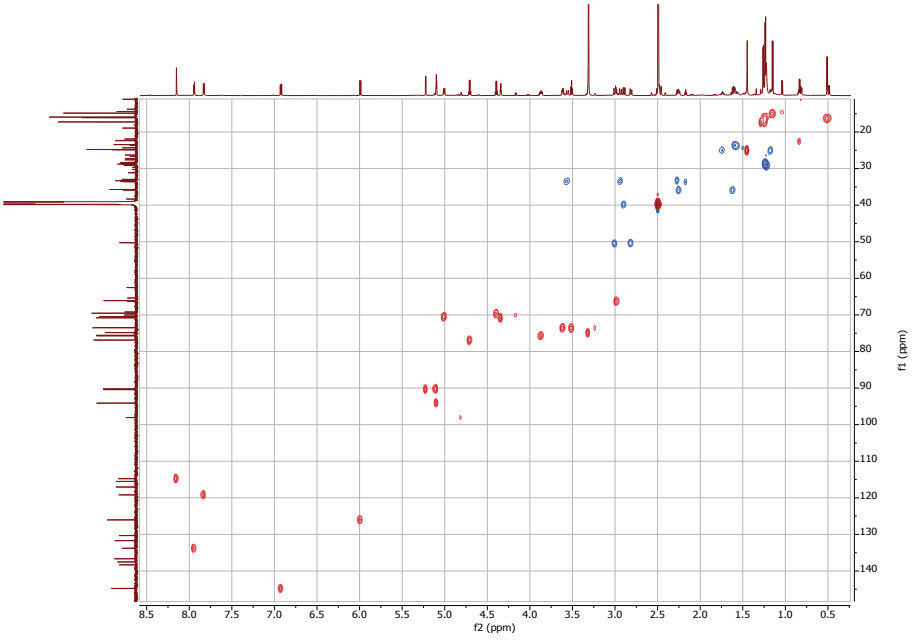


Figure S18. Multiplicity-edited HSQC spectrum of 8 (850 MHz, in DMSO-d_6)

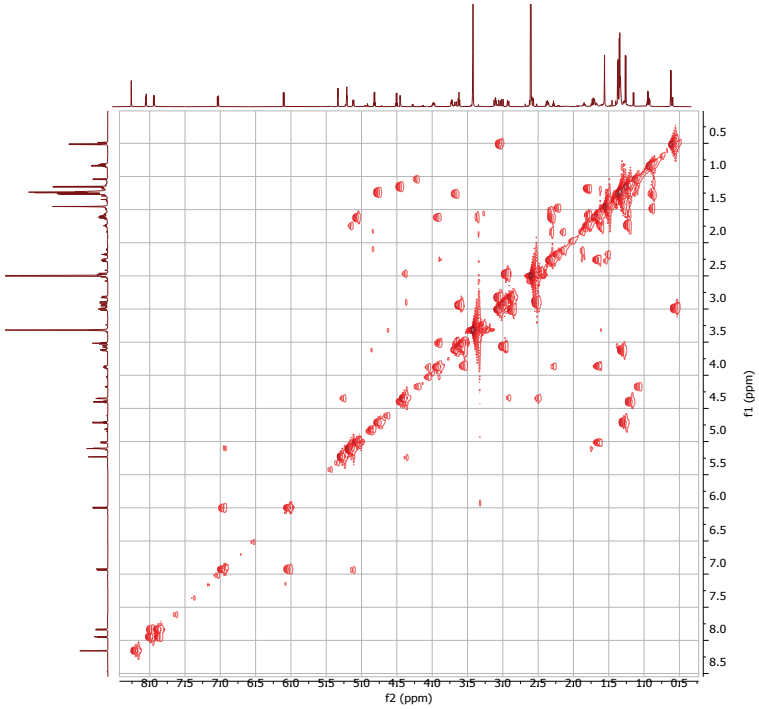


Figure S19. ^{1}H–^{1}H COSY spectrum of 8 (850 MHz, in DMSO-d_6)

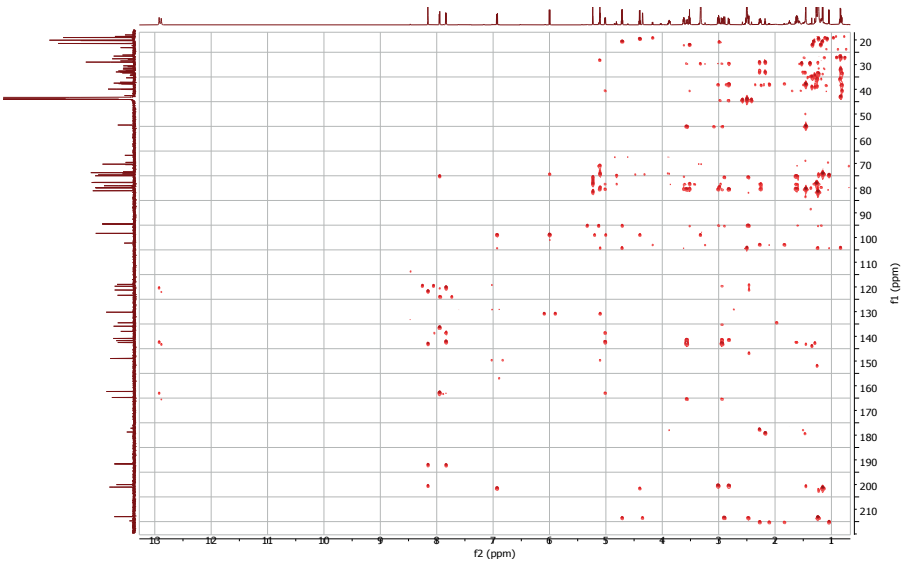


Figure S20. HMBC spectrum of 8 (850 MHz, in DMSO-d_6)

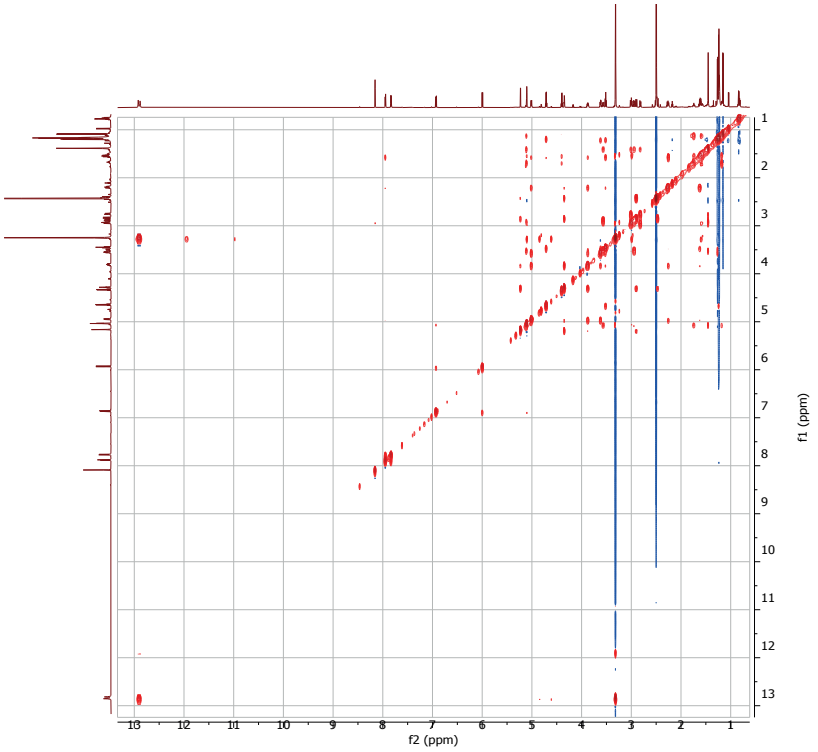


Figure S21. NOESY spectrum of 8 (850 MHz, in DMSO-d_6)