

## Supplementary Information

Toward safer and more sustainable by design biocatalytic amide-bond coupling

Elisabeth Söderberg <sup>a,b</sup>, Kerstin Felicitas von Borries <sup>c</sup>, Ulf Norinder <sup>d</sup>, Mark Petchey <sup>e</sup>, Ganapathy Ranjani <sup>a,b</sup>, Swapnil Chavan <sup>f</sup>, Hanna Holmquist <sup>g</sup>, Magnus Johansson <sup>h</sup>, Ian Cotgreave <sup>f</sup>, Martin A. Hayes <sup>e</sup>, Peter Fantke <sup>c</sup>, Per-Olof Syrén <sup>a,b</sup>

---

<sup>a</sup>School of Chemistry, Biotechnology and Health, Science for Life Laboratory, KTH Royal Institute of Technology, Tomtebodavägen 23 171 65, Solna, Sweden

<sup>b</sup>School of Engineering Sciences in Chemistry, Biotechnology and Health, Department of Fibre and Polymer Technology, KTH Royal Institute of Technology, 114 28 Stockholm, Sweden

<sup>c</sup>Quantitative Sustainability Assessment, Department of Environmental and Resource Engineering, Technical University of Denmark, Bygningstorvet 115, 2800 Kgs. Lyngby, Denmark

<sup>d</sup>Department of Computer and Systems Sciences, Stockholm University, 164 25 Kista, Sweden

<sup>e</sup>Discovery Sciences, BioPharmaR&D, AstraZeneca, 431 83 Mölndal, Sweden

<sup>f</sup>The Research Institute of Sweden RISE, Chemical and pharmaceutical toxicology, 151 36 Södertälje, Sweden

<sup>g</sup>IVL Swedish Environmental Research Institute, Aschebergsgatan 44, 411 33 Göteborg, Sweden

<sup>h</sup>Early CVRM Medicinal Chemistry, BioPharmaR&D, AstraZeneca, 431 83 Mölndal, Sweden

## Table of Contents

Supplementary Notes.....	4
Literature study ATP-dependent amide bond synthetases .....	4
USEtox average CF:s.....	4
Table S1. Average CF values for safechems and their corresponding amides.....	4
Table S2. Percentile of average CF values for the subset of safechems used in the experimental coupling compared to the percentile of the average CFs for the 16 acids and 17 amines with lowest CF values. ....	4
Methods .....	5
Protein expression and purification.....	5
Thermostability .....	5
UPLC-MS .....	5
Synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide.....	5
Synthesis of 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide and N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide .....	5
Table S3. Sequence identity (%) between McbA and its ancestors A1-A4.....	7
Table S4. Melting temperature ( $T_m$ ) of McbA and ancestors A1-A4 measured by nano differential scanning fluorimetry. ....	7
Table S5. Protein expression levels of McbA and ancestors. ....	7
Figure S1. Crude protein content of McbA and ancestor A1-A4.....	7
Figure S2. Thermal unfolding by McbA and ancestors A1-A4 .....	8
Figure S3. Phylogenetic tree of McbA used to compute ancestral sequences .....	8
Figure S4. Alignment used for the phylogenetic tree construction. ....	12
Figure S5. Alignment of McbA and ancestors A1-A4.....	12
Figure S6. $^1\text{H-NMR}$ of synthesised 3-Hydroxy-N-(3-phenylpropyl)benzamide (standard). ....	13
Figure S7. $^{13}\text{C-NMR}$ of synthesised 3-Hydroxy-N-(3-phenylpropyl)benzamide (standard). ....	13
Figure S8. $^1\text{H-NMR}$ of the enzymatic synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide (crude NMR). ....	14
Figure S9. $^{13}\text{C-NMR}$ of the enzymatic synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide (crude NMR). ....	14
Figure S10 $^{13}\text{C-NMR}$ (left) and $^1\text{H-NMR}$ (right) of 3-hydroxybenzoic acid, 3-phenyl-1-propylamine, and the enzymatic synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide (crude NMR).....	15
Figure S11. $^1\text{H-NMR}$ of synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide (standard). ....	16
Figure S12. $^{13}\text{C-NMR}$ of synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide (standard). ....	16
Figure S13. $^1\text{H-NMR}$ of biocatalytic synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide .....	17
Figure S14. $^{13}\text{C-NMR}$ of biocatalytic synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide .....	17
Figure S15. DEPT spectra of biocatalytic synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide .....	18
Figure S16. $^1\text{H-NMR}$ of synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide (standard). ....	19
Figure S17. $^{13}\text{C-NMR}$ of synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide (standard). ....	19
Figure S18 $^1\text{H-NMR}$ of biocatalytic synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide .....	20
Figure S19 $^{13}\text{C-NMR}$ of biocatalytic synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide .....	20
Figure S20. DEPT spectra of biocatalytic synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide .....	21
Figure S21. 305 nm UV absorption calibration curve of different 3-hydroxybenzoic acid concentrations. ....	22

Figure S22. HPLC-MS chromatogram of a sample taken from the upscaled biocatalytic synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide .....	22
Figure S23. HPLC-MS chromatogram of 3-Hydroxy-N-(3-phenylpropyl)benzamide standard.....	23
Figure S24. 305 nm UV absorption calibration curve of different 3-acetylbenzoic acid concentrations.....	23
.....	23
Figure S25. HPLC-MS chromatogram of a sample taken from the upscaled biocatalytic synthesis of 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide.....	24
Figure S26. HPLC-MS chromatogram of 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide standard.....	24
Figure S27. 305 nm UV absorption calibration curve of different benzoic acid concentrations.....	25
Figure S28. HPLC-MS chromatogram of a taken from the upscaled biocatalytic synthesis of N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide.....	25
Figure S29. HPLC-MS chromatogram of N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide standard.....	26
Table S6. Derivation of USEtox input parameters' mean and standard deviation from in-silico model prediction and uncertainty. ....	27
Table S7. Overview of uncertainty factors applied onto the model-specific uncertainty.....	28
Figure S30. USEtox characterization factors vs <i>in-silico</i> screening scores of safechems and subset of filtered out aromatic acids and amines. ....	28
Figure S31. Predicted toxicity and environmental score of detected amides from the experimental coupling considering only a subset of possible safechems in comparison to all possible amides from the whole safechem panel. ....	28
Figure S32. Toxic effects coupled to the molecular weight of amines and acids.....	29
Figure S33. Scatter plot of the USEtox CFs of amides with <i>in silico</i> hazard score equal to or lower than ten (figure 3) compared to the sum of the CFs of their constituent acid and amine moieties. ....	29
Table S8. Conversion by McbA (in %)..	30
Table S9. Conversion by A1 (in %). ....	31
Table S10. Conversion by A2 (in %).. ....	32
Table S11. List of safechem amines from <i>in silico</i> screening.....	33
Table S12. List of safechem acids from in silico screening.....	35
Table S13. The exact masses of detected and potential amides in the experimental coupling in figure 4. ....	37
Notes and references.....	38

## Supplementary Notes

### Literature study ATP-dependent amide bond synthetases

Many of the characterised ABS enzymes have been identified in genome mining projects of gene clusters responsible for the biosynthesis of various medically relevant secondary metabolites. NovL, (1) CloL, (2) SimL, (3) and CouL (4,5) are involved in the biosynthesis of aminocoumarin antibiotics novobiocin, clorobiocin, simocyclinone D8, and coumermycin A1 (figure 2A, panel C). Another group of ABS-reminiscent enzymes that share sequence motifs and similar reaction mechanisms to ABS-enzymes, are Ann1, (6) AsuD1 (7) and ORF33 (figure 2A, panels B, D, E) (8,9). These enzymes are involved in the synthesis of 2-amino-3-hydroxycyclopent-2-enone (C5N) and tailoring of polyketides. Recently, a group of CfaL enzymes responsible for the ligation of coronafacic acid (CFA) and coronamic acid (CMA) in the bacterial plant hormone-mimic phytotoxin coronatine (COR) was characterised and screened against different acyl and amine substrates (figure 2A, panel F). The enzymes accepted a broad range of both acyl and amine substrates; aliphatic, aryl, and heteroaryl carboxylic acids, and non-proteogenic amino acids. (10) The enzymes also managed to utilize carboxylic acids carrying other reactive functionalities, groups that would have needed to be protected if the substrate were to be used in chemical synthesis; again, emphasizing the potential of biocatalysts in amide bond synthesis.

### USEtox average CF:s

The average values of the CFs of the subset of arbitrarily selected safechems that were used in the experimental coupling did not differ much compared to the average values of all safechems (table S1). The average human CFs of the acids used were slightly lower than those of all safechem acids, while the opposite was true for the amines. For ecotoxicity, amines and acids in the subset were predicted slightly more toxic compared to all safechems. As for the amides made by the subset safechems, their average CFs did not differ much from the average of all amides made by safechems passing the established threshold (figure 3, table S1). The percentiles of the average CF values of the subset were between 44.3 to 55.4 %. The average CF percentile value of an equally large subset of safechem acids and amines could go as low as 19.9 and 4.7%, respectively (table S2). If we had implemented USEtox after the *in silico* filtering and before the experimental coupling, we could have reduced the human and ecotoxicity of the panel used in the experimental coupling even further.

Table S1. Average CF values for safechems and their corresponding amides. For reference, the average CFs are presented separately for all safechems, their responding amides and for the subset that was used in the experimental coupling, respectively. Amides (row 5) are the amides passing the established threshold in the *in silico* filtering (figure 3).

Type	Freshwater ecotoxicity (CF's)		Human toxicity (CF's)	
	Air	Freshwater	Air	Freshwater
Safechem acids	-1.34	0.34	-4.07	-4.48
Safechem acids used in the experimental coupling	-1.18	0.43	-4.15	-4.60
Safechem amines	-0.72	1.16	-3.97	-4.17
Safechem amines used in the experimental coupling	-0.67	1.15	-3.94	-4.13
All amides	-0.11	1.50	-3.02	-3.50
Amides from safechems used in the experimental coupling	0.01	1.64	-3.05	-3.53

Table S2. Percentile of average CF values for the subset of safechems used in the experimental coupling compared to the percentile of the average CFs for the 16 acids and 17 amines with lowest CF values. Data is shown for each of the four categories.

Type	Percentile of average CF			
	Freshwater ecotoxicity		Human toxicity	
Type	Air	Freshwater	Air	Freshwater
Subset of 16 acids used in the coupling	50.8	46.9	52.2	41.4
16 acids with the lowest CF value	19.9	23.9	26.4	30.8
Subset of 17 amines used in the coupling	47.4	44.3	54.4	46.4
17 amines with the lowest CF values	5.3	4.7	6.3	6.1

## Methods

### Protein expression and purification

The four consensus sequences were codon optimised for expression in *Escherichia Coli* by GeneArt (Thermo Fisher Scientific). Kanamycin resistance gene and N-terminal His-tag were included in all constructs, and each construct was delivered in pET-28a(+) plasmids. BL21(DE3) competent cells were transformed with approximately 2 ng plasmid, and selection was made by growing the cells on 30 µg/mL kanamycin lysogeny broth (LB) agar plates. For protein expression, overnight cultures grown in LB supplemented with 30 µg/mL kanamycin were added to LB with 30 µg/mL kanamycin to a start OD600 of 0.1. When the OD600 reached 0.6, 1 mM IPTG was added. McbA, A3, and A4 were set to grow overnight at 16 °C and 180 rpm, while A2 and A3 at 30 °C. The cells were harvested by centrifugation at 5,000 g for 20 min, and pellets were resuspended in 50 mM phosphate buffer (pH 7.5) with 10 % glycerol (w/v), 300 mM NaCl, and 40 mM imidazole. The cell suspensions were kept on ice while being sonicated at 30s x 3 bursts with 1.5 minute breaks using Branson Sonifier 450 (35% duty cycle and output control 5). Afterwards, the sonicated cell suspensions were centrifuged for 30 min at 4 °C and 14,000 g. The supernatants were injected into the ÄKTA explorer protein purification system (Cytiva) at 4 ml/min, equipped with a 5 mL HP-HisTrap column (Cytiva) equilibrated with ten column volumes of binding buffer. After washing the column with 20 column volumes of binding buffer, the proteins were eluted by a gradient going from 40 mM to 300 mM imidazole over 20 column volumes. Elution fractions were analysed by Invitrogen NuPAGE 4-12% Bis-Tris gel (Thermo Fisher Scientific), and desalted with PD10 columns (Sigma-Aldrich) equilibrated with 50 mM phosphate buffer (pH 7.5) and 300 mM NaCl. Concentration of protein samples was made with MacroSep 10 kDa cut-off centrifugation filters (Pall Laboratory). Protein concentration was measured with NanoDrop. 80% glycerol solution in pH 7.5 phosphate buffer with 300 mM NaCl was added to the protein samples to a final concentration of 10%. The proteins were aliquoted to 1.5 mL microcentrifuge tubes, snap-frozen in liquid nitrogen, and stored at -80 °C.

### Thermostability

The thermostability of the proteins was assessed by nano differential scanning fluorimetry using Prometheus NT.48 nanoDSF (NanoTemper Technologies). Samples of 2 mg/mL protein solution were prepared in glass capillaries. The fluorescence of the proteins at 330 and 350 nm was measured at a temperature gradient going from 20 to 95 °C at 1 °C/min. The derivate of the ratio between the two measured wavelengths over the temperature interval was used to determine the melting temperature of the proteins.

### UPLC-MS

The reactions were quenched with 30 µL acetonitrile supplemented with 1 mg/mL 1,2,3-tribromobenzene, and the plates were sealed by Velocity11's Plateloc. The plates were shaken in a short burst followed by centrifugation at 4000 g for 10 min. Waters Acquity UPLC equipped with a photodiode array detector, and a 3100 mass spectrometer was used for analysis. 3 µL samples were injected into an Acquity UPLC HSS C18 column (1.8 µm, 2.1 mm X 50 mm). The mobile phases were prepared by adding 0.5 mL of acidic solution (pH 3 solution consisting of 126.3 g water, 151.8 g formic acid, and 21.8 g ammonium hydroxide solution 25-30%) to 1 L of water (mobile phase A), and 1 mL of the acidic solution was added to 1 L 95% acetonitrile (mobile phase B). The mobile phase went from 10% to 99% B in 2 minutes. The conversions of the reactions were measured as the amide peak area percentage of the total amide and acid peak area using MassLynx (Waters).

### Synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide

3-hydroxybenzoic acid (0.511 g, 3.70 mmol) was dissolved in DCM (10 mL). To this, 1H-benzo[d][1,2,3]triazol-1-ol hydrate (0.623 g, 4.07 mmol) and 3-phenylpropan-1-amine (0.526 ml, 3.70 mmol) were added. Following this, diisopropylmethanediimine (0.573 ml, 3.70 mmol) was added to the mixture and the reaction was mixed for 4 h. Upon completion of the reaction, the diisopropylurea was filtered away and the DCM layer was washed with saturated bicarbonate solution (1x15 mL) and 2 M HCl (1x15 mL). The organic layer was dried over anhydrous sodium sulfate and the solvent was then removed under reduced pressure. The crude product was purified by flash column chromatography (Ethyl acetate: Heptane gradient 5 – 50% over 12 CV) to yield the desired product as a white solid in 40% yield. <sup>1</sup>H NMR (500 MHz, DMSO) δ 9.60 (s, 1H), 8.37 (t, *J* = 5.6 Hz, 1H), 7.34 – 7.14 (m, 8H), 6.92 – 6.86 (m, 1H), 3.25 (td, *J* = 7.1, 5.6 Hz, 2H), 2.65 – 2.58 (m, 2H), 1.81 (qd, *J* = 7.7, 6.4 Hz, 2H). <sup>13</sup>C NMR (126 MHz, DMSO) δ 166.3, 157.3, 141.8, 136.2, 129.2, 128.3, 128.3, 125.7, 117.9, 117.6, 114.2, 38.8, 32.6, 30.9.

### Synthesis of 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide and N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide

In a round bottom flask the acid (1.0 equiv.), EDC (1.5 equiv.), HOBt (1.0 equiv.), the amine (1.2 equiv.), and DCM: DMF (10:1) were added. The reaction mixture was stirred at room temperature overnight, and diluted with DCM. The organic phase was washed with saturated NaHCO<sub>3</sub>, and brine, dried over MgSO<sub>4</sub>, filtered, and concentrated in vacuo. The crude was purified by preparative

TLC (2%-10% MeOH-CHCl<sub>3</sub>) to afford the corresponding amide compound.

**3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide:**  $R_f = 0.44$  (MeOH/CHCl<sub>3</sub> (10:90)), Colourless solid, Yield (16 mg, 44%)

<sup>1</sup>H NMR, 400 MHz, DMSO-*d*<sub>6</sub>: 8.80 (t,  $J=5.6$  Hz, 1H), 8.4 (s, 1H), 8.08 (t,  $J=8.0$  Hz, 2H), 7.62 (t,  $J=7.8$  Hz, 1H), 7.22 (dd,  $J=16, 8.0$  Hz, 4H), 5.12 (t,  $J=5.7$  Hz, 1H), 4.45 (d,  $J=5.7$  Hz, 2H), 3.49 (dd,  $J=14.0, 6.6$  Hz, 2H), 2.84 (t,  $J=7.4$  Hz, 2H), 2.63 (s, 3H). <sup>13</sup>C NMR, 400 MHz, DMSO-*d*<sub>6</sub>: 197.7, 165.4, 140.4, 137.8, 136.8, 135.0, 131.8, 130.8, 128.9, 128.4, 126.8, 126.6, 62.8, 41.1, 34.8, 26.9.

**N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide:**  $R_f = 0.5$  (MeOH/CHCl<sub>3</sub> (2:98)), Colourless solid, Yield (5.0 mg, 34%)

<sup>1</sup>H NMR, 400 MHz, DMSO-*d*<sub>6</sub>: 9.07 (t,  $J=5.5$  Hz, 1H), 7.95 (d,  $J=7.2$  Hz, 2H), 7.84 (d,  $J=7.7$  Hz, 1H), 7.65 (t,  $J=7.5$  Hz, 1H), 7.58 (t,  $J=7.1$  Hz, 1H), 7.53-7.47 (m, 4H), 4.85 (d,  $J=5.9$  Hz, 2H), 3.10 (t,  $J=5.2$  Hz, 4H), 1.57 (br, 4H), 1.47 (d,  $J=4.3$  Hz, 2H)

<sup>13</sup>C NMR, 400 MHz, DMSO-*d*<sub>6</sub>: 166.6, 138.5, 134.5, 134.0, 133.2, 131.5, 129.8, 128.4, 127.7, 127.3, 127.1, 45.7, 39.6, 24.9, 23.1

Table S3. Sequence identity (%) between McbA and its ancestors A1-A4.

	<b>McbA</b>	<b>A1</b>	<b>A2</b>	<b>A3</b>	<b>A4</b>
<b>McbA</b>	100.0	70.0	62.4	50.4	40.0
<b>A1</b>	70.0	100.0	85.2	65.0	48.8
<b>A2</b>	62.4	85.2	100.0	75.4	54.8
<b>A3</b>	50.4	65.0	75.4	100.0	66.8
<b>A4</b>	40.0	48.8	54.8	66.8	100.0

Table S4. Melting temperature ( $T_m$ ) of McbA and ancestors A1-A4 measured by nano differential scanning fluorimetry.

<b>Enzyme</b>	<b><math>T_m</math> 1 (°C)</b>	<b><math>T_m</math> 2 (°C)</b>
McbA	48.1	
A1	73.1	
A2	66.0	79.3
A3	43.8	73.3
A4	67.9	

Table S5. Protein expression levels of McbA and ancestors

<b>Enzyme</b>	<b>Approximate mg protein per g pellet</b>
McbA	10
A1	7
A2	8
A3	14
A4	8

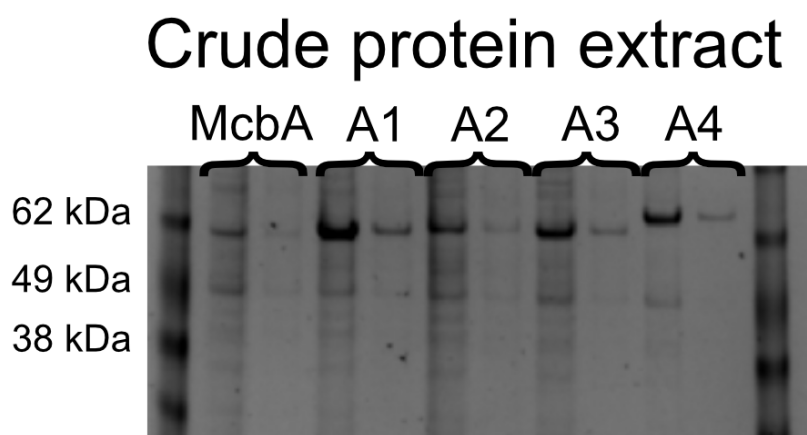


Figure S1. Crude protein content of McbA and ancestor A1-A4. Samples were normalized by OD600, two dilutions of each sample were loaded onto the gel. The molecular weight marker is SeeBlue Plus2.

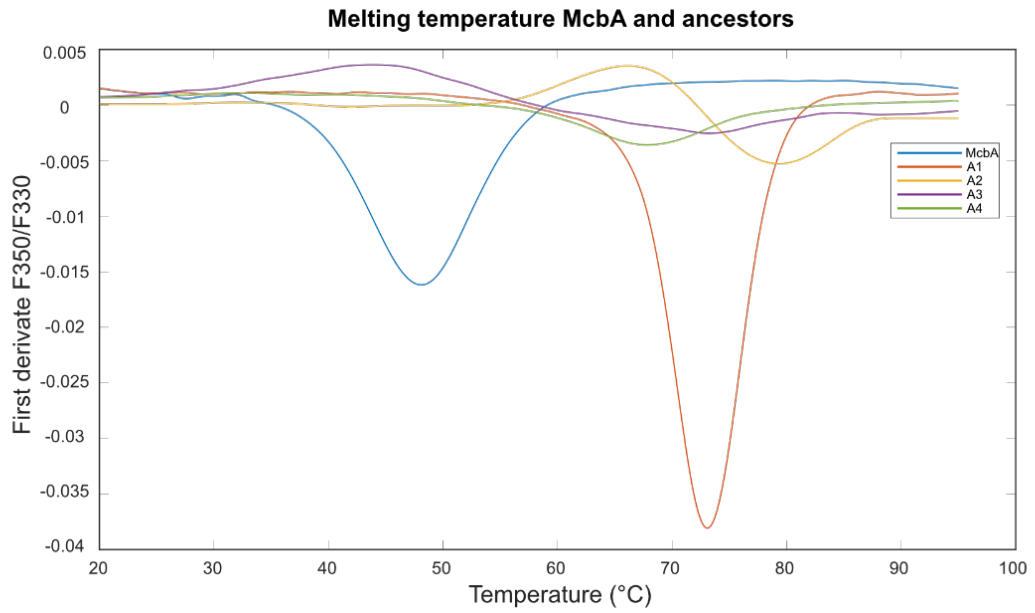


Figure S2. Thermal unfolding by McbA and ancestors A1-A4, measured by change in fluorescence ratio at wavelengths 350 and 330 nm by Prometheus NT.48 nanoDSF (NanoTemper Technologies).

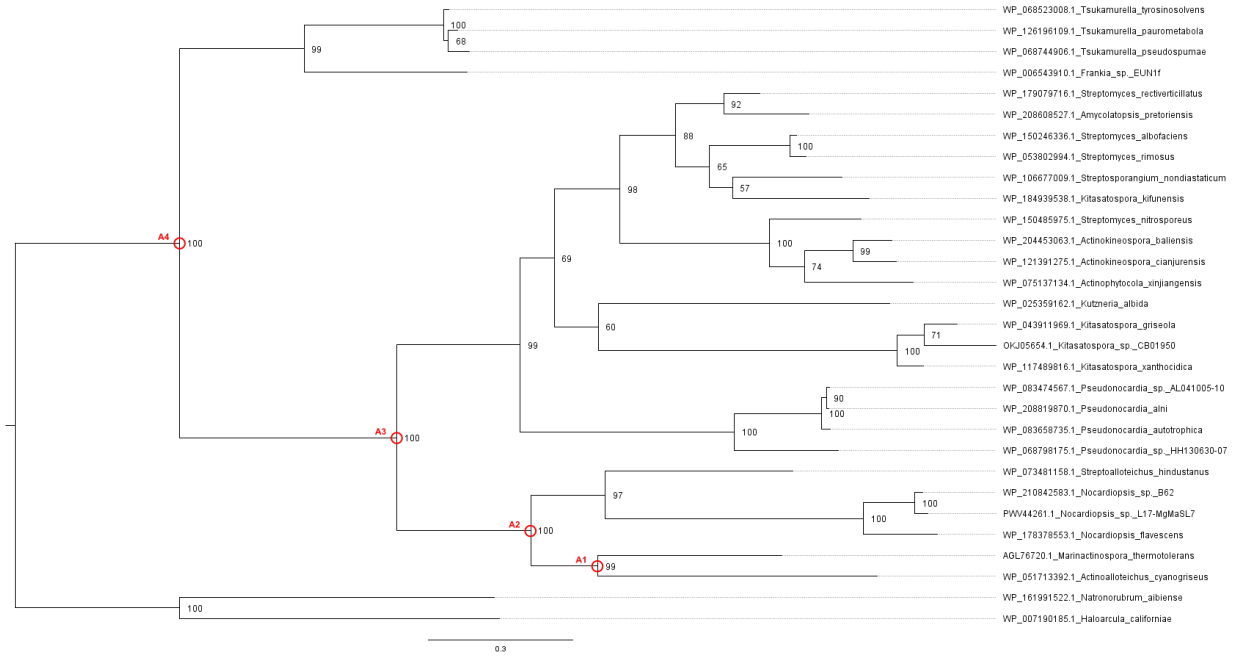


Figure S3. Phylogenetic tree of McbA used to compute ancestral sequences. The numbers shown on the nodes are the bootstrap values (1000 bootstrap).



AGL76720.1\_Marinactinospora\_thermotolerans  
WP\_051713392.1\_Actinocalloteichus\_cyanogriseus  
WP\_073481158.1\_Streptoalloteichus\_hindustanus  
WP\_210842583.1\_Nocardioopsis\_sp.\_B62  
PWV44261.1\_Nocardioopsis\_sp.\_L17-MgMaSL7  
WP\_178378553.1\_Nocardioopsis\_flavescens  
WP\_179079716.1\_Streptomyces\_rectiverticillatus  
WP\_208608527.1\_Amycolatopsis\_pretoriensis  
WP\_150246336.1\_Streptomyces\_albofaciens  
WP\_053802994.1\_Streptomyces\_rimosus  
WP\_106677009.1\_Streptosporangium\_nondiastaticum  
WP\_184939538.1\_Kitasatospora\_kifunensis  
WP\_150485975.1\_Streptomyces\_nitrosporeus  
WP\_204453063.1\_Actinokineospira\_baliensis  
WP\_121391275.1\_Actinokineospira\_cianjurensis  
WP\_075137134.1\_Actinophytocola\_xinjiangensis  
WP\_025359162.1\_Kutzneria\_albida  
WP\_083474567.1\_Pseudonocardia\_sp.\_AL041005-10  
WP\_208819870.1\_Pseudonocardia\_alni  
WP\_083658735.1\_Pseudonocardia\_autotrophica  
WP\_068798175.1\_Pseudonocardia\_sp.\_HH130630-07  
WP\_043911969.1\_Kitasatospora\_griseola  
WP\_117489816.1\_Kitasatospora\_xanthocidica  
OKJ05654.1\_Kitasatospora\_sp.\_CB01950  
WP\_068523008.1\_Tsukamurella\_tyrosinosolvans  
WP\_126196109.1\_Tsukamurella\_paurometabola  
WP\_068744906.1\_Tsukamurella\_pseudospumae  
WP\_006543910.1\_Frankia\_sp.\_EUN1f  
WP\_161991522.1\_Natronocrubrum\_aibiense  
WP\_007190185.1\_Haloarcula\_californiae

AGL76720.1\_Marinactinospora\_thermotolerans  
WP\_051713392.1\_Actinocalloteichus\_cyanogriseus  
WP\_073481158.1\_Streptoalloteichus\_hindustanus  
WP\_210842583.1\_Nocardioopsis\_sp.\_B62  
PWV44261.1\_Nocardioopsis\_sp.\_L17-MgMaSL7  
WP\_178378553.1\_Nocardioopsis\_flavescens  
WP\_179079716.1\_Streptomyces\_rectiverticillatus  
WP\_208608527.1\_Amycolatopsis\_pretoriensis  
WP\_150246336.1\_Streptomyces\_albofaciens  
WP\_053802994.1\_Streptomyces\_rimosus  
WP\_106677009.1\_Streptosporangium\_nondiastaticum  
WP\_184939538.1\_Kitasatospora\_kifunensis  
WP\_150485975.1\_Streptomyces\_nitrosporeus  
WP\_204453063.1\_Actinokineospira\_baliensis  
WP\_121391275.1\_Actinokineospira\_cianjurensis  
WP\_075137134.1\_Actinophytocola\_xinjiangensis  
WP\_025359162.1\_Kutzneria\_albida  
WP\_083474567.1\_Pseudonocardia\_sp.\_AL041005-10  
WP\_208819870.1\_Pseudonocardia\_alni  
WP\_083658735.1\_Pseudonocardia\_autotrophica  
WP\_068798175.1\_Pseudonocardia\_sp.\_HH130630-07  
WP\_043911969.1\_Kitasatospora\_griseola  
WP\_117489816.1\_Kitasatospora\_xanthocidica  
OKJ05654.1\_Kitasatospora\_sp.\_CB01950  
WP\_068523008.1\_Tsukamurella\_tyrosinosolvans  
WP\_126196109.1\_Tsukamurella\_paurometabola  
WP\_068744906.1\_Tsukamurella\_pseudospumae  
WP\_006543910.1\_Frankia\_sp.\_EUN1f  
WP\_161991522.1\_Natronocrubrum\_aibiense  
WP\_007190185.1\_Haloarcula\_californiae

AGL76720.1\_Marinactinospora\_thermotolerans  
WP\_051713392.1\_Actinocalloteichus\_cyanogriseus  
WP\_073481158.1\_Streptoalloteichus\_hindustanus  
WP\_210842583.1\_Nocardioopsis\_sp.\_B62  
PWV44261.1\_Nocardioopsis\_sp.\_L17-MgMaSL7  
WP\_178378553.1\_Nocardioopsis\_flavescens  
WP\_179079716.1\_Streptomyces\_rectiverticillatus  
WP\_208608527.1\_Amycolatopsis\_pretoriensis  
WP\_150246336.1\_Streptomyces\_albofaciens  
WP\_053802994.1\_Streptomyces\_rimosus  
WP\_106677009.1\_Streptosporangium\_nondiastaticum  
WP\_184939538.1\_Kitasatospora\_kifunensis  
WP\_150485975.1\_Streptomyces\_nitrosporeus  
WP\_204453063.1\_Actinokineospira\_baliensis  
WP\_121391275.1\_Actinokineospira\_cianjurensis  
WP\_075137134.1\_Actinophytocola\_xinjiangensis  
WP\_025359162.1\_Kutzneria\_albida  
WP\_083474567.1\_Pseudonocardia\_sp.\_AL041005-10  
WP\_208819870.1\_Pseudonocardia\_alni  
WP\_083658735.1\_Pseudonocardia\_autotrophica  
WP\_068798175.1\_Pseudonocardia\_sp.\_HH130630-07  
WP\_043911969.1\_Kitasatospora\_griseola  
WP\_117489816.1\_Kitasatospora\_xanthocidica  
OKJ05654.1\_Kitasatospora\_sp.\_CB01950  
WP\_068523008.1\_Tsukamurella\_tyrosinosolvans  
WP\_126196109.1\_Tsukamurella\_paurometabola  
WP\_068744906.1\_Tsukamurella\_pseudospumae  
WP\_006543910.1\_Frankia\_sp.\_EUN1f  
WP\_161991522.1\_Natronocrubrum\_aibiense  
WP\_007190185.1\_Haloarcula\_californiae

130 140 150 160 170 180  
AGL76720.1\_Marinactinospora\_thermotolerans G R D I L A A S V P E G T F L R Y R E H E G T A V V A F T S G T T C P K G V A H S S T A M S A C V . D . . .  
WP\_051713392.1\_Actinoalloteichus\_cyanogriseus G E D L V L A A A V A D A P A L P A A L D R P A S I A Y T S G T T C S P K G V H G T T A M A A C L . D . . .  
WP\_073481158.1\_Streptoalloteichus\_hindustanus G E D L V L A A G A Q S A E P V E P T A D E R A A T A V G F T G C T G R A K G V C R A P F D L E A C L . D . . .  
WP\_210842583.1\_Nocardiopsis\_sp.\_B62 G E N L L E L A A Q X P D A P M E C V H R D D A L S M V T F T S S G S T G E P K G I A H F A S S A G F L . D . . .  
PWV44261.1\_Nocardiopsis\_sp.\_L17-MgMaSL7 G E N L L E L A A Q X P D V P M E C V H R D D A L S M V T F T S S G S T G E P K G I A H F A S S A G F L . D . . .  
WP\_178378553.1\_Nocardiopsis\_flavescens G E N L L A D G R P D A P L D C A D R D D V S M V T F T S S G S T G E P K G I T H F A S S S G F F . D . . .  
WP\_179079716.1\_Streptomyces\_rectiverticillatus G G D L L E L A A T E S T E P I S G L A R F E D I C S I R H T G G T T G H P K G I C T S F E R A G R M R . P . . .  
WP\_208608527.1\_Amycolatopsis\_pretoriensis G T D L L E L A A A P T E R I S G L A R F G D I C S I R H T G G T T G H P K G I C T D F E R A G R M R . P . . .  
WP\_150246336.1\_Streptomyces\_albofaciens G T D L L E L A S A E S D E P V P S G A R F D D I C T I R H T G G T T G H P K G I C T Y G Q V R T F V . D D E D  
WP\_053802994.1\_Streptomyces\_rimosus G T D L L E L A S A E S D E P V P S G A R F E D I C T I R H T G G T T G H P K G I C T Y G Q V R T F V A D D E D  
WP\_106677009.1\_Streptosporangium\_nondiastaticum G T D L L E L A A G O P A E P L S G L A R F E D I C M I R H T G G T T G H P K A I C T A F E R L P G F Q . D D V D  
WP\_184939538.1\_Kitasatospora\_kifunensis G K D L L E L A A Q E S A E P L A G L A R A E D I A T I R H T G G T T G H P K G I C S T F G G G A A F R . N G L Q  
WP\_150485975.1\_Streptomyces\_nitrosoporus G E D L L L A A A P O P E T P V E S Q A R F E D L C M I R H S G G T T G H P K G V R S T F A R M R N T G . L . . .  
WP\_204453063.1\_Actinokineospora\_baliensis G E D L L L A A A G L S D E P V A S A A R F G D L C F I R H S G G T T G H P K G I R S T F A R M R G L D . R . . .  
WP\_121391275.1\_Actinophytocola\_cianjurgensis G E D L L L A A A H S D L P V E S T A R F G D L C F I R H S G G T T G H P K G I R S T F A R M R G L D . R . . .  
WP\_075137134.1\_Actinophytocola\_xinjiangensis A E D L L L A A A H S P D D P V T S Q A R F D D L C F I R H S G G T T G H P K G I R S T F A R M R G L D . Q . . .  
WP\_025359162.1\_Kutzneria\_albida G S D L L L A A A A D S P V P P H P V A F G D T W S V R H T G G T T G H P K G I L M P S G G Y R A M L . E . . .  
WP\_083474567.1\_Pseudonocardia\_sp.\_AL041005-10 D T D V T A E A A T A S T E P V R G R A R F D A E Q I R H T G G T T G H P K G I T Y T F D H H R R G A . A . . .  
WP\_208819870.1\_Pseudonocardia\_alni D T D V T A E A A T A S T E P V R G R A R F D A E Q I R H T G G T T G H P K G I T Y T F D H H R R G A . A . . .  
WP\_083658735.1\_Pseudonocardia\_autotrophica D T D V T A E A A T A S T E P V R G R A R F D A E Q I R H T G G T T G H P K G I T Y T F D H H R R G A . A . . .  
WP\_068798175.1\_Pseudonocardia\_sp.\_HH130630-07 R T D V L A A A H A S T A P I T G R A R L S D V E Q I R H T G G T T G H P K G I T Y T F G H H V R A A . A . . .  
WP\_043911969.1\_Kitasatospora\_griseola G E D V I A A A H R P A E P D V R I R F E D D L S I R H T G G T T G H P K G I L L T H G P Y R A I F . D . . .  
WP\_117489816.1\_Kitasatospora\_xanthocidica G E D V I A A A H R P A E P D V R V G F E D D L S I R H T G G T T G H P K G I L S L H G P Y R I F . D . . .  
OKJ05654.1\_Kitasatospora\_sp.\_CB01950 G E D V I A A A H R P A E P D V R I G F D D I S I R H T G G T T G H P K G I L T R H G Y R L I F . D . . .  
WP\_126196109.1\_Tsukamurella\_paurometabola G V D L T A V A A G E P P G L A A H T A D F S S T A S L A Y T G G T T G K S K G V M L S Y S G G A A L L . R . . .  
WP\_068744906.1\_Tsukamurella\_pseudospumae G V D L T A V A A G E P P G L A A H T A D F S S T A S L A Y T G G T T G K S K G V M L S Y S G G A A L L . R . . .  
WP\_006543910.1\_Frankia\_sp.\_EUN1f G A D L V L A A G Y E P R L T V P R L D F E A T A A I A Y T G G T T G Q P K G V M Y T R G S A M T . Q . . .  
WP\_161991522.1\_Natronorubrum\_aibiense W R S A A R L E A D Q A P D S P S V S V E P D A T A G H F Y T G G T T G D P K G V C M T Y S C L A T N L . L . . .  
WP\_007190185.1\_Haloarcula\_californiae V D S F D Q L L E T A E P V A P D V S V S K S T L A G V F H T G G T T G D P K G V K H T Q E N L A L N A . L . . .

190 200 210 220  
AGL76720.1\_Marinactinospora\_thermotolerans . . . . . A A V S M Y G . . . . . R G P W R F L I P I P S D I G G E L A Q C T L A T G G T V V L E E F  
WP\_051713392.1\_Actinoalloteichus\_cyanogriseus . . . . . V A R A M Y G . . . . . P P P W R F L V A I P I S D I G G E L A Q W I L A C G G V V V R E D L F  
WP\_073481158.1\_Streptoalloteichus\_hindustanus . . . . . A S L T I F G . . . . . E G P W R F L V C I P I A D I G G E M A E W T L A A G G T V R E D F  
WP\_210842583.1\_Nocardiopsis\_sp.\_B62 . . . . . N A R H M Y G . . . . . P G P W R F L V A I P I S D I G G E I V Q W T L A V G G T V L D D F  
PWV44261.1\_Nocardiopsis\_sp.\_L17-MgMaSL7 . . . . . N A R H M Y G . . . . . P G P W R F L V A I P I S D I G G E I V Q W T L A V G G T V L D D F  
WP\_178378553.1\_Nocardiopsis\_flavescens . . . . . N A L H M Y G . . . . . P G P W R F L V A I P I S D I G G E I V Q W T L A G G T V L M D D F  
WP\_179079716.1\_Streptomyces\_rectiverticillatus . . . . . Q P G E A A . . . . . A P P R Q L V S T P L A H A A G L V A D H L L A E G G T V L E D F  
WP\_208608527.1\_Amycolatopsis\_pretoriensis . . . . . T P D R R K P . . . . . D P P R Q L V S T P L A H A A G L V A D H V L A E G G T V L E D F  
WP\_150246336.1\_Streptomyces\_albofaciens P G D A K V L W E R P A D Q . . . . . E R T R L V C T T L A H A A G M M A D V T L R E H G L V L V L E D F  
WP\_053802994.1\_Streptomyces\_rimosus A G D A K V L W E R P A D Q . . . . . E G N R L V C T T L A H A A G M M A D V T L R E H G L V L V L E D F  
WP\_106677009.1\_Streptosporangium\_nondiastaticum . . . . . E I L E P E G K A D . . . . . D G M R Q L V C T T L A H A A G L L A D V T L R T D G G T V L L D D F  
WP\_184939538.1\_Kitasatospora\_kifunensis . . . . . D S E E A E G E E . . . . . K V E R R Q L V C T T L A H A A G H L A D H T L G E G G A V L L D D F  
WP\_150485975.1\_Streptomyces\_nitrosoporus . . . . . M R E L S D . . . . . E V R R D L V C T P L S H A G G F L A D T L A S G G T V L H D G F  
WP\_204453063.1\_Actinokineospora\_baliensis . . . . . V S G T F G G . . . . . A Q P D L V C T P L S H A G G F L A D T L V A G S G T V L H Y A F  
WP\_121391275.1\_Actinophytocola\_cianjurgensis . . . . . V R A T N D V . . . . . D E P R D L V C T P L S H A G G F L A D A T L A A G G T V L H Y A F  
WP\_075137134.1\_Actinophytocola\_xinjiangensis . . . . . V S A L F G G . . . . . H T P R D L V C T P V S H A A G F I A D A T L A A G G T V L H R T F  
WP\_025359162.1\_Kutzneria\_albida . . . . . H Q A T A G G . . . . . A D V V Y L S I T P L A H A G A C N L A D G V F A A G G T V I H R F  
WP\_083474567.1\_Pseudonocardia\_sp.\_AL041005-10 . . . . . M R R G T G G F G S G . . . . . D A G A R L L V A T P V A H V G G S L A D R T L T D G G T V L Q D D F  
WP\_208819870.1\_Pseudonocardia\_alni . . . . . M R R G T G G F G S G . . . . . D A G A R L L V A T P V A H V G G S L A D R T L T D G G T V L Q D D F  
WP\_083658735.1\_Pseudonocardia\_autotrophica . . . . . M R R G T G G F G S G . . . . . D A G A R L L V A T P V A H V G G S L A D R T L A D G G T V L Q D G F  
WP\_068798175.1\_Pseudonocardia\_sp.\_HH130630-07 . . . . . M R R S M G S P G T D A G E S G P P R M L A A T P V A H A G G A L G D R M L A G G G T V L L R F  
WP\_043911969.1\_Kitasatospora\_griseola . . . . . T P V G D D R . . . . . A P R L L A A T P L A H I A G V L S D L V L H H G G T V L Q R S F  
WP\_117489816.1\_Kitasatospora\_xanthocidica . . . . . A P I G S D P A . . . . . A G A P R L L A A T P L A H I A G F I A D L T L Q H G G T V L Q R S F  
OKJ05654.1\_Kitasatospora\_sp.\_CB01950 . . . . . D P V G N E E . . . . . A P R L L A A T P L A H I A G V F S D L V L Y H G G L V L L R S F  
WP\_126196109.1\_Tsukamurella\_tyrosinosolvens . . . . . I Q R T E W E W . . . . . P E E I R F L V C A P L S H A G G A F W N P T S M T G S M V L P R F  
WP\_068744906.1\_Tsukamurella\_pseudospumae . . . . . I Q R T E W E W . . . . . P E E I R F L V C A P L S H A G G A F W N P T S M G S S M V L P R F  
WP\_006543910.1\_Frankia\_sp.\_EUN1f . . . . . I L L T E W Q W . . . . . P K E L R R F L V C T P L S H A G A P F I P V P L H Q G G T V L P K F  
WP\_161991522.1\_Natronorubrum\_aibiense . . . . . A H P A D L G F . . . . . S S D D T G L V A T P I S H A G T F L L A T L A G G T V L R R F  
WP\_007190185.1\_Haloarcula\_californiae . . . . . A H A V E L D T . . . . . Q P R V L L M T P L P S A G L M A G G L A Q G S H H V I T Q G F

230 240 250 260 270 280  
AGL76720.1\_Marinactinospora\_thermotolerans Q P D A V L E A I E R E R A T H V F L A P N W L Y Q L A E H P A L P R S D L S S L R R V V Y G G A P A P S R V A A A  
WP\_051713392.1\_Actinoalloteichus\_cyanogriseus D P L A T L A L L E R E R I T H L F T A P S S L Y Q L A E H P D L D H F D L T A L R L A V Y G G A P A P A R T S A A  
WP\_073481158.1\_Streptoalloteichus\_hindustanus E P A D I L A T I G A E R I T H V F C A P G W V Y Q L A E H P A L A D A D L S S L T Q I P Y G G A P S P A R T A D A  
WP\_210842583.1\_Nocardiopsis\_sp.\_B62 R A D E V A R I L G S E H I T H F F G S P M V H A L V H E P G L R N A D L S D L R L V A Y G G A P S P S R T A D A  
PWV44261.1\_Nocardiopsis\_sp.\_L17-MgMaSL7 R A D E A V R I L G R E R I T H F F G S P M V H A L V H E P G L R N R L P D L R L V A Y G G A V S P S R T A D A  
WP\_178378553.1\_Nocardiopsis\_flavescens D P G Q V L A A I E R E R I T H L F L L P P L L Y Q L M D H P H A D R T D T S S L R Q V I Y G G C Q T S P A R T A D A  
WP\_179079716.1\_Streptomyces\_rectiverticillatus D P G Q V L A A I E R E R I T H L F L L P P L L Y S L V D H P D A V R T D L S S L R Q V N Y G G C K A S P A R T A D A  
WP\_208608527.1\_Amycolatopsis\_pretoriensis D P A E V L A T I E R E R I T H L F L L P P L L Y Q L T D H P D A A T T D T S S L R C L T Y G G C Q S S P A R T A D A  
WP\_150246336.1\_Streptomyces\_albofaciens D P A E V L A T I E R E K I T H L F L L P P L L Y R L T D H P D A A T T D T S S L R C L T Y G G C Q S S P A R T A D A  
WP\_053802994.1\_Streptomyces\_rimosus D A A B A L A A I E R E R I T H F L F L L P P L L Y Q L M D H P D V D S T P T S S L R N W T Y G G C Q S S P A R T A D A  
WP\_106677009.1\_Streptosporangium\_nondiastaticum D P G V L A T I E R E R I T R L V L P P L M Y Q L D H P D S K D A D T S S L R T L L Y G G C L A S P A R T A D A  
WP\_184939538.1\_Kitasatospora\_kifunensis D A G E V L A T I A R E E I T H V L L L P P H L Y Q V L D H P D E T T T S S M R R I T Y G G C Q S S P A R T A D A  
WP\_204453063.1\_Actinokineospora\_baliensis D P G E V L A T I E R E R I T H A F L L P P L L Y Q V L D H P . G E F D T S S I Q R I T Y G G T P S A P V R T A Q A  
WP\_121391275.1\_Actinophytocola\_cianjurgensis D P G E V L A T I E R E R I T H V F L L P P L L Y R V L D H P . G D H T S S L S R I T Y G G T P S A P V R T A Q A  
WP\_075137134.1\_Actinophytocola\_xinjiangensis D P G E V L A T I E R E R I T H L F L L P P L L Y Q V L D H P . A R V D T S S L R Q L T Y G G T P S P T R T A E A  
WP\_025359162.1\_Kutzneria\_albida A P A Q V L A D I E R H R V D F L L P P L L Y R V L D D P A S A S T D T S S L R R F L Y G G C A A S P T R T A E A  
WP\_083474567.1\_Pseudonocardia\_sp.\_AL041005-10 E P G A Y L A A V E R E R I T H G F L L P L I Y R L D H P D L D R T D L S S L R S L V Y G G A P A P P R R T A E A  
WP\_208819870.1\_Pseudonocardia\_alni E P G A Y L A A V E R E R I T H G F L L P L I Y R L D H P D L D R T D L S S L R S L V Y G G A P A P P R R T A E A  
WP\_083658735.1\_Pseudonocardia\_autotrophica E A G A F L A A V E R E R I T H F L L P L I Y R L D H P D L D R T D L S S L R S L V Y G G A P A P P R R T A E A  
WP\_068798175.1\_Pseudonocardia\_sp.\_HH130630-07 D A G A V L A A V E R E R I T H M L L P P L L Y R V L D H P G L D R T D L S S L R T L Y G G A P A P P R R T A E A  
WP\_043911969.1\_Kitasatospora\_griseola D P G A V L A A V E R E R I T D L W V L P P L L Y Q L D H P A I G R T D L S S L R Q I M Y G G A A S P T R T A E A  
WP\_117489816.1\_Kitasatospora\_xanthocidica E P G A V L A A I E R E R I T D L W V L P P L L H Q L D H P A I G R T D L S S L R Y L S Y G G S A A S P T R T A E A  
OKJ05654.1\_Kitasatospora\_sp.\_CB01950 E P G A V L S A V E R E R I T E F W V L P A L L H Q L D H P A I A R T D L S S V R Q I I Y G G S A A S P T R T A E A  
WP\_126196109.1\_Tsukamurella\_paurometabola E P E A V L A A I E K Y R I T A T M L V P T M I Y A L L D H P K F D E Y D L S S L E T V F Y G A S A I S P T R T A E A  
WP\_068744906.1\_Tsukamurella\_pseudospumae E P E A V L A A I E K Y R I T A T M L V P T M I Y A L L D H P K F D E Y D L S S L E T V F Y G A S A I S P T R T A E A  
WP\_006543910.1\_Frankia\_sp.\_EUN1f D A G R V L A I E Q H R I T S V F L V P T S M I Y A L L D H P R E D E T P L S S L E T V F Y G S S P M S P R T A E A  
WP\_161991522.1\_Natronorubrum\_aibiense D E T D F C A A V E T H G V T W T L V P T M L Y R L D G P L A A H V S S L E N V Y G A A P I Q P D R L Q D A  
WP\_007190185.1\_Haloarcula\_californiae D A E R A L E L V E Q E D I S W T F M V P T M V Y R V L D V L E R E T H D T S L E T L V Y G A A P M K E D R L R E G



```

                                450      460      470      480      490
AGL76720.1_Marinactinospora_thermotolerans  . . DDPD EHLRA LVRDHLGDLHVR RRVVRFVRS PVPV PAKRDKVKVRFWFTD . . . . .
WP_051713392.1_Actinocalloteichus_cyanogriseus . . ADAD EHLREHVAQLGEPHVRRLIRFVDTIPLTDGCKRDKSALRALFFADTFSR . .
WP_073481158.1_Streptoalloteichus_hindustanus  RAVAFSALAAAVRTILGQVHEPRIDLLASPLTPRCKRDKKTAIRGQAADRACSPHV
FWW44261.1_Nocardopsis_sp._B62                RTVAFSALAAAVRTILGQVHEPRIDLLASPLTPRCKRDKKTAIRGQAADRACSPHI
WP_178378553.1_Nocardopsis_flavescens         RTAEFTTAGLAAVRTILGQVHEPRIDLLASPLTPRCKRDKKTAIRGQAEATASPT
WP_179079716.1_Streptomyces_rectiverticillatus . . KVGAD EHLAMVRQRGAMYEPHITITFAEQPLTDAGKRDKKRLSRATPTQERSQG
WP_208608527.1_Amycolatopsis_pretoriensis     . . EVDPDALRAMVRQRGAMYVPHITITFVDRPLTDGCKRDKKRLSRHAAQAEGR . .
WP_150246336.1_Streptomyces_albifaciens       . . DVADQLRAMVREARGAMYEPHIEFVEQPLTDAGKRDKKQLRAMAARTAV . .
WP_053802994.1_Streptomyces_rimosus          . . DVADQLRAMVREARGAMYEPHIEFVEQPLTDAGKRDKKQLRAQATRSAA . .
WP_106677009.1_Streptosporangium_nondiasticum . . ELDEKELRELFVGNRERGAMYEPHVSFIDAPLTDAGKRDKKFLRREAEAAAATR . .
WP_184939538.1_Kitasatospora_kifunensis      . . TVEQELRAMVREERGAMYEPARITFADAPLTDVGRDKKLRQWVQAAV . .
WP_150485975.1_Streptomyces_nitrosporeus     . . TTTEDELRLALVRDRRGLVYVPLIEFVADPLTDAGKRDKKLRRAAA . .
WP_204453063.1_Actinokineospora_baliensis    . . TVTAE SLR TMVRDRGRDMYVDPHIVLVDAPLTDVGRDKKLRRAADRSTAR.G
WP_121391275.1_Actinokineospora_cianjurenensis . . TTTEELR TMVRDRGRDMYVDPHIVLVDAPLTDVGRDKKLRRAATSTP . .
WP_075137134.1_Actinophytocola_xinjiangensis . . AVTEDELRLQMVSTGRGAMYVDRVFEVAAPLTDVGRDKKLRRAAVG . .
WP_025359162.1_Kutzneria_albida              . . TLTEDELRLQMVSTGRGALYTPHHVFEVDRPLTDAGKRDKKVLRAARH . .
WP_083474567.1_Pseudonocardia_sp._AL041005-10 . . PSPDELTLGLVADRAGAMYVDTVEFVEPLTDGCKRDKKRLRAELVQQR . .
WP_208819870.1_Pseudonocardia_alni          . . PSPDELTLGLVADRAGAMYVDTVEFVEPLTDGCKRDKKRLRAELVQQR . .
WP_083658735.1_Pseudonocardia_autotrophica   . . PSPDELTLGLVADRAGAMYVDTVQFVEPLTDGCKRDKKRLRAELVQQR . .
WP_068798175.1_Pseudonocardia_sp._HH130630-07 . . PCDDELTLGLVADRAGAMYVRSISFVAEPLTDGCKRDKKRLRAELTA . .
WP_043911969.1_Kitasatospora_griseola       . . RPDELRLRAFVTERKGRIFAPALHLVDEPLTFVGRDKKRLRAEAATR . .
WP_117489816.1_Kitasatospora_xanthocidica   . . RPDLTLRAFVTERKGRIFAPVLLHVDPEPLTFVGRDKKRLRAAATA . .
OKJ05654.1_Kitasatospora_sp._CB01950        . . QPHDLRLRAFVTERKGRIFAPALHLVDEPLTFVGRDKKRLRAAVAG . .
WP_068523008.1_Tsukamurella_tyrosinosolvans . . TVDVEELRLDRRKGAVYTPKTVDFVESPVPVSLGKRDKKALRAQYA . .
WP_126196109.1_Tsukamurella_paurometabola   . . TVDVEELRLDRRKGAVYTPKTVDFVESPVPVSLGKRDKKALRAQYA . .
WP_068744906.1_Tsukamurella_pseudosummae   . . TVDVEELRLDRRKGAVYTPKTVDFVESPVPVSLGKRDKKALRAIYA . .
WP_006543910.1_Frankia_sp._EUN1f            . . SVAAEELALVKEHKGSHHAPKGVDFVADPLTAVGKRDKKALRAHWSNDRRVG
WP_161991522.1_Natronorubrum_aibiense       . . DIDVDELLAFAGEKRSYKPKSVDFVETPTTFYKIDREALRERYSWDEERRIN
WP_007190185.1_Haloarcula_californiae       . . TVDRANITSEFCRDLSDYEVVPSKIDVVEITPTTFYKIDREALRERYSWDEERRIN

```

Figure S4. Alignment used for the phylogenetic tree construction. Alignment illustration made by using ESPrnt v 3.0, Robert, X. and Gouet, P. (2014) "Deciphering key features in protein structures with the new ENDscript server". Nucl. Acids Res. 42(W1), W320-W324 - doi: 10.1093/nar/gku316 (freely accessible online).

```

1      10      20      30      40      50
McbA  MGYARRVMDG...GEVAVTGAAGGSGTGARLRHQVRLAHALTEACHPGRGVAC
A1    MGYVRRVLDGAAADPGFELVSGRRLLTGAEPREQVHLLAALAEQCHGCDGVAC
A2    MGYVRRVLDGAAADPGFELVSGRRLLTGAEPREQVHLLAALAEQCHGCDGVAC
A3    MDVLDLGLGAAADPGFVAHLCDSTLTGSEDRVSLAALAEQCHGCDGVAC
A4    MLYELLLTTLSSREERVAITFDGTLTYGEEERRSSSLAALAEQCHGCDGVAC

60     70     80     90     100
McbA  LHNATWRAIALRLDAVQALIGCHYVGLRPTAAVTEQARATAAADSAALVFEPSVEARA
A1    LHGNTPEAVVLRDLAVQLIGCRVYGLRPMFSTTEQARFLAEAPPALVFDMDMAERA
A2    LSGNTEPAVAVLRDLAVQLIGARYVGLRPMLSADEQARILADAPALVFDRLAERA
A3    LSGNREELVAVQLVMTGARVYALRPMLSADDHAYVLADASTALVFDRFWATERA
A4    LLSNREPAVAVLRDLAVQIARVGLRPMLSVDTLWVLSADAAALVFDSPFTVTV

110    120    130    140    150    160
McbA  ADLDELRVSVPLVSLGPTSRGRDILA.ASVPEGTFLRYREHPGIAVVARTSCTGT
A1    ADLDELRVSVPLVSLGPAPLGEDLLALAAQAPAELEFRARPEAVAVTGGTTG
A2    ADLDELRVSVPLVSLGPAPTGEDLLALAAAPAEPLSARARPEAIASIAVTTGTTG
A3    ADLDELRVSVPLVSLGPTDTCEDLVLAAGAEPEPLTVHADPEAIASIAVTTGTTG
A4    EDLDELRVSVPLVSLDADTPWESLAQLAEGAEPDSSVSVPLDIAVGVVTTGTTG

170    180    190    200    210
McbA  TPKGVHHSSTAMSAQVDAVSMYGRFPWRFLPIPLSDLGGELAQCTLARGGTVVL
A1    QPKGVGHFTTAMAAQLDAALSMYGGPWRFLVCIPLSDLGGELAQCTLARGGTVVL
A2    HPKGVGHFTGGMAAMLDAQRSAAGGPWRFLVCTPLSHAGGALAQCTLARGGTVVL
A3    QPKGVGHYSGMAAMLQVQRTENG.GELRFLVCTPLSHAGGAFMNPQLARGGTVVL
A4    DPKGVGHHSGLATNLLAHPAELG.EEVHFLVATPLSHSAGCTFLAQLARGGTVVL

220    230    240    250    260    270
McbA  LEFQDQDRLVLAERERATQVFLNPNWLYQLAEHPALRSDHSSLRVRVYGGAPAV
A1    LEFQDQDRLVLAERERITHVFLNPNWLYQLAEHPALRSDHSSLRVRVYGGAPAT
A2    LEFQDQDRLVLAERERITHVFLNPNWLYQLAEHPALRSDHSSLRVRVYGGAPAS
A3    LEFQDQDRLVLAERERITHVFLNPNWLYQLAEHPALRSDHSSLRVRVYGGAPMS
A4    RQFQDQDRLVLAERERITHVFLNPNWLYQLAEHPALRSDHSSLRVRVYGGAPMS

280    290    300    310    320    330
McbA  PSLVLAARERQCAVLMNYCQEAAPFARITPDDHARRDLTAVGRPHVVEVTR
A1    PSLVLAARERQCAVLMNYCQEAAPFARITPDDHARRDLTAVGRPHVVEVTR
A2    PSLVLAARERQCAVLMNYCQEAAPFARITPDDHARRDLTAVGRPHVVEVTR
A3    PSLVLAARERQCAVLMNYCQEAAPFARITPDDHARRDLTAVGRPHVVEVTR
A4    PDLVLAARERQCAVLMNYCQEAAPFARITPDDHARRDLTAVGRPHVVEVTR

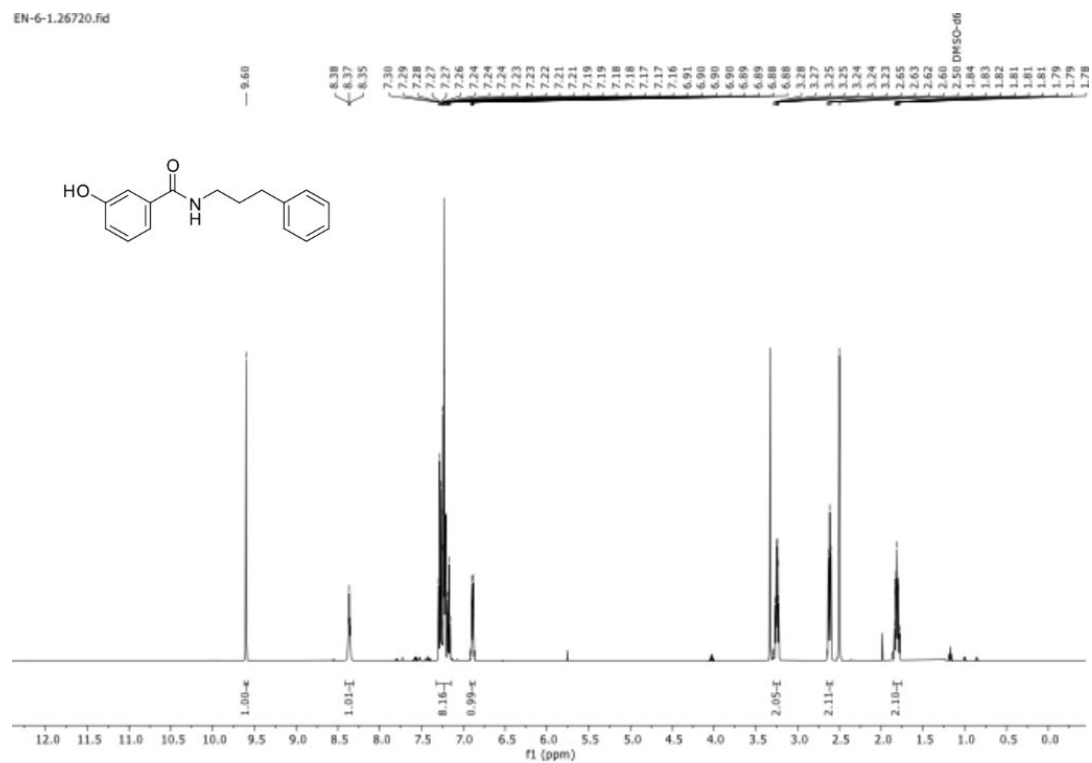
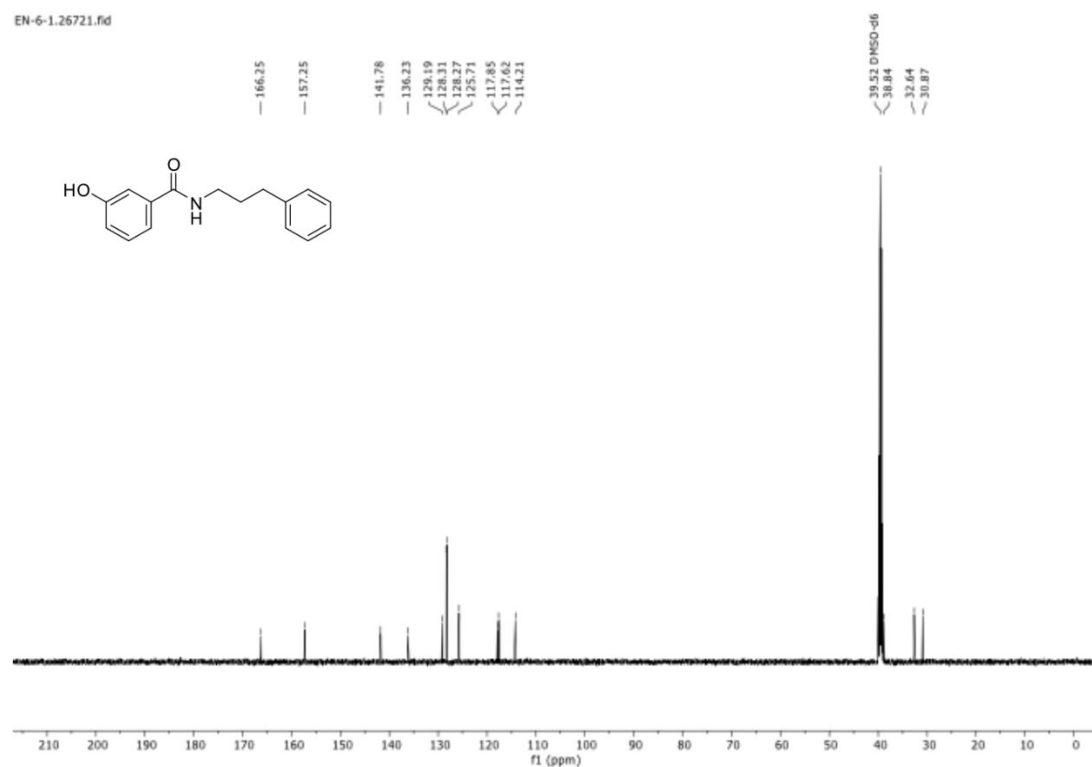
340    350    360    370    380
McbA  DDSGGTLPKCAVGEVWVRSPMTMSGYWRDPERTAQVLSGWLRTGDVGTDEDGHL
A1    DDDGGDLKPCAVGEVWVRSPMTMSGYWRDPERTAQVLSGWLRTGDVGTDEDGHL
A2    DDDGGDLKPCAVGEVWVRSPMTMSGYWRDPERTAQVLSGWLRTGDVGTDEDGHL
A3    DDDGGDLKPCAVGEVWVRSPMTMSGYWRDPERTAQVLSGWLRTGDVGTDEDGHL
A4    DDSGGDVPKPCAVGEVWVRSPMTMSGYWRDPERTAQVLSGWLRTGDVGTDEDGHL

390    400    410    420    430    440
McbA  HLLDRLQDDIIVVEAVNVYSRREVEHLVLSHDDVRAAVVGVDPDPSGEAVCAAVVVA
A1    YLVDKRVKDDIIVVEAVNVYSRREVEHLVLSHDDVRAAVVGVDPDPSGEAVCAAVVVA
A2    YLVDKRVKDDIIVVEAVNVYSRREVEHLVLSHDDVRAAVVGVDPDPSGEAVCAAVVVA
A3    YLVDKRVKDDIIVVEAVNVYSRREVEHLVLSHDDVRAAVVGVDPDPSGEAVCAAVVVA
A4    YLVDKRVKDDIIVVEAVNVYSRREVEHLVLSHDDVRAAVVGVDPDPSGEAVCAAVVVA

450    460    470    480    490
McbA  DGAADPPEHLRALVRDRLGDLHVRVRFVRSIPVTPAGKDKKALRAFAD
A1    DGAEDVDAELRALVRDRLGDLHVRVRFVRSIPVTPAGKDKKALRAFAD
A2    DGAEDVDAELRALVRDRLGDLHVRVRFVRSIPVTPAGKDKKALRAFAD
A3    DGAEDVDAELRALVRDRLGDLHVRVRFVRSIPVTPAGKDKKALRAFAD
A4    EGRDVEDDLRALVRDRLGDLHVRVRFVRSIPVTPAGKDKKALRAFAD

```

Figure S5. Alignment of McbA and ancestors A1-A4, derived from the phylogenetic tree in fig S1. Alignment illustration made by ESPrnt v 3.0, Robert, X. and Gouet, P. (2014) "Deciphering key features in protein structures with the new ENDscript server". Nucl. Acids Res. 42(W1), W320-W324 - doi: 10.1093/nar/gku316 (freely accessible online).

Figure S6. <sup>1</sup>H-NMR of synthesised 3-Hydroxy-N-(3-phenylpropyl)benzamide (standard).Figure S7. <sup>13</sup>C-NMR of synthesised 3-Hydroxy-N-(3-phenylpropyl)benzamide (standard).

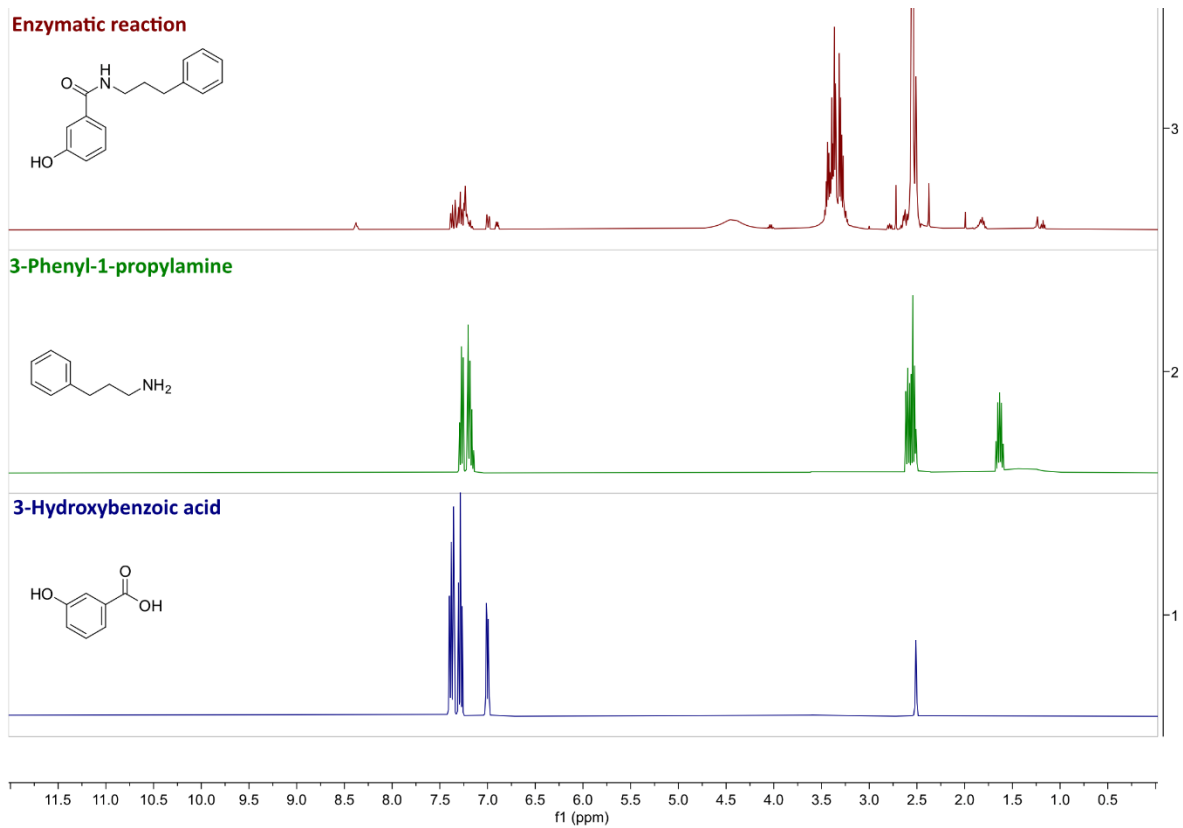


Figure S8.  $^1\text{H-NMR}$  of the enzymatic synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide (crude NMR). As a control,  $^1\text{H-NMR}$  of 3-phenyl-1-propylamine and 3-hydroxybenzoic acid are included.

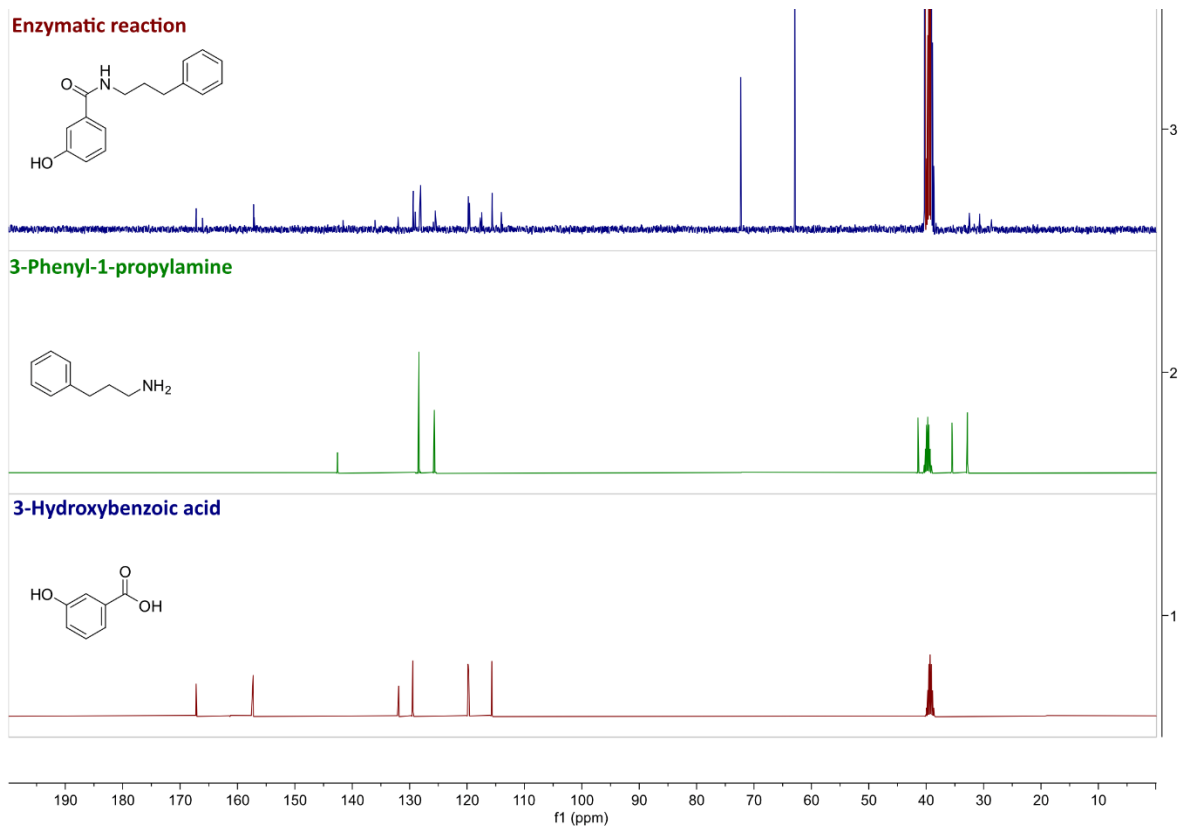


Figure S9.  $^{13}\text{C-NMR}$  of the enzymatic synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide (crude NMR). As a control,  $^{13}\text{C-NMR}$  of 3-phenyl-1-propylamine and 3-hydroxybenzoic acid are included.

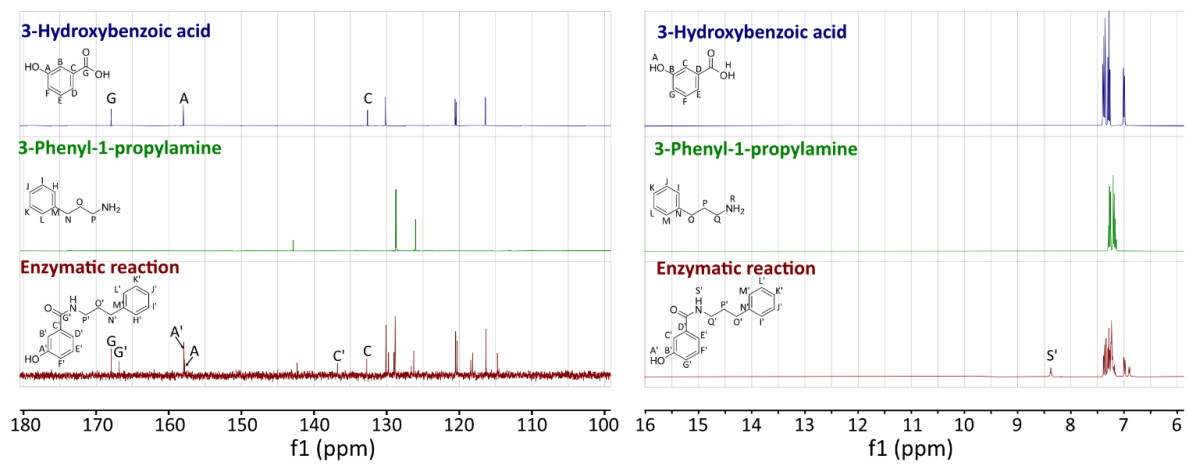


Figure S10  $^{13}\text{C}$ -NMR (left) and  $^1\text{H}$ -NMR (right) of 3-hydroxybenzoic acid, 3-phenyl-1-propylamine, and the enzymatic synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide (crude NMR). In the left chromatogram, a shift of the carbonyl carbon ( $\text{G}'$ ), the  $\alpha$ -carbon ( $\text{C}'$ ) and the aromatic carbon ( $\text{A}'$ ) is consistent with the formation of an amide. In the right chromatogram the peak at around 8.4 ppm ( $\text{S}'$ ) shows the presence of an amide bond.

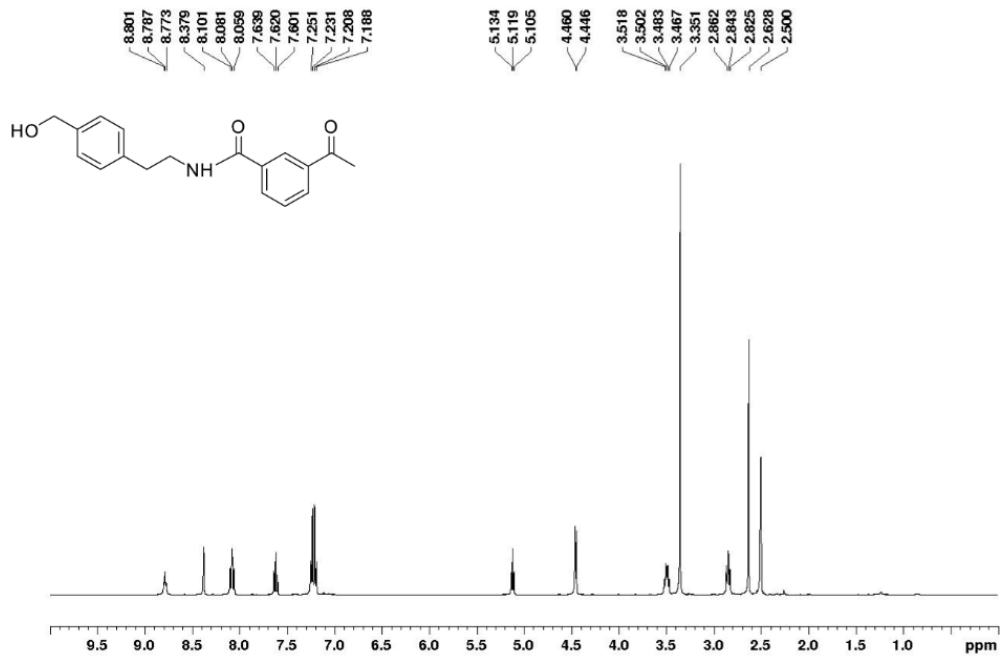


Figure S11. <sup>1</sup>H-NMR of synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide (standard).

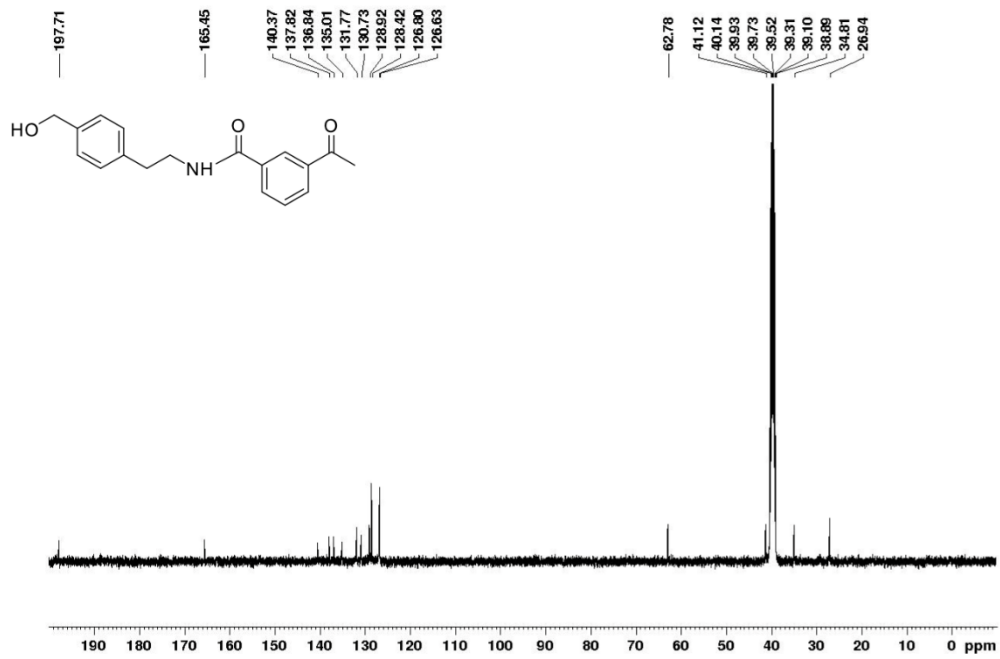


Figure S12. <sup>13</sup>C-NMR of synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide (standard).



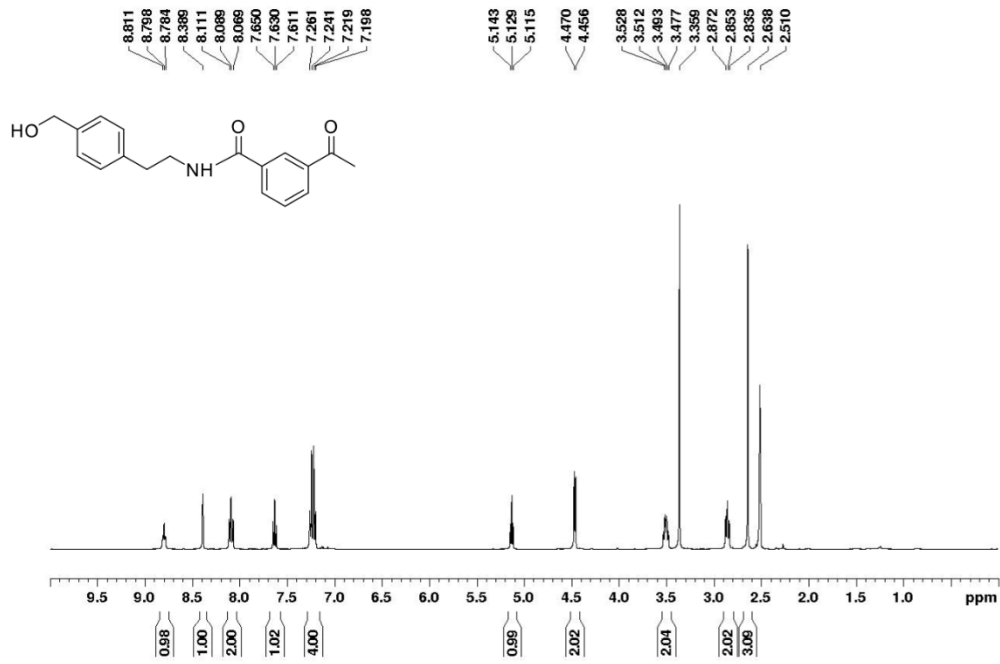


Figure S13. <sup>1</sup>H-NMR of biocatalytic synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide

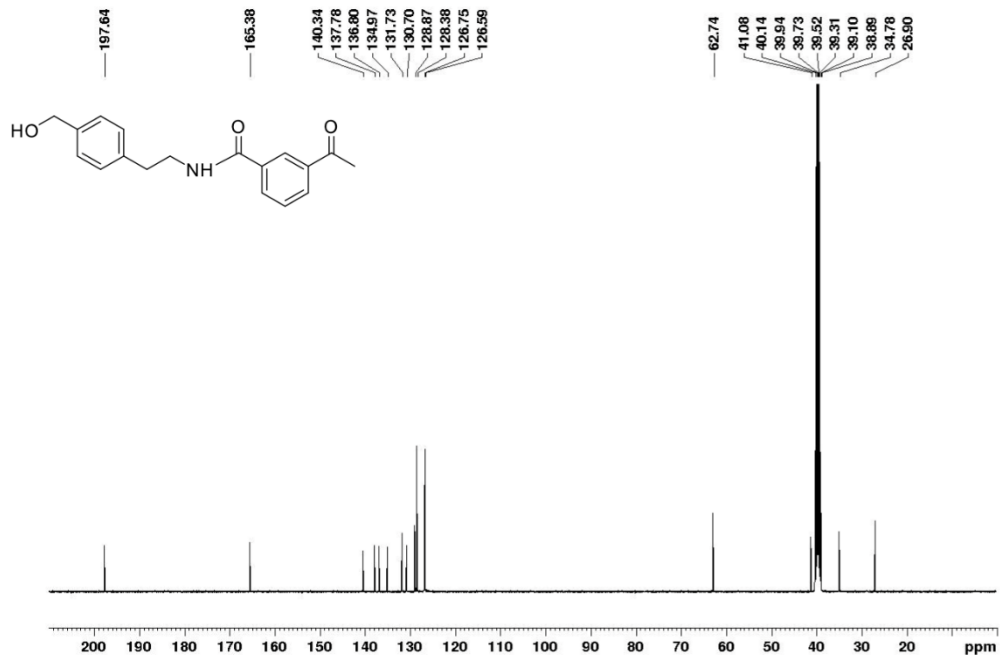


Figure S14. <sup>13</sup>C-NMR of biocatalytic synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide

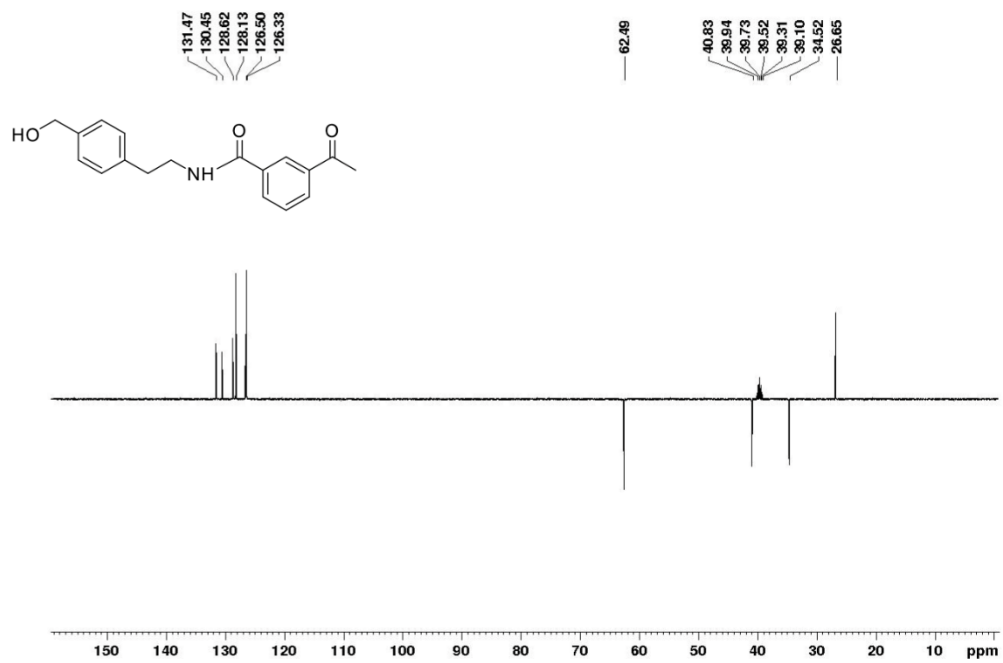


Figure S15. DEPT spectra of biocatalytic synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide

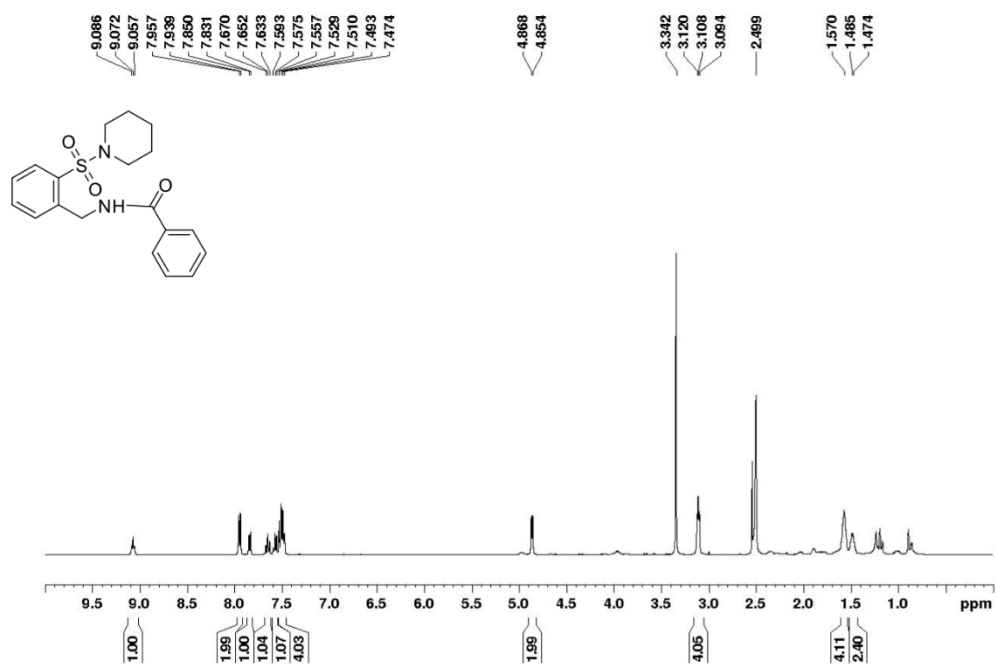


Figure S16. <sup>1</sup>H-NMR of synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide (standard).

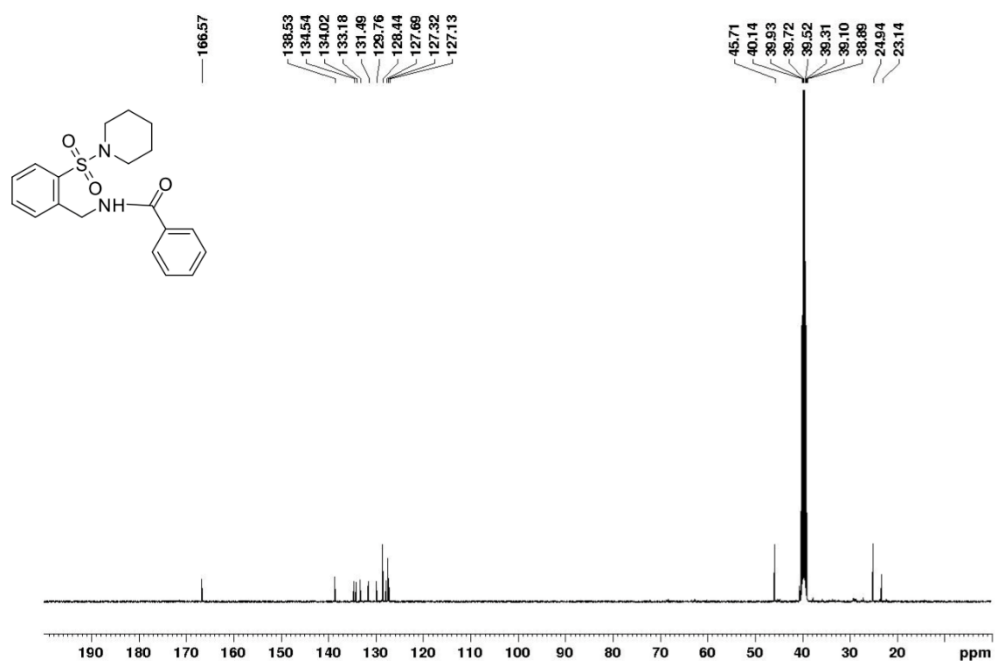


Figure S17. <sup>13</sup>C-NMR of synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide (standard).

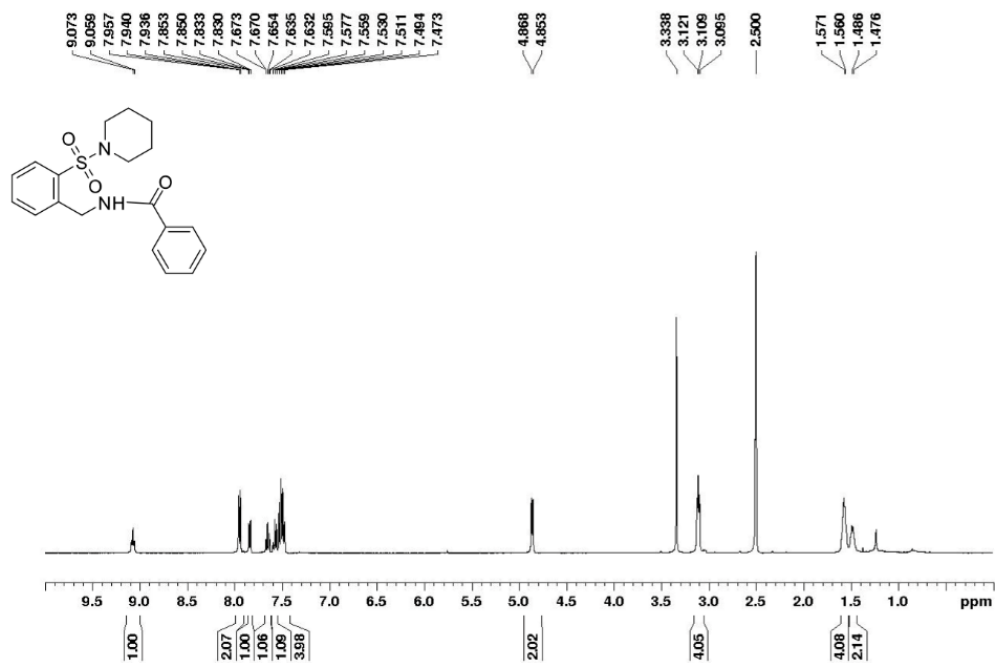


Figure S18 <sup>1</sup>H-NMR of biocatalytic synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide

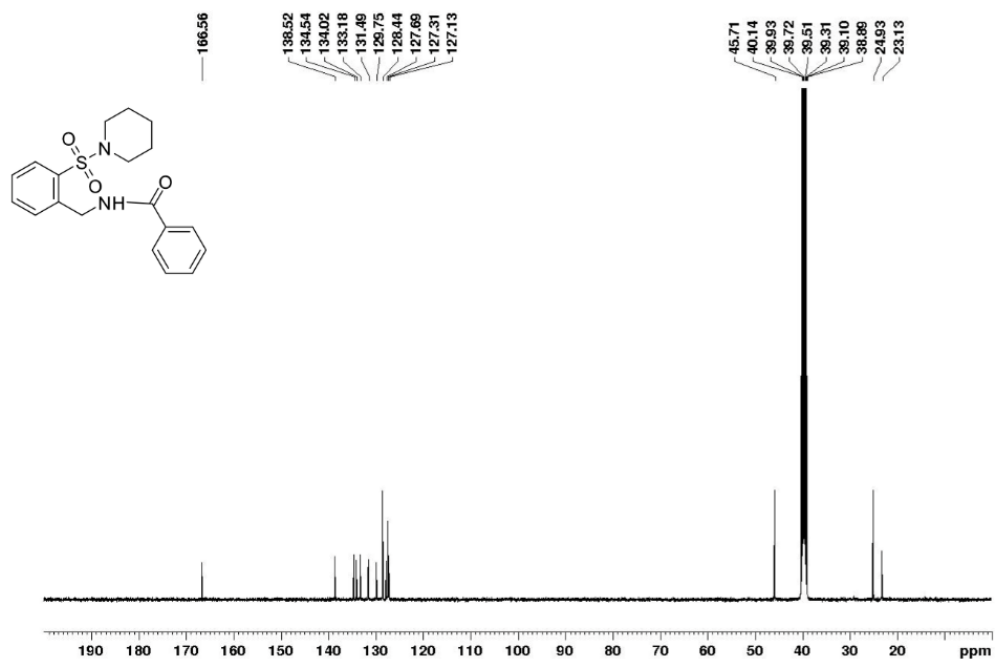


Figure S19 <sup>13</sup>C-NMR of biocatalytic synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide

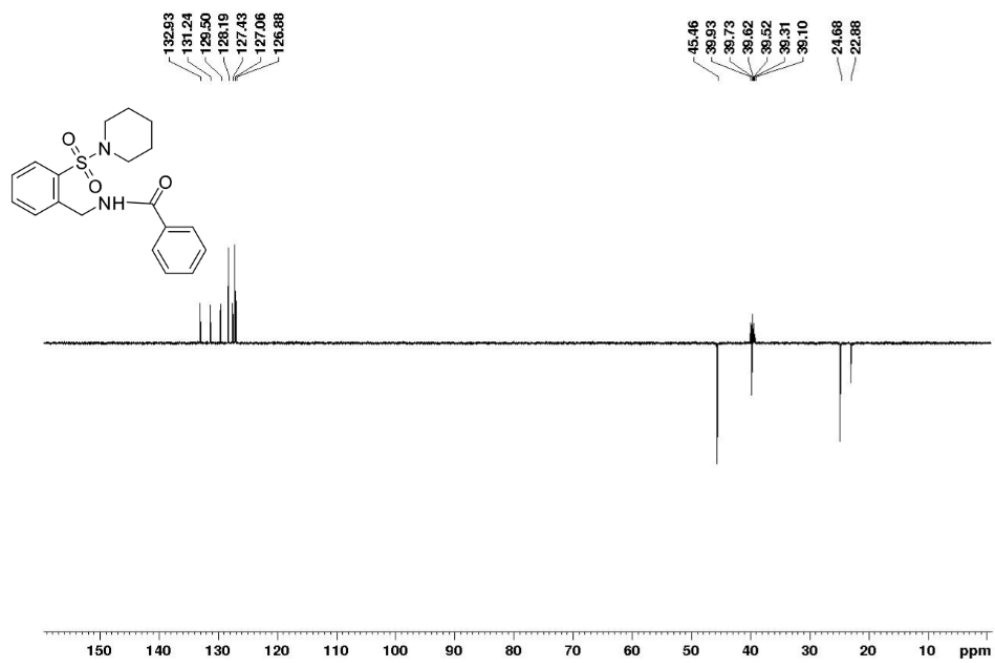


Figure S20. DEPT spectra of biocatalytic synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide

### 305 nm UV absorption 3-hydroxybenzoic acid dilution series

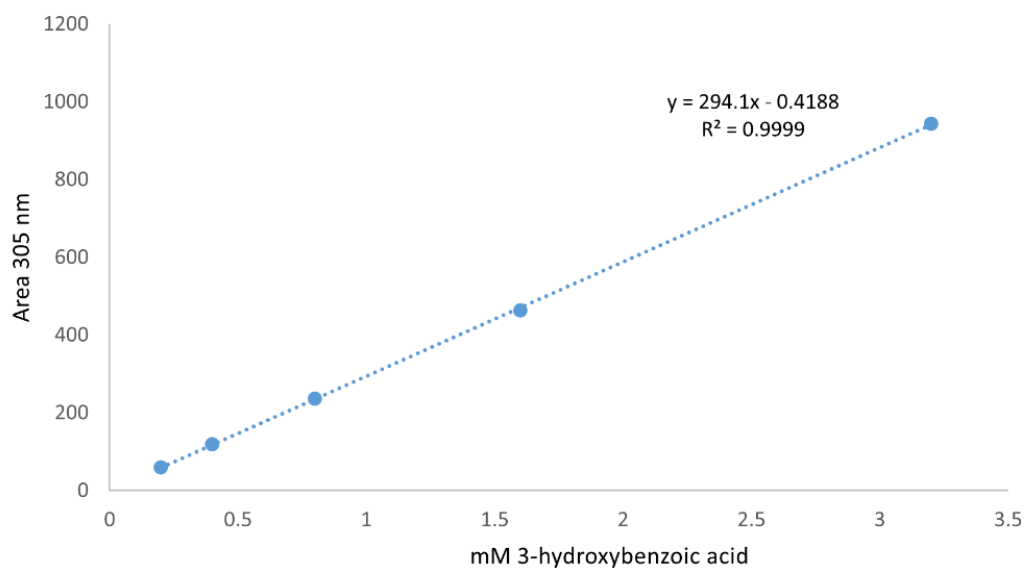


Figure S21. 305 nm UV absorption calibration curve of different 3-hydroxybenzoic acid concentrations. In the equation, y is the 305 nm peak area of the acid, and x is the mM of the acid in the HPLC sample.

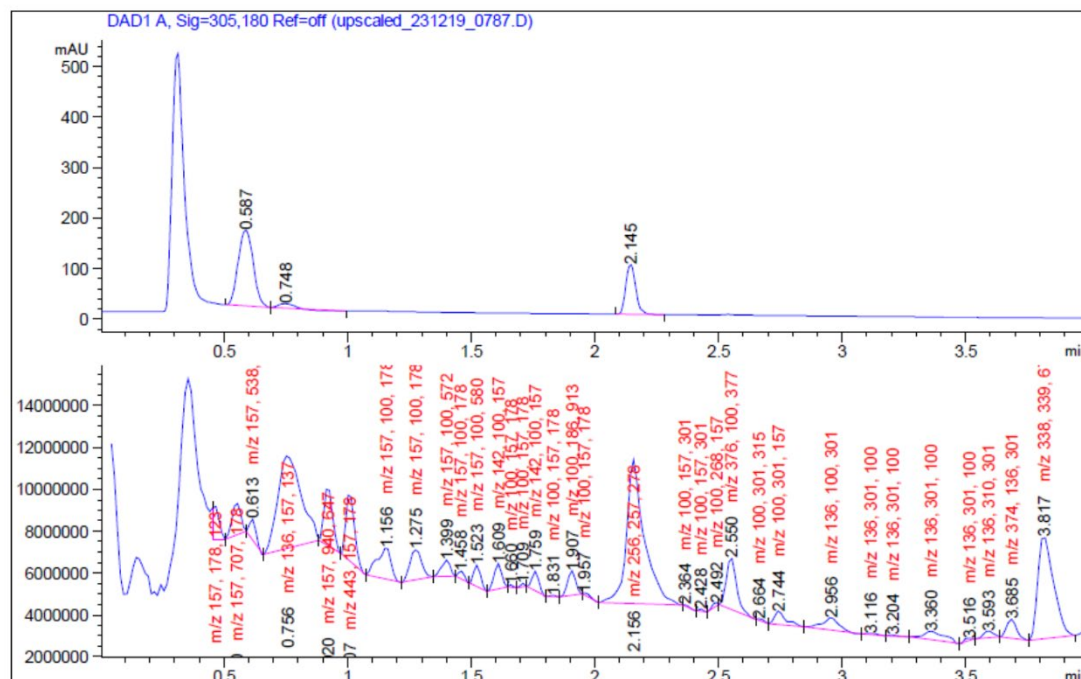


Figure S22. HPLC-MS (positive ionization mode, 305 nm) chromatogram of a sample taken from the upscaled biocatalytic synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide 60 h after the start. At 0.587 min is 3-hydroxybenzoic acid, 0.748 min is 3-phenylpropylamine, and at 2.145 min is the 3-Hydroxy-N-(3-phenylpropyl)benzamide, with its expected mass of 256 m/z. For reference, all mass peaks throughout the elution are shown.

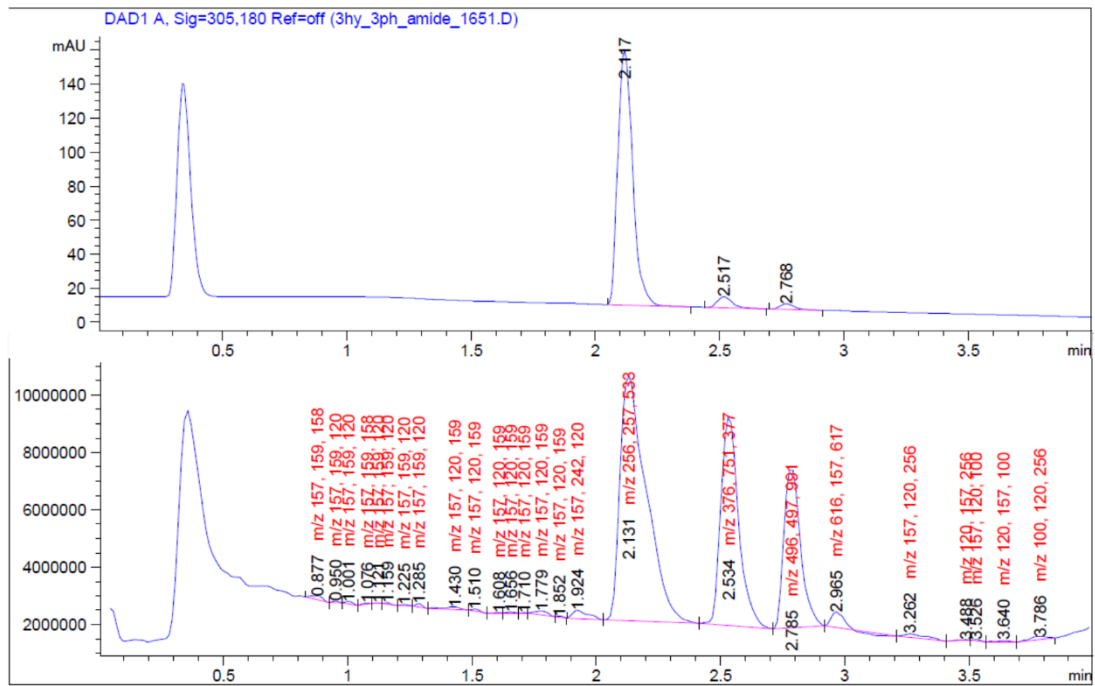


Figure S23. HPLC-MS (positive ionization mode, 305 nm) chromatogram of 3-Hydroxy-N-(3-phenylpropyl)benzamide standard. At 2.117 min is the amide, with its expected mass of 256 *m/z*. For reference, all mass peaks throughout the elution are shown.

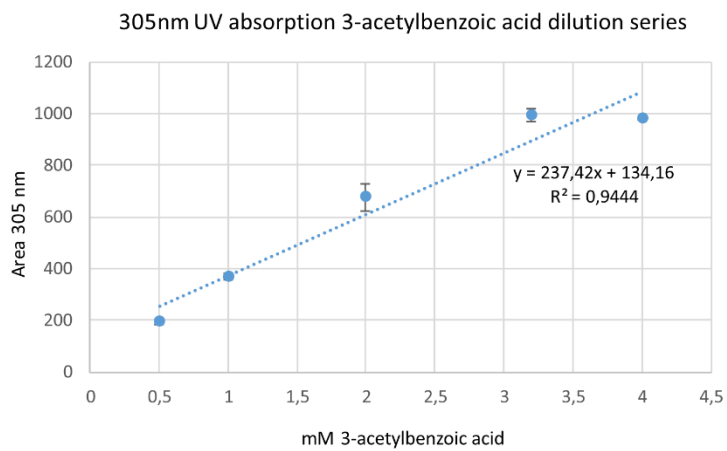


Figure S24. 305 nm UV absorption calibration curve of different 3-acetylbenzoic acid concentrations. In the equation, *y* is the 305 nm peak area of the acid, and *x* is the mM of the acid in the HPLC sample.

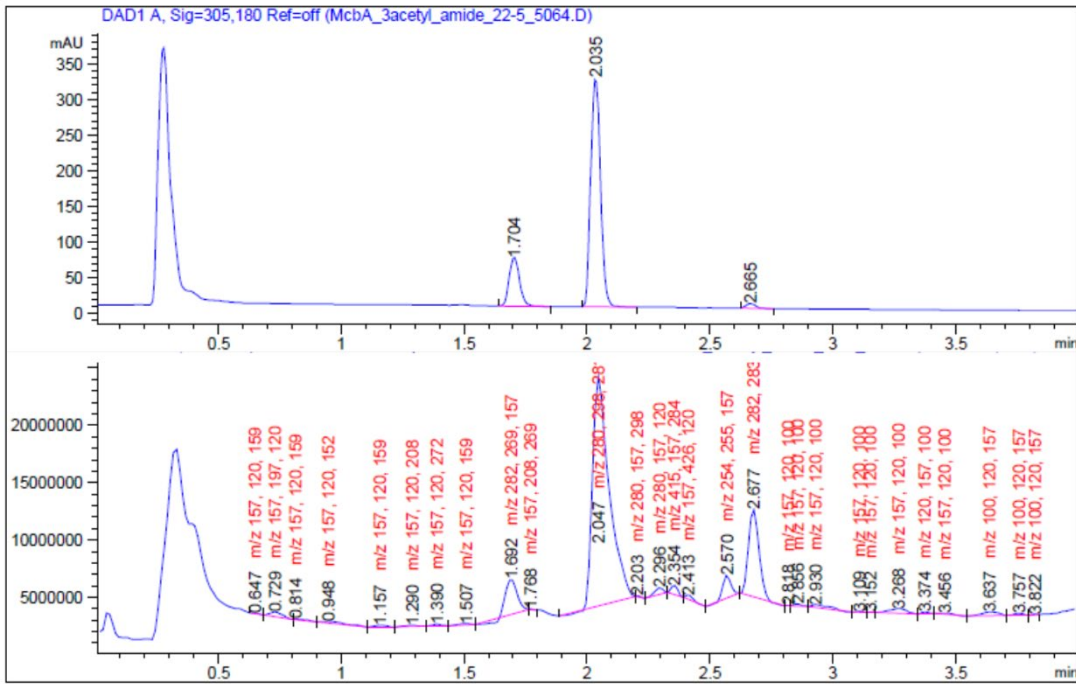


Figure S25. HPLC-MS (positive ionization mode, 305 nm) chromatogram of a sample taken from the upscaled biocatalytic synthesis of 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide 40 h after the start. At 1.704 min is 3-acetylbenzoic acid, and at 2.035 min is 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide, with its expected mass of 298  $m/z$ . For reference, all mass peaks throughout the elution are shown.

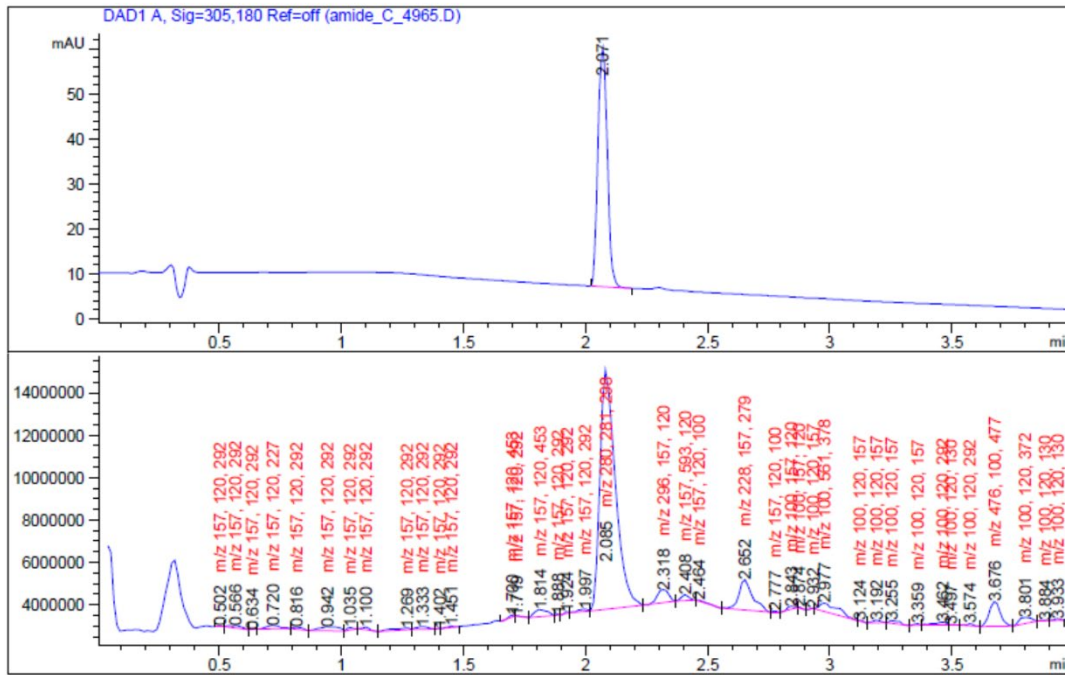


Figure S26. HPLC-MS (positive ionization mode, 305 nm) chromatogram of 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide standard. At 2.071 min is the amide, with its expected mass of 298  $m/z$ . For reference, all mass peaks throughout the elution are shown.



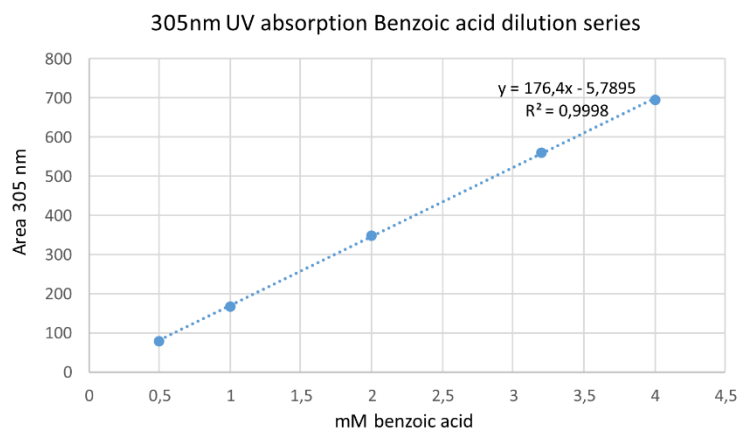


Figure S27. 305 nm UV absorption calibration curve of different benzoic acid concentrations. In the equation, y is the 305 nm peak area of the acid, and x is the mM of the acid in the HPLC sample.

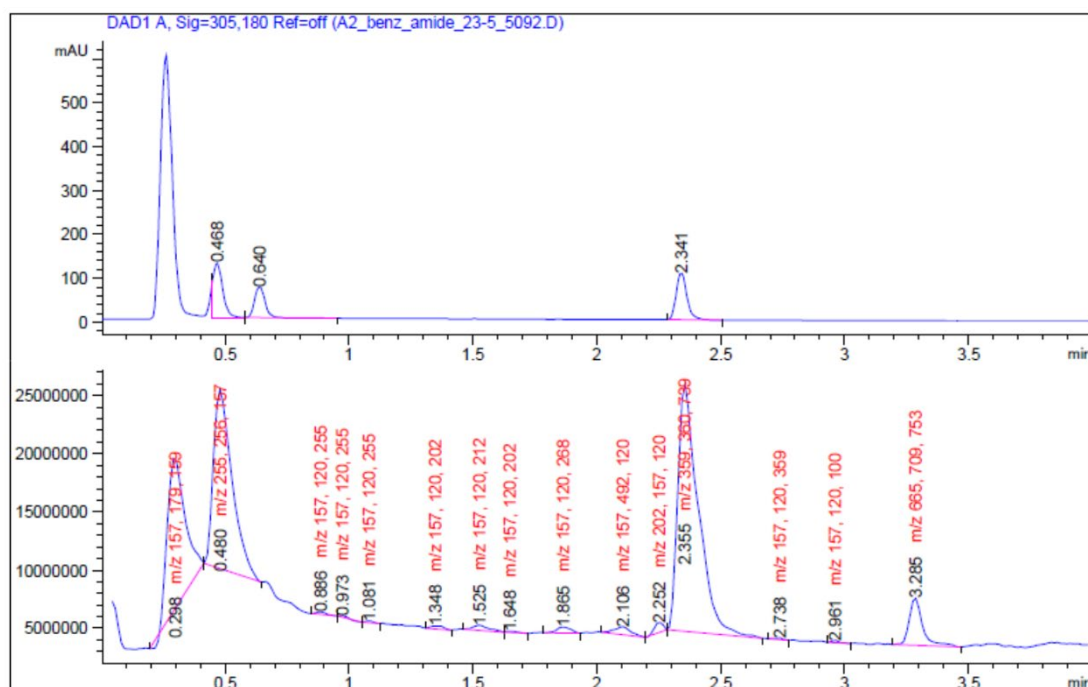


Figure S28. HPLC-MS (positive ionization mode, 305 nm) chromatogram of a taken from the upscaled biocatalytic synthesis of N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide 60 h after the start. At 0.640 min is benzoic acid, and at 2.341 min is N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide, with its expected mass of 359 m/z. For reference, all mass peaks throughout the elution are shown.

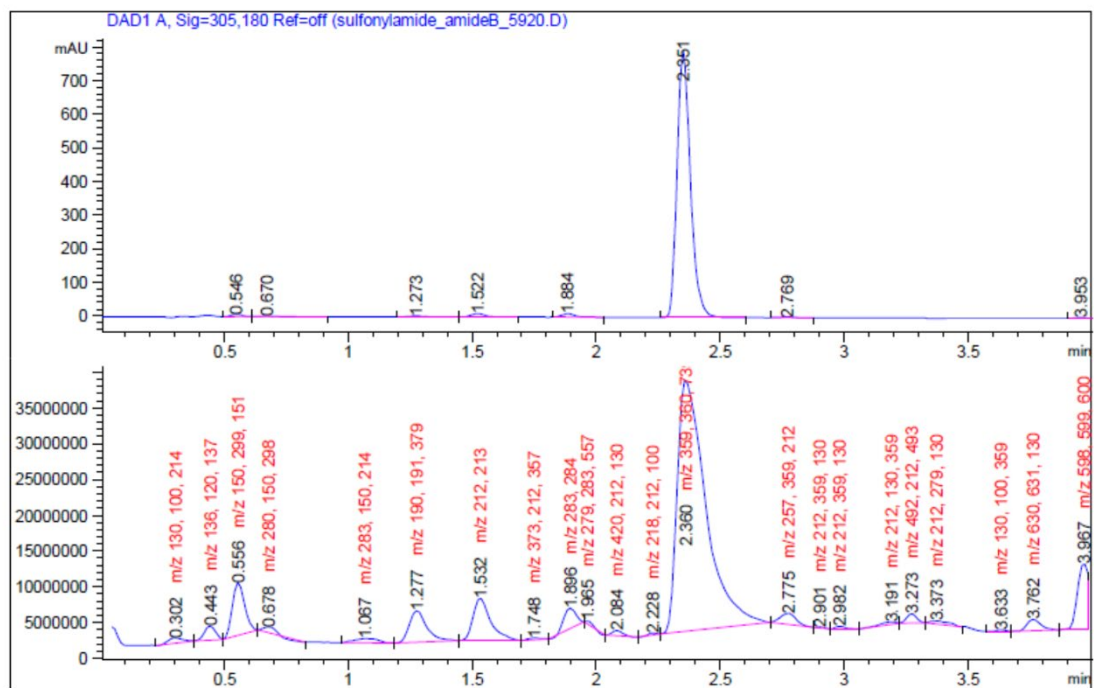


Figure S29. HPLC-MS (positive ionization mode, 305 nm) chromatogram of N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide standard. At 2.360 min is the amide, with its expected mass of 359  $m/z$ . For reference, all mass peaks throughout the elution are shown.

Table S6. Derivation of USEtox input parameters' mean and standard deviation from in-silico model prediction and uncertainty (described by the reported log RMSE).

USEtox 2.0 input	Description	Prediction model	Mean value	Standard deviation	Comment
pKa.gain [-]	Acid dissociation constant (acid reaction)	OPERA pKa_a [-]	$pKa_a$	2.0	
pKa.loss [-]	Acid dissociation constant (basic reaction)	OPERA pKa_b [-]	$pKa_b$	1.7	
Kow [-]	Octanol-water partition ratio	OPERA LogP [-]	$10^{LogP}$	0.69	
Pvap25 [Pa]	Vapor pressure at 25C	OPERA VP [log(mm/Hg)]	$133.3 \cdot 10^{LogVP}$	1.08	with 133.3 Pa/mm.Hg <sup>-1</sup>
Sol25 [mg/L]	Water solubility at 25C	OPERA WS [log(mol/L)]	$MW \cdot 10^{LogWS} \cdot 1000$	0.81	MW = molecular weight
kdegA [1/s]	Degradation rate in air	OPERA LogOH [log(cm <sup>3</sup> /mol.s)]	$\frac{1.5e^6}{2} \cdot 10^{LogOH}$	1.14	with 1.5e <sup>6</sup> mol OH/cm <sup>3</sup>
kdegW [1/s]	Degradation rate in water	OPERA BioHL [log(d)]	$\frac{\ln(2)}{10^{BioHL} \cdot 86400}$	0.26	
kdegSd [1/s]	Degradation rate in sediment	OPERA BioHL [log(d)]	$\frac{\ln(2)}{9 \cdot 10^{BioHL} \cdot 86400}$	1.5 · 0.26	with water-to-sediment extrapolation factor = 9 <sup>a)</sup> , with uncertainty increased by a factor 1.5 for intermedia extrapolation
kdegSI [1/s]	Degradation rate in soil	OPERA BioHL [log(d)]	$\frac{\ln(2)}{2 \cdot 10^{BioHL} \cdot 86400}$	1.5 · 0.26	with water-to-soil extrapolation factor = 2 <sup>a)</sup> , with uncertainty increased by a factor 1.5 for intermedia extrapolation
LogHC50 [log(mg/L)]	50% hazard concentration for freshwater species	ECOSAR ChV Fish, ChV Daphnid, ChV Green Algae [mg/L]	$\frac{1}{n} \sum_{i=1}^n \log_{10} \frac{ChV_i}{0.4}$ <i>i</i> = [1: fish, 2: daphnid, 3: algae]	1.4 <sup>b)</sup>	with EC50 = ChV/0.4 <sup>c)</sup> assuming ChV~NOEC
ED50 non-cancer, ingestion [kg/lifetime]	50% chronic non-cancer effect dose via ingestion	CTV RfD NOAEL [mg/kg.d]	$4.1 \cdot 9 \cdot 10^{RfD NOAEL} \cdot 70 \cdot 70 \cdot 365 \cdot 1e^{-6}$	0.9	with interspecies human-to-rat extrapolation factor = 4.1 <sup>d)</sup> , NOAEL-to-EC50 extrapolation factor = 9 <sup>d)</sup> , human bodyweight = 70 kg, human lifetime = 70 years, 1 year = 365 days, 1 mg = 1e <sup>-6</sup> kg
ED50 non-cancer, inhalation [kg/lifetime]	50% chronic non-cancer effect dose via inhalation	CTV RfD NOAEL [mg/kg.d]	$ED50_{ingestion}^{non-cancer}$	1.5 · 0.9	with uncertainty increased by a factor 1.5 for 1:1 extrapolation from ingestion to inhalation
ED50 cancer, ingestion [kg/lifetime]	50% chronic cancer effect dose via ingestion	CTV OSF [risk per mg/kg.d]	$0.8 \cdot \frac{1}{OSF} \cdot 70 \cdot 70 \cdot 365 \cdot 1e^{-6}$	1.2	with 1/q*-to-ED50 conversion factor = 0.8 <sup>d)</sup> , human bodyweight = 70 kg, human lifetime = 70 years, 1 year = 365 days, 1 mg = 1e <sup>-6</sup> kg
ED50 cancer, inhalation [kg/lifetime]	50% chronic cancer effect dose via inhalation	CTV IUR [risk per mg/m <sup>3</sup> ]	$0.8 \cdot \frac{1}{IUR} \cdot 13 \cdot 70 \cdot 365 \cdot 1e^{-6}$	1.23	with 1/q*-to-ED50 conversion factor = 0.8 <sup>d)</sup> , inhalation rate = 13 m <sup>3</sup> /d, human lifetime = 70 years, 1 year = 365 days, 1 mg = 1e <sup>-6</sup> kg
BAF fish [-]	Bioaccumulation factor in fish	OPERA BCF [-]	$10^{LogBCF}$	0.55	

a) US EPA EPISuite (2012), b) Ping Hou et al., "Estimate Ecotoxicity Characterization Factors for Chemicals in Life Cycle Assessment Using Machine Learning Models," Environment International 135 (2020): 105393-, c) Nicolò Aurisano et al., "Extrapolation Factors for Characterizing Freshwater Ecotoxicity Effects," Environmental Toxicology and Chemistry 38, no. 11 (2019): 2568–82., d) Peter Fantke (Ed.) et al., "USEtox® 2.0 Documentation (Version 1.1)", <http://usetox.org>.

Table S7. Overview of uncertainty factors applied onto the model-specific uncertainty (described by the log RMSE) based on assessing the model applicability for each chemical.

Model suite	Applicability domain	Uncertainty factors
OPERA	Local applicability domain based on structural similarity of query chemical to five nearest neighbors in training data set, ranging from 0 to 1	Inside domain ( $AD \geq 0.6$ ): 1 Inside extended domain ( $0.4 \leq AD < 0.6$ ): 1.5 Outside domain ( $AD < 0.4$ ): 2
CTV	Global applicability domain based on Z-score, which corresponds to the number of standard deviations	Inside domain ( $AD \leq 1$ ): 1 Inside extended domain ( $1 < AD \leq 3$ ): 1.5 Outside domain ( $AD \geq 3$ ): 2 Imputed value (contains metals/metalloids): 3
ECOSAR	"DomainOfApplicability" flag (no continuous value)	Inside domain (no AD flag): 1 Outside domain (AD flag): 2

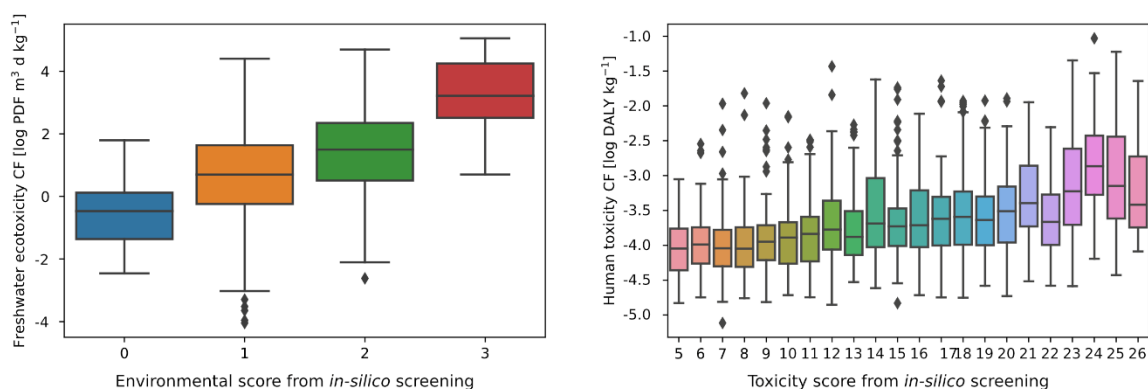


Figure S30. USEtox characterization factors vs *in-silico* screening scores of safechems and subset of filtered out aromatic acids and amines A) Median freshwater ecotoxicity USEtox characterization factor of the in-silico screening environmental score of safechems and sample of 408 aromatic amines and 448 aromatic acids that were filtered out. B) Median human toxicity USEtox characterization factor of in-silico screening toxicity score of safechems and sample of 408 aromatic amines and 448 aromatic acids that were filtered out.

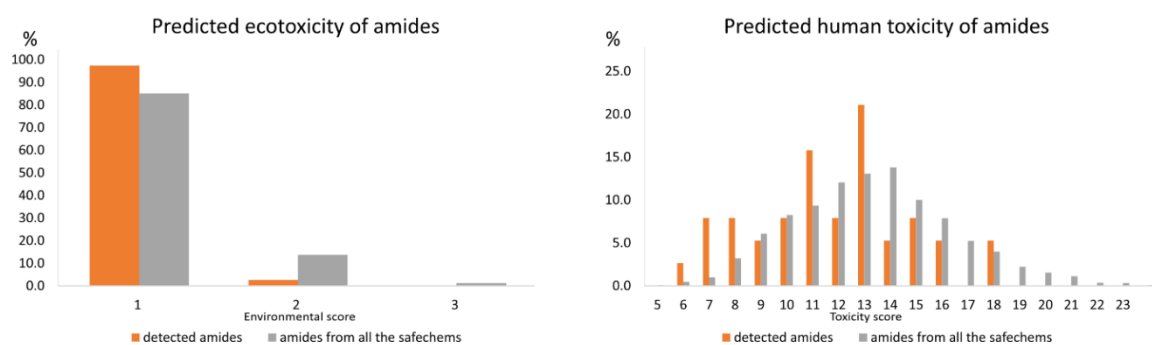


Figure S31. Predicted toxicity and environmental score of detected amides from the experimental coupling considering only a subset of possible safechems in comparison to all possible amides from the whole safechem panel.

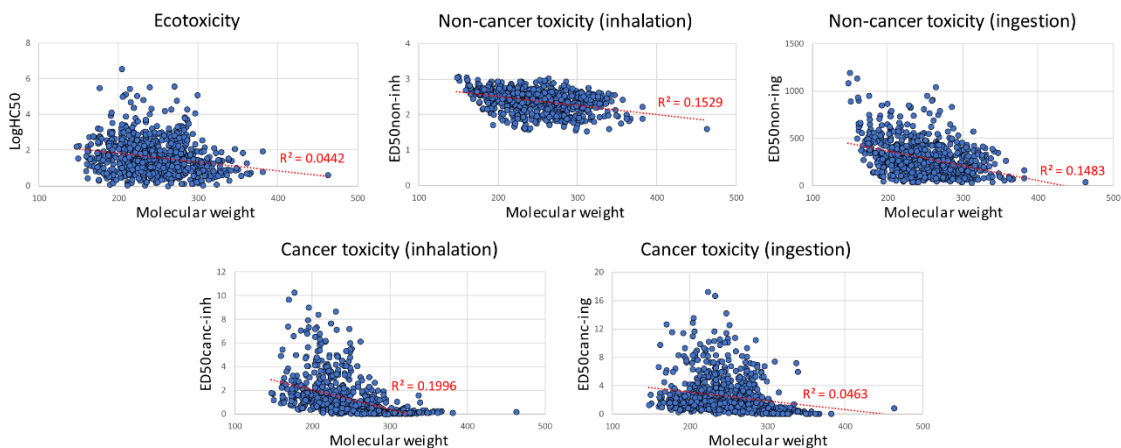


Figure S32. Toxic effects coupled to the molecular weight of amines and acids which were filtered out in the *in silico* filtering process (figure 3). The lower the value (y-axis), the more toxic. HC50 is the hazardous concentration of a chemical at which 50% of the species in an aquatic ecosystem are exposed to the chemical above their tolerance concentration, LogHC50 (log mg/L). ED50non = Human-equivalent lifetime dose per person that causes a non-cancer disease probability of 50% via either inhalation or digestion(kg/lifetime), ED50canc = Human-equivalent lifetime dose per person that causes a cancer disease probability of 50% via either inhalation or digestion (kg/lifetime).

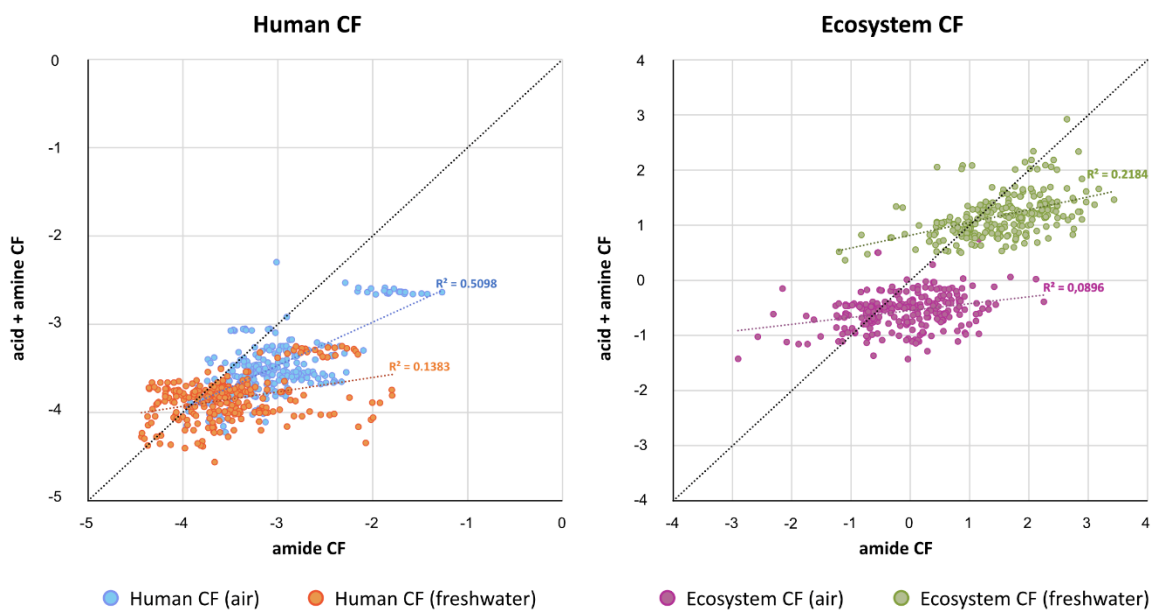


Figure S33. Scatter plot of the USEtox CFs of amides with *in silico* hazard score equal to or lower than ten (figure 3) compared to the sum of the CFs of their constituent acid and amine moieties. The results demonstrate a clear deviation between the amide CF and the combined CFs of the acid and amine. This observation underscores the importance of directly evaluating the environmental impact of the final product, rather than solely relying on the CFs of its building blocks. The amides in the plot passed the initial filtering, and the acids and amines are their safechem building blocks.

Table S8. Conversion by McbA (in %). The first column refers to the acids, and the first row refers to the amines. Conversion is calculated from the amide peak area divided by total acid and amide peak area (DAD) and based on two replicates.

McbA												
	1b	2b	3b	4b	5b	6b	7b	8b	9b	10b	11b	12b
1a	23.8 ± 0.5	31.2 ± 0.9	8.0 ± 0.1	27.6 ± 0.5	0 ± 0	0 ± 0	41.7 ± 0.0	0 ± 0	29.0 ± 3.2	34.3 ± 0.2	1.5 ± 0.1	34.7 ± 9.0
2a	5.0 ± 0.1	2.5 ± 0.0	0.4 ± 0.0	1.5 ± 0.0	0 ± 0	0.1 ± 0.0	0.8 ± 0.1	1.5 ± 0.0	1.7 ± 0.1	2.6 ± 0.7	0.2 ± 0.0	1.1 ± 0.1
3a	2 ± 0.0	1.8 ± 0.0	0.2 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.6 ± 0.0	0.4 ± 0.0	0.2 ± 0.0	0 ± 0	0.2 ± 0.0
4a	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
5a	0 ± 0	0.5 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.4 ± 0.0	0 ± 0	0.2 ± 0.0	0 ± 0	0.2 ± 0.0
6a	29.9 ± 0.5	38.1 ± 1.9	7.0 ± 0.2	12.4 ± 0.3	2.3 ± 0.2	0.4 ± 0.1	10.4 ± 0.3	10.2 ± 0.1	11.6 ± 0.2	6.7 ± 0.3	10.6 ± 0.3	10.4 ± 0.7
7a	1.6 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
8a	0 ± 0	4.7 ± 0.3	0.6 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	2.7 ± 0.0	1 ± 0.0	0.8 ± 0.1	0 ± 0	0.5 ± 0.0
9a	1.4 ± 0.1	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
10a	15.1 ± 0.9	37.0 ± 1.0	10.1 ± 0.3	19.0 ± 0.2	1.1 ± 0.2	0.1 ± 0.0	8.2 ± 0.7	12.0 ± 0.1	13.0 ± 0.1	9.8 ± 0.3	4.8 ± 0.4	9.1 ± 0.7
11a	7.3 ± 0.3	4.9 ± 0.2	0 ± 0	1.2 ± 0.1	0 ± 0	0 ± 0	0.3 ± 0.0	1.6 ± 0.0	1.3 ± 0.0	0.5 ± 0.0	0 ± 0	0.5 ± 0.0
12a	1.5 ± 0.1	0.3 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.4 ± 0.0	0 ± 0	0 ± 0
13a	4.5 ± 0.0	3.8 ± 0.1	1.9 ± 0.1	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.8 ± 0.0	0.8 ± 0.0	0.4 ± 0.0	0 ± 0	0 ± 0
14a	6.3 ± 0.0	4.5 ± 0.2	0.3 ± 0.0	1.1 ± 0.0	0 ± 0	0 ± 0	0.6 ± 0.0	1.2 ± 0.1	1.9 ± 0.1	0.8 ± 0.0	0.6 ± 0.1	0.9 ± 0.2
15a	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.6 ± 0.2	0 ± 0	0 ± 0	0 ± 0	0 ± 0

Table S9. Conversion by A1 (in %). The first column refers to the acids, and the first row refers to the amines. Conversion is calculated from the amide peak area divided by total acid and amide peak area (DAD) and based on two replicates.

A1												
	1b	2b	3b	4b	5b	6b	7b	8b	9b	10b	11b	12b
1a	21.9 ± 0.0	17.8 ± 0.2	3.3 ± 0.0	8.0 ± 0.2	0 ± 0	0.7 ± 0.0	0 ± 0	0 ± 0	11.2 ± 0.5	3.4 ± 0.1	18.3 ± 0.3	9.6 ± 2.9
2a	1.0 ± 0.0	0.2 ± 0	0 ± 0	0 ± 0	0 ± 0	0.1 ± 0.0	0 ± 0	0 ± 0	0.2 ± 0.0	0 ± 0	0.9 ± 0.0	0 ± 0
3a	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
4a	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
5a	0 ± 0	0.2 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
6a	4.2 ± 0.0	2.4 ± 0.0	0.3 ± 0.0	0.6 ± 0.3	0.2 ± 0.0	0.7 ± 0.2	0 ± 0	0.7 ± 0.2	0 ± 0	0.2 ± 0.0	4.8 ± 0.8	0 ± 0
7a	1.5 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
8a	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
9a	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
10a	5.4 ± 0.0	1.9 ± 0.2	0.8 ± 0.0	0.9 ± 0.1	0 ± 0	0.2 ± 0.0	0.8 ± 0.0	1.3 ± 0.1	1.4 ± 0.1	0.1 ± 0.0	3.4 ± 0.0	3.6 ± 0.1
11a	0.3 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
12a	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
13a	0.5 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
14a	0.2 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.1 ± 0.0	0 ± 0
15a	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.6 ± 0.6	0 ± 0	0 ± 0	0 ± 0	0 ± 0

Table S10. Conversion by A2 (in %). The first column refers to the acids, and the first row refers to the amines. Conversion is calculated from the amide peak area divided by total acid and amide peak area (DAD) and based on two replicates.

A2												
	1b	2b	3b	4b	5b	6b	7b	8b	9b	10b	11b	12b
1a	6.2 ± 0.6	1.7 ± 0.0	22.4 ± 1.2	13.4 ± 0.3	0 ± 0	0 ± 0	16.5 ± 1.1	0 ± 0	1.2 ± 0.0	35.9 ± 0.7	0.9 ± 0.0	51.5 ± 15.6
2a	0.3 ± 0.0	0 ± 0	3.2 ± 0.7	1.2 ± 0.1	0 ± 0	0 ± 0	2.3 ± 0.1	5.1 ± 0.2	0.3 ± 0.0	6.5 ± 0.1	0 ± 0	14.1 ± 0.3
3a	0 ± 0	0 ± 0	0.8 ± 0.2	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.2 ± 0.0	0 ± 0	0.4 ± 0.0	0 ± 0	2.6 ± 0.2
4a	0 ± 0	0 ± 0	0.5 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
5a	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
6a	1.3 ± 0.0	0.9 ± 0.1	6.1 ± 0.0	2.9 ± 0.0	0 ± 0	0 ± 0	7.8 ± 0.0	6.4 ± 1.0	0 ± 0	5.9 ± 0.1	0 ± 0	24.6 ± 2.3
7a	1.3 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
8a	0 ± 0	0 ± 0	12.2 ± 1.7	0 ± 0	0 ± 0	0 ± 0	4.8 ± 0.2	9.9 ± 0.5	0 ± 0	7.7 ± 0.0	0 ± 0	8.9 ± 0.6
9a	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
10a	1.2 ± 0.0	0.2 ± 0.0	10.7 ± 0.5	3.2 ± 0.0	0.1 ± 0.0	0 ± 0	5.7 ± 0.0	4.9 ± 0.3	0.6 ± 0.0	5.9 ± 0.1	0 ± 0	9.1 ± 0.6
11a	1.1 ± 0.0	0 ± 0	9.0 ± 0.6	0.9 ± 0.0	0 ± 0	0 ± 0	10.6 ± 1.4	1.5 ± 0	0.6 ± 0.0	8.0 ± 1.8	0 ± 0	4.6 ± 0.0
12a	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.5 ± 0.1	0 ± 0	0 ± 0
13a	0.8 ± 0.0	0 ± 0	10.5 ± 0.9	0 ± 0	0 ± 0	0 ± 0	18.4 ± 2.2	1.9 ± 0.0	0.6 ± 0.0	6.1 ± 0.1	0 ± 0	6.5 ± 0.0
14a	0.7 ± 0.0	0 ± 0	4.1 ± 0.5	1.2 ± 0.4	0 ± 0	0 ± 0	5.0 ± 0.0	3.5 ± 0.1	1.5 ± 0.1	10.9 ± 0.1	0 ± 0	9.7 ± 0.8
15a	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	2.1 ± 0.1	0 ± 0	0 ± 0	0 ± 0	0 ± 0



Table S11. List of safechem amines from *in silico* screening.

Compound	SMILES	Hazard factor	Ecohazard factor	used in experimental synthesis
Aniline	<chem>NC1=CC=CC=C1</chem>	2	0	yes
amine8634	<chem>NCc1cccc1S(=O)(=O)N1CCCC1</chem>	3	1	yes
amine729	<chem>NC1(c2ccccc2)CCNC1</chem>	3	1	yes
amine6672	<chem>NC1CCN(S(=O)(=O)c2ccccc2)C1</chem>	3	1	yes
amine6328	<chem>NCCCc1cccc1</chem>	1	1	yes
amine516	<chem>NCCCN1Cc2ccccc2C1</chem>	3	1	yes
amine3327	<chem>NCc1cccc1CN1CCCC1</chem>	2	1	yes
amine3289	<chem>NCCc1ccc(CCN)cc1</chem>	4	1	yes
amine2540	<chem>NCc1cccc1CN1CCCC1CO</chem>	3	1	yes
amine2119	<chem>NCCc1ccc(CO)cc1</chem>	2	1	yes
amine14086	<chem>Cc1cccc1CCCN</chem>	2	1	yes
amine13521	<chem>CN(CCN)Cc1cccc1</chem>	1	1	yes
amine13034	<chem>NCCc1cccc1O</chem>	3	0	yes
amine12828	<chem>NCc1ccc(S(=O)(=O)NCC2CCCO2)cc1</chem>	2	1	yes
amine1217	<chem>COCCNC(=O)C(N)Cc1cccc1</chem>	4	1	yes
amine11946	<chem>CN(C)S(=O)(=O)Cc1ccc(CN)cc1</chem>	3	1	yes
2-aminopyridine	<chem>NC1=CC=CC=N1</chem>	2	1	yes
amine9676	<chem>NC1(c2ccccc2)CCC(=O)NC1</chem>	4	1	
amine8732	<chem>NC1CCCS(=O)(=O)c2ccccc21</chem>	3	1	
amine8721	<chem>COc1ccc(S(=O)(=O)N2CCC(N)CC2)cc1</chem>	2	1	
amine8666	<chem>NC1CCc2cc(C(=O)O)ccc2C1</chem>	4	1	
amine8369	<chem>NC1CCCN(c2ccc(C(=O)O)cc2)C1</chem>	3	1	
amine7654	<chem>CC(c1ccc(S(C)(=O)=O)cc1)N1CCC(N)CC1</chem>	3	1	
amine7539	<chem>NCCCCc1cccc1</chem>	1	1	
amine6912	<chem>C=Cc1ccc(CCN)cc1</chem>	2	1	
amine6247	<chem>Cc1cc(C)cc(C(O)CN)c1</chem>	2	1	
amine5745	<chem>CN(CCN)S(=O)(=O)c1ccc(F)cc1</chem>	4	1	
amine5247	<chem>NCCCCc1cccc(O)c1</chem>	2	1	

amine5086	<chem>COC1ccc(CN)cc1</chem>	3	1	
amine4400	<chem>COC1ccc(CCN)c(F)c1</chem>	3	1	
amine429	<chem>NCC(O)c1ccccc1</chem>	1	0	
amine4167	<chem>Cc1cccc(N2CCCC(N)C2=O)c1</chem>	2	1	
amine3884	<chem>NCCN(CC(=O)O)Cc1ccccc1</chem>	4	1	
amine3444	<chem>CC(N)c1ccc(N2CCCS2(=O)=O)cc1</chem>	3	1	
amine3280	<chem>CCN(CC)Cc1ccccc1CN</chem>	4	1	
amine3274	<chem>Cc1ccc(CCN)cc1</chem>	2	1	
amine3222	<chem>CCN(CC)Cc1ccc(CN)cc1</chem>	3	1	
amine3084	<chem>CCN(CCCN)c1ccccc1</chem>	4	1	
amine3076	<chem>CN(CCCN)Cc1ccccc1</chem>	1	1	
amine3005	<chem>COC(=O)C(N)c1cc(C)cc(C)c1</chem>	3	1	
amine2998	<chem>COC(=O)C(N)c1ccc2c(c1)CCO2</chem>	3	1	
amine1943	<chem>NC(c1ccccc1)C1CCOC1</chem>	4	1	
amine14452	<chem>NC1CCCN(C(=O)c2ccccc2)C1</chem>	4	1	
amine14127	<chem>CN(C)CCOc1ccccc1CN</chem>	4	1	
amine14103	<chem>CN(C)c1ccc(CCN)cc1</chem>	4	1	
amine14084	<chem>NCCc1cccc(O)c1</chem>	2	1	
amine13680	<chem>NCCN1Cc2ccccc2C1</chem>	3	1	
amine1335	<chem>NCCc1ccccc1F</chem>	2	1	
amine13327	<chem>COC1ccc(CCN)cc1</chem>	3	1	
amine13246	<chem>NCCc1ccc(O)cc1</chem>	2	1	
amine13174	<chem>NCCCOCCc1ccccc1</chem>	1	1	
amine12952	<chem>NC(c1ccccc1)C1CCCO1</chem>	3	1	
amine12947	<chem>CCN(CC)Cc1ccccc1CN</chem>	3	1	
amine12928	<chem>CN(C)CCOc1ccccc1CN</chem>	4	1	
amine12902	<chem>CNS(=O)(=O)c1cc(F)c(Cl)c(CN)c1</chem>	4	2	
amine12829	<chem>Nc1ccccc1C(N)CC(=O)O)c1</chem>	3	1	
amine1276	<chem>CCc1ccccc1CCN</chem>	1	1	
amine11623	<chem>CC(N)(CO)c1ccccc1N)c1</chem>	3	1	

amine11341	<chem>NCc1cccc2c1CCNC2</chem>	4	1	
amine10961	<chem>NCc1cccc(F)c1</chem>	2	0	
amine10593	<chem>NCC1CCc2ccccc2CC1</chem>	4	1	
amine10438	<chem>NCc1cccc2c1NCCC2</chem>	4	1	

Table S12. List of safechem acids from in silico screening.

Compound	SMILES	Hazard factor	Ecohazard factor	used in experimental synthesis
acid9360	<chem>O=C(O)c1cccc(C(=O)O)c1</chem>	2	0	yes
acid8305	<chem>O=C(O)c1ccccc1</chem>	1	0	yes
acid7954	<chem>O=C(O)c1cccc(O)c1</chem>	2	0	yes
acid6946	<chem>O=C(O)c1ccccc1O</chem>	2	0	yes
acid25277	<chem>O=C(O)c1ccc(F)c(O)c1</chem>	2	0	yes
acid24861	<chem>CC(=O)c1cccc(C(=O)O)c1</chem>	1	0	yes
acid24207	<chem>Cc1cccc(C)c1C(=O)O</chem>	2	0	yes
acid24135	<chem>O=C(O)C(=O)c1ccccc1</chem>	2	0	yes
acid23854	<chem>O=C(O)c1cccc(B(O)O)c1</chem>	3	0	yes
acid20978	<chem>O=C(O)c1ccccc1CO</chem>	3	0	yes
acid18918	<chem>O=C(O)C1CC(O)CN1C(=O)c1ccc2c(c1)CCC2</chem>	2	1	yes
acid18600	<chem>O=C(O)c1ccc2c(c1)COC2</chem>	3	0	yes
acid14665	<chem>CC(=O)c1ccc(C(=O)O)cc1</chem>	2	0	yes
acid11933	<chem>CC(=O)c1ccccc1C(=O)O</chem>	2	0	yes
acid10321	<chem>O=C(O)c1ccc(O)cc1</chem>	2	0	yes

4-methoxybenzoic acid	<chem>COc1ccc(C(=O)O)cc3</chem>	4	0	yes
acid22013	<chem>NC1Cc2ccc(C(=O)O)cc2C1</chem>	2	1	
acid9525	<chem>O=C(O)c1c(F)cccc1F</chem>	2	0	
acid8330	<chem>O=C(O)c1ccc(C(=O)O)cc1</chem>	2	0	
acid7972	<chem>O=C(O)c1ccc2c(c1)CCNCC2</chem>	4	1	
acid7295	<chem>O=C(O)c1ccc2c(c1)CC2</chem>	3	0	
acid4638	<chem>O=C(O)c1ccc2c(c1)CNC2</chem>	4	1	
acid6413	<chem>CC(=O)C(NS(=O)(=O)c1ccc(C)C(C)c1)C(O)O</chem>	4	1	
acid7343	<chem>CCc1cccc1CC(=O)O</chem>	3	0	
acid2256	<chem>CCc1cccc1C(=O)O</chem>	2	0	
acid6407	<chem>CC(=O)C(NS(=O)(=O)c1ccc(C)cc1)C(O)O</chem>	4	1	
acid8776	<chem>Cc1cc(O)cc(C(O)O)c1C(O)O</chem>	4	1	
acid6280	<chem>COc1cc(C=O)ccc1C(O)O</chem>	4	0	
Nicotinic acid	<chem>O=C(O)c1cccnc3</chem>	2	0	
hydrocinnamic acid	<chem>O=C(O)CCc1cccc3</chem>	0	0	

Table S13. The exact masses of detected and potential amides in the experimental coupling in figure 4.

Exact Mass													
		amine11946	amine1217	amine14086	amine2119	amine2540	amine3327	amine6328	amine6672	amine729	amine8634	aniline	Phenethyl-amine
		1b	2b	3b	4b	5b	6b	7b	8b	9b	10b	11b	12b
2-Naphthoic acid	1a	382.1	376.2	303.2	305.1	374.2	358.2	289.2	380.1	294.2	408.2	247.1	275.1
4-methoxy-benzoic acid	2a	362.1	356.2	283.2	285.1	354.2	338.2	269.1	360.1	296.2	388.1	227.1	255.1
acid10321	3a	348.1	342.2	269.1	271.1	340.2	324.2	255.1	346.1	282.1	374.1	213.1	241.1
acid11933	4a	374.1	368.2	295.2	297.1	366.2	350.2	281.1	372.1	308.2	400.1	239.1	267.1
acid14665	5a	374.1	368.2	295.2	297.1	366.2	350.2	281.1	372.1	308.2	400.1	239.1	267.1
acid18600	6a	374.1	368.2	295.2	297.1	366.2	350.2	281.1	372.1	308.2	400.1	239.1	267.1
acid18918	7a	485.2	479.2	406.2	408.2	477.3	461.3	392.2	483.2	419.2	511.2	350.2	378.2
acid23854	8a	376.1	370.2	297.2	299.1	368.2	352.2	283.1	374.1	310.1	402.1	241.1	269.1
acid24207	9a	360.2	354.2	281.2	283.2	352.2	336.2	267.2	358.1	294.2	386.2	225.1	253.1
acid24861	10a	374.1	368.2	295.2	297.1	366.2	350.2	281.1	372.1	308.2	400.1	239.1	267.1
acid25277	11a	366.1	360.1	287.1	289.1	358.2	342.2	273.1	364.1	300.1	392.1	231.1	259.1
acid6946	12a	348.1	342.2	269.1	271.1	352.3	324.2	255.1	346.1	282.1	374.1	213.1	241.1
acid7954	13a	348.1	342.2	269.1	271.1	340.2	324.2	255.1	346.1	282.1	374.1	213.1	241.1
acid8305	14a	332.1	326.2	253.1	255.1	324.2	308.2	239.1	330.1	266.1	358.1	197.1	225.1
acid9360	15a	376.1	370.2	297.1	299.1	368.2	352.2	283.1	374.1	310.1	402.1	241.1	269.1

## Notes and references

1. Steffensky M, Mühlenweg A, Wang ZX, Li SM, Heide L. Identification of the Novobiocin Biosynthetic Gene Cluster of *Streptomyces spheroides* NCIB 11891. *Antimicrob Agents Chemother*. 2000 May;44(5):1214–22.
2. Pojer F, Li SM, Heide L. Molecular cloning and sequence analysis of the clorobiocin biosynthetic gene cluster: new insights into the biosynthesis of aminocoumarin antibiotics. *Microbiology (Reading)*. 2002 Dec;148(Pt 12):3901–11.
3. Galm U, Schimana J, Fiedler HP, Schmidt J, Li SM, Heide L. Cloning and analysis of the simocyclinone biosynthetic gene cluster of *Streptomyces antibioticus* Tü 6040. *Archives of Microbiology*. 2002 Aug 1;178(2):102–14.
4. Schmutz E, Steffensky M, Schmidt J, Porzel A, Li SM, Heide L. An unusual amide synthetase (CouL) from the coumermycin A<sub>1</sub> biosynthetic gene cluster from *Streptomyces rishiriensis* DSM 40489. *Eur J Biochem*. 2003 Nov;270(22):4413–9.
5. Wang ZX, Li SM, Heide L. Identification of the Coumermycin A<sub>1</sub> Biosynthetic Gene Cluster of *Streptomyces rishiriensis* DSM 40489. *Antimicrob Agents Chemother*. 2000 Nov;44(11):3040–8.
6. Kalan L, Gessner A, Thaker MN, Waglechner N, Zhu X, Szawiola A, et al. A Cryptic Polyene Biosynthetic Gene Cluster in *Streptomyces calvus* Is Expressed upon Complementation with a Functional bldA Gene. *Chemistry & Biology*. 2013 Oct;20(10):1214–24.
7. Rui Z, Petříčková K, Škanta F, Pospíšil S, Yang Y, Chen CY, et al. Biochemical and Genetic Insights into Asukamycin Biosynthesis. *Journal of Biological Chemistry*. 2010 Aug;285(32):24915–24.
8. McAlpine JB, Bachmann BO, Pirae M, Tremblay S, Alarco AM, Zazopoulos E, et al. Microbial Genomics as a Guide to Drug Discovery and Structural Elucidation: ECO-02301, a Novel Antifungal Agent, as an Example †. *J Nat Prod*. 2005 Apr;68(4):493–6.
9. Zhang W, Bolla ML, Kahne D, Walsh CT. A Three Enzyme Pathway for 2-Amino-3-hydroxycyclopent-2-enone Formation and Incorporation in Natural Product Biosynthesis. *J Am Chem Soc*. 2010 May 12;132(18):6402–11.
10. Winn M, Rowlinson M, Wang F, Bering L, Francis D, Levy C, et al. Discovery, characterization and engineering of ligases for amide synthesis. *Nature*. 2021 May 20;593(7859):391–8.