Supplementary Information (SI) for Green Chemistry.
This journal is © The Royal Society of Chemistry 2024

Supplementary Information

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

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

[a.] School of Chemistry, Biotechnology and Health, Science for Life Laboratory, KTH Royal Institute of Technology, Tomtebodavägen 23 171 65, Solna, Sweden

[b.] School of Engineering Sciences in Chemistry, Biotechnology and Health, Department of Fibre and Polymer Technology, KTH Royal Institute of Technology, 114 28 Stockholm, Sweden

[c.] Quantitative Sustainability Assessment, Department of Environmental and Resource Engineering, Technical University of Denmark, Bygningstorvet 115, 2800 Kgs. Lyngby, Denmark

[d.] Department of Computer and Systems Sciences, Stockholm University, 164 25 Kista, Sweden

[e.] Discovery Sciences, BioPharmaR&D, AstraZeneca, 431 83 Mölndal, Sweden

[f.] The Research Institute of Sweden RISE, Chemical and pharmaceutical toxicology, 151 36 Södertälje, Sweden

[g.] IVL Swedish Environmental Research Institute, Aschebergsgatan 44, 411 33 Göteborg, Sweden

[h.] Early CVRM Medicinal Chemistry, BioPharmaR&D, AstraZeneca, 431 83 Mölndal, Sweden

Table of Contents

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) ClоL, (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).

	Freshwater ecotoxicity (CF's)		Human toxicity (CF's)	
Type	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.

	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 (1×15 mL) and 2 M HCl (1×15 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. ^{1}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). ^{13}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 $NaHCO_3$, and brine, dried over $MgSO_4$, filtered, and concentrated in vacuo. The crude was purified by preparative

TLC (2%-10% MeOH-$CHCl_3$) to afford the corresponding amide compound.
3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide: R_f= 0.44 (MeOH/$CHCl_3$ (10:90)), Colourless solid, Yield (16 mg, 44%)
^{1}H NMR, 400 MHz, DMSO-d_6: 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). ^{13}C NMR, 400 MHz, DMSO-d_6: 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/$CHCl_3$ (2:98)), Colourless solid, Yield (5.0 mg, 34%)
^{1}H NMR, 400 MHz, DMSO-d_6: 9.07 (t, *J*=5.5Hz, 1H), 7.95 (d, *J*=7.2Hz, 2H), 7.84 (d, *J*= 7.7 Hz, 1H), 7.65 (t, *J*=7.5 Hz, 1H), 7.58 (t, *J*=7.1Hz, 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*=4.3Hz, 2H)
^{13}C NMR, 400 MHz, DMSO-d_6: 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.

	McbA	**A1**	**A2**	**A3**	**A4**
McbA	100.0	70.0	62.4	50.4	40.0
A1	70.0	100.0	85.2	65.0	48.8
A2	62.4	85.2	100.0	75.4	54.8
A3	50.4	65.0	75.4	100.0	66.8
A4	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.

Enzyme	**T_m 1 (°C)**	**T_m 2 (°C)**
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

Enzyme	**Approximate mg protein per g pellet**
McbA	10
A1	7
A2	8
A3	14
A4	8

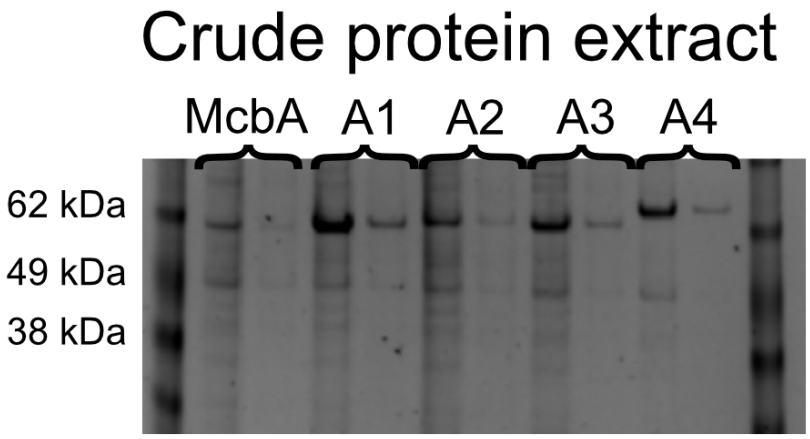


Figure S1. Crude protein content of McbA and ancestor A1-A4. Samples were normalized by OD600, two dilutions of each sample were loaded onto the gel. The molecular weight marker is SeeBlue Plus2.

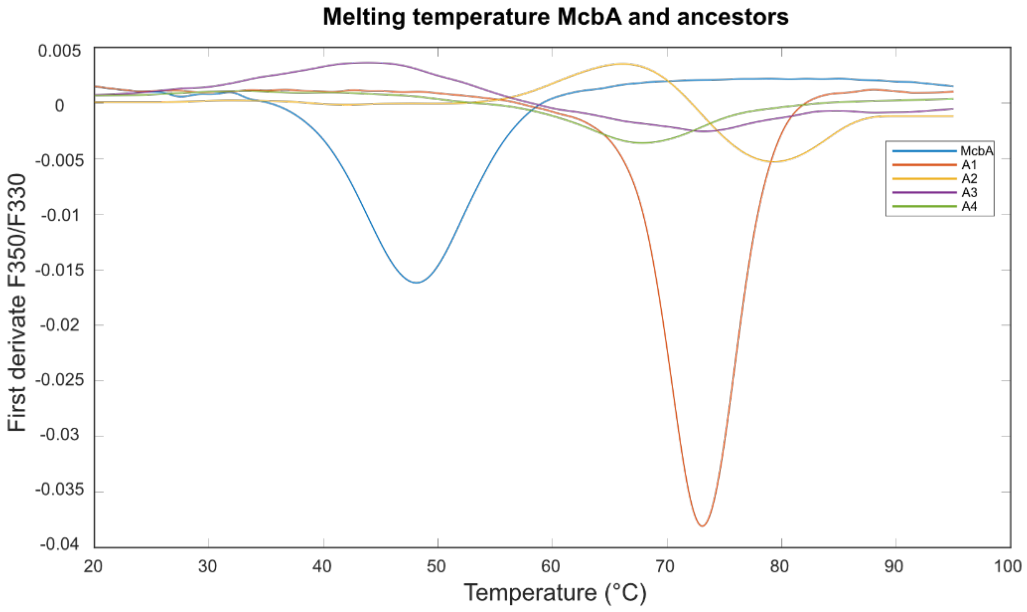


Figure S2**.** Thermal unfolding by McbA and ancestors A1-A4, measured by change in fluorescence ratio at wavelengths 350 and 330 nm by Prometheus NT.48 nanoDSF (NanoTemper Technologies).

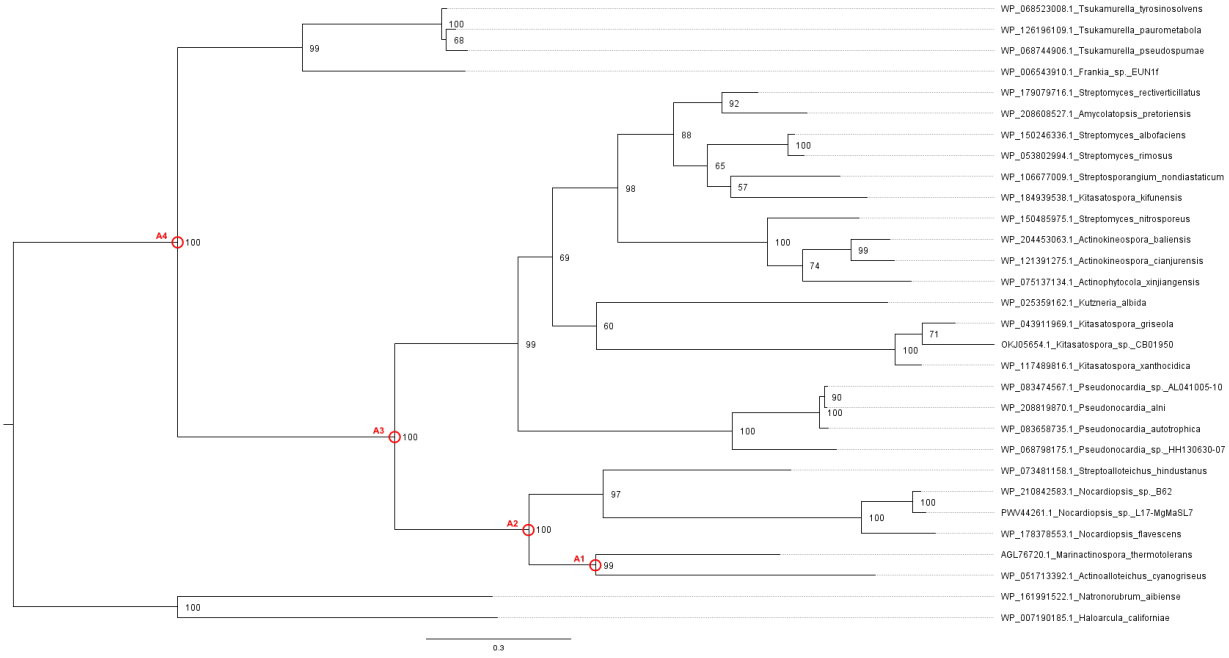
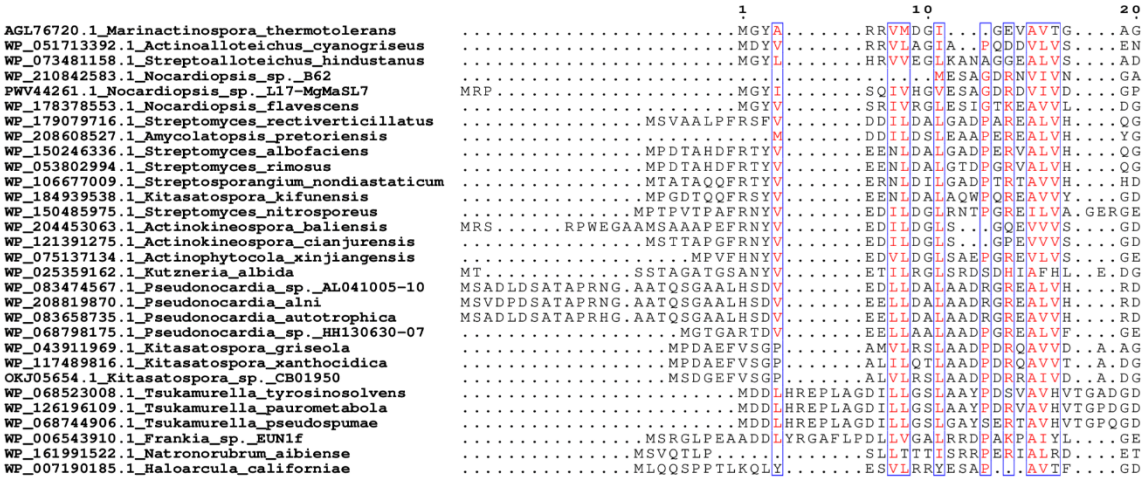
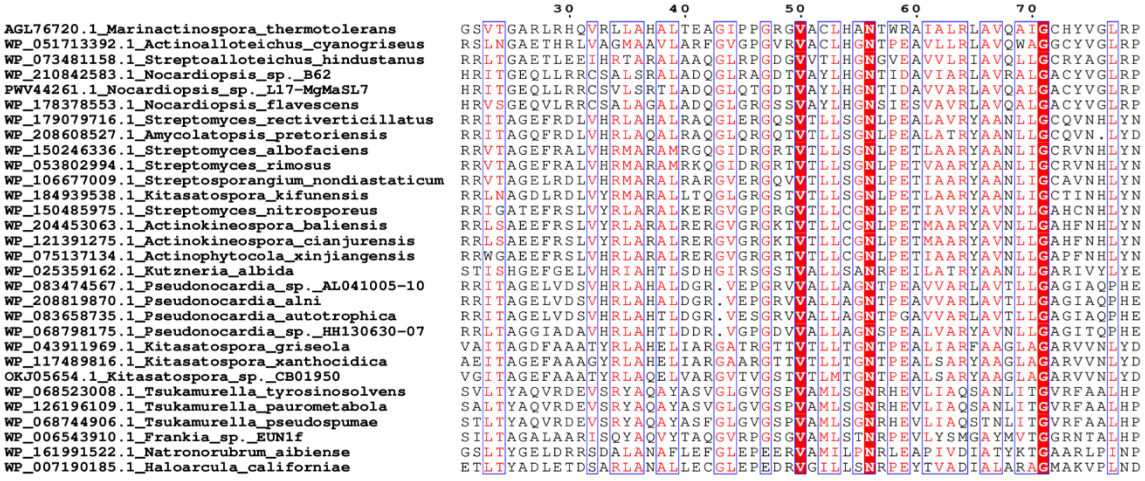
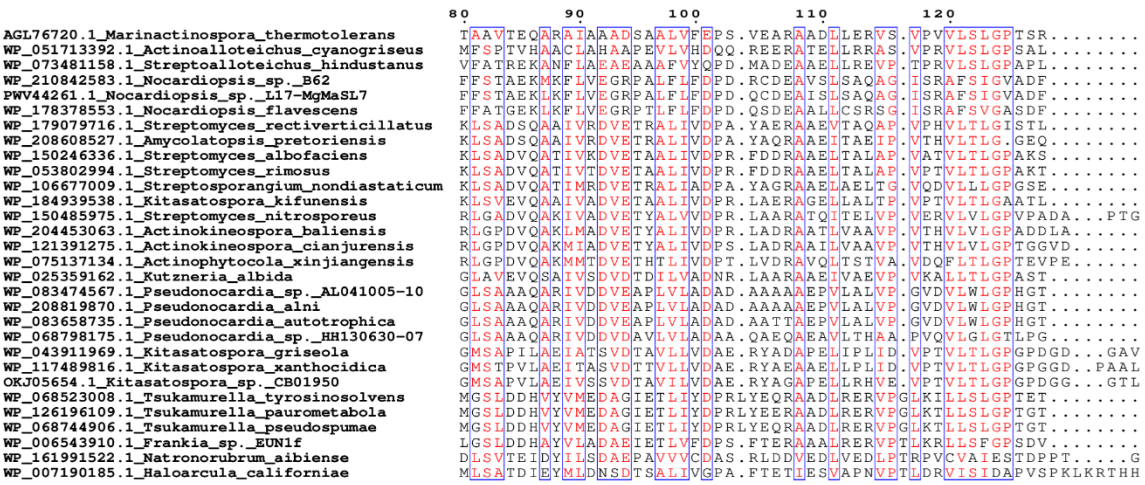
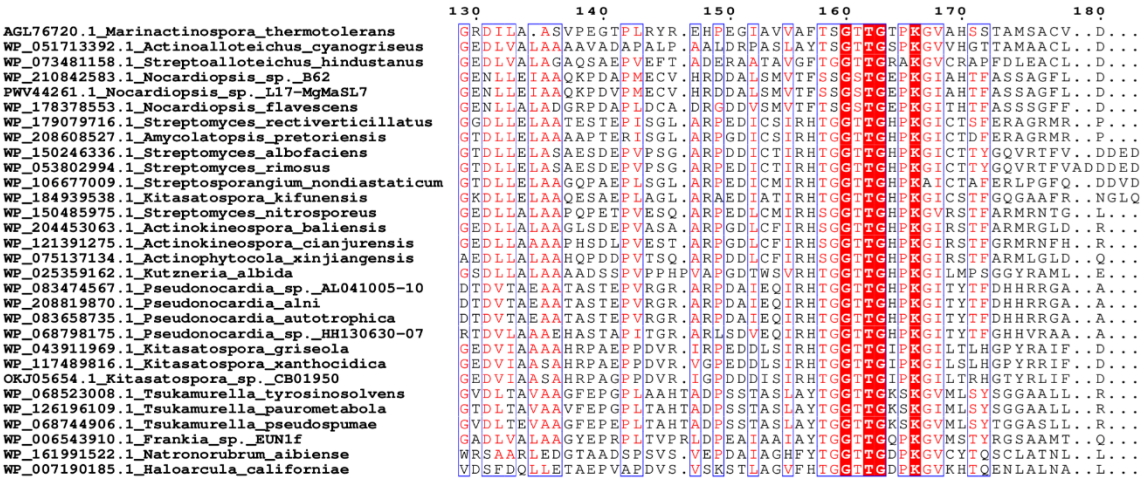


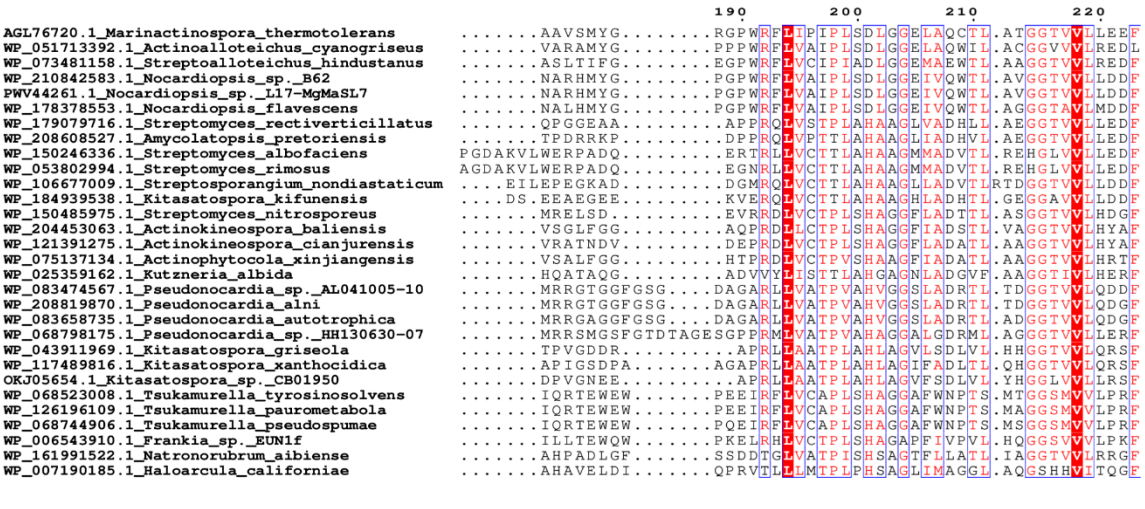
Figure S3**.** Phylogenetic tree of McbA used to compute ancestral sequences. The numbers shown on the nodes are the bootstrap values (1000 bootstrap).

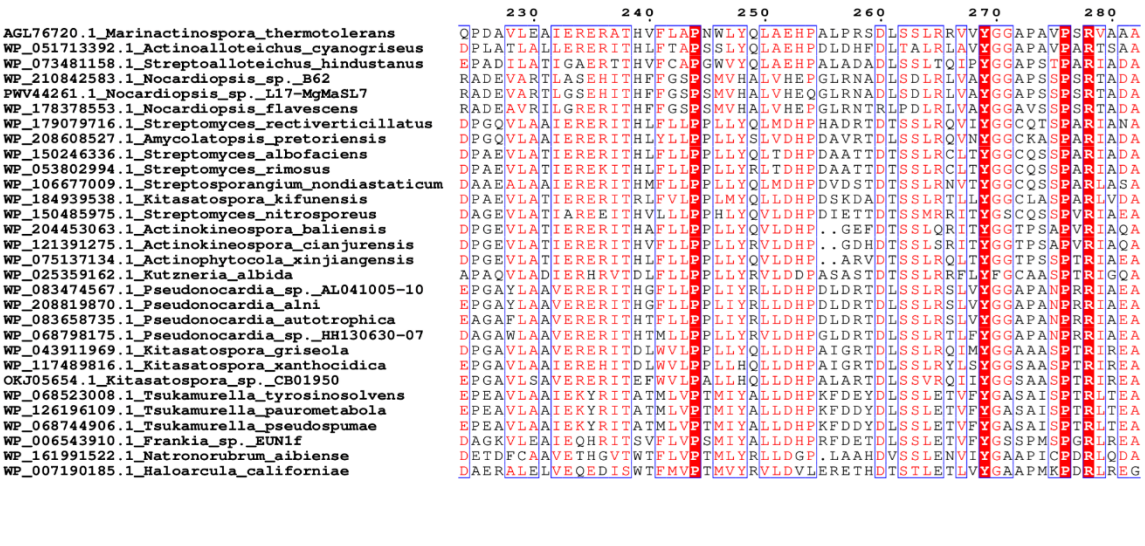

AGL76720.1_Marinactinospora_thermotolerans
WP_051713392.1_Actinoalloteichus_cyanogriseus
WP_073481158.1_Streptoalloteichus_hindustanus
WP_210842583.1_Nocardiopsis_sp._B62
PWV44261.1_Nocardiopsis_sp._L17-MgMaSL7
WP_178378553.1_Nocardiopsis_flavescens
WP_179079716.1_Streptomyces_rectiverticillatus
WP_208608527.1_Amycolatopsis_pretoriensis
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_Actinokineospora_baliensis
WP_121391275.1_Actinokineospora_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_tyrosinosolvens
WP_126196109.1_Tsukamurella_paurometabola
WP_068744906.1_Tsukamurella_pseudospumae
WP_006543910.1_Frankia_sp._EUN1f
WP_161991522.1_Natronorubrum_aibiense
WP_007190185.1_Haloarcula_californiae

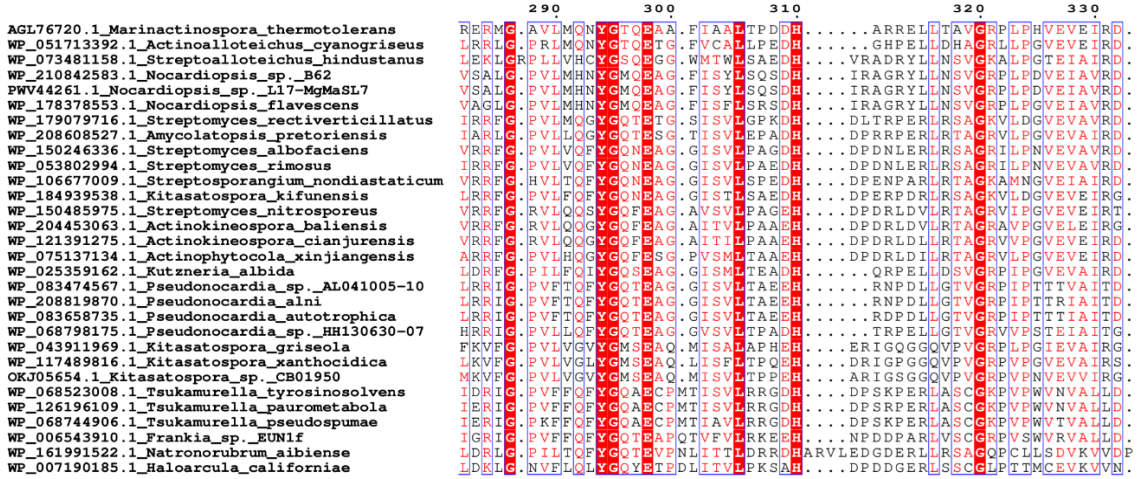

AGL76720.1_Marinactinospora_thermotolerans
WP_051713392.1_Actinoalloteichus_cyanogriseus
WP_073481158.1_Streptoalloteichus_hindustanus
WP_210842583.1_Nocardiopsis_sp._B62
PWV44261.1_Nocardiopsis_sp._L17-MgMaSL7
WP_178378553.1_Nocardiopsis_flavescens
WP_179079716.1_Streptomyces_rectiverticillatus
WP_208608527.1_Amycolatopsis_pretoriensis
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_Actinokineospora_baliensis
WP_121391275.1_Actinokineospora_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_tyrosinosolvens
WP_126196109.1_Tsukamurella_paurometabola
WP_068744906.1_Tsukamurella_pseudospumae
WP_006543910.1_Frankia_sp._EUN1f
WP_161991522.1_Natronorubrum_aibiense
WP_007190185.1_Haloarcula_californiae

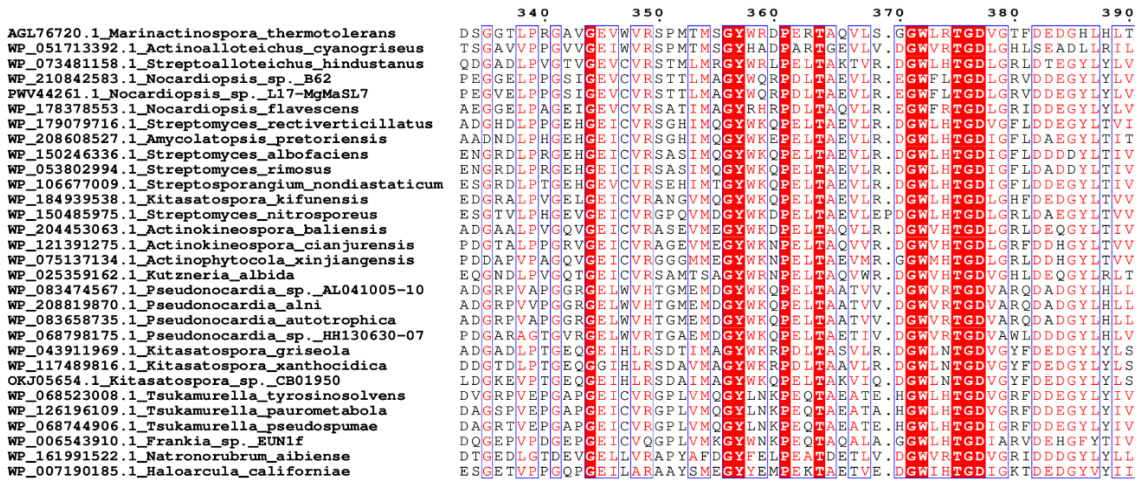

AGL76720.1_Marinactinospora_thermotolerans
WP_051713392.1_Actinoalloteichus_cyanogriseus
WP_073481158.1_Streptoalloteichus_hindustanus
WP_210842583.1_Nocardiopsis_sp._B62
PWV44261.1_Nocardiopsis_sp._L17-MgMaSL7
WP_178378553.1_Nocardiopsis_flavescens
WP_179079716.1_Streptomyces_rectiverticillatus
WP_208608527.1_Amycolatopsis_pretoriensis
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_Actinokineospora_baliensis
WP_121391275.1_Actinokineospora_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_tyrosinosolvens
WP_126196109.1_Tsukamurella_paurometabola
WP_068744906.1_Tsukamurella_pseudospumae
WP_006543910.1_Frankia_sp._EUN1f
WP_161991522.1_Natronorubrum_aibiense
WP_007190185.1_Haloarcula_californiae

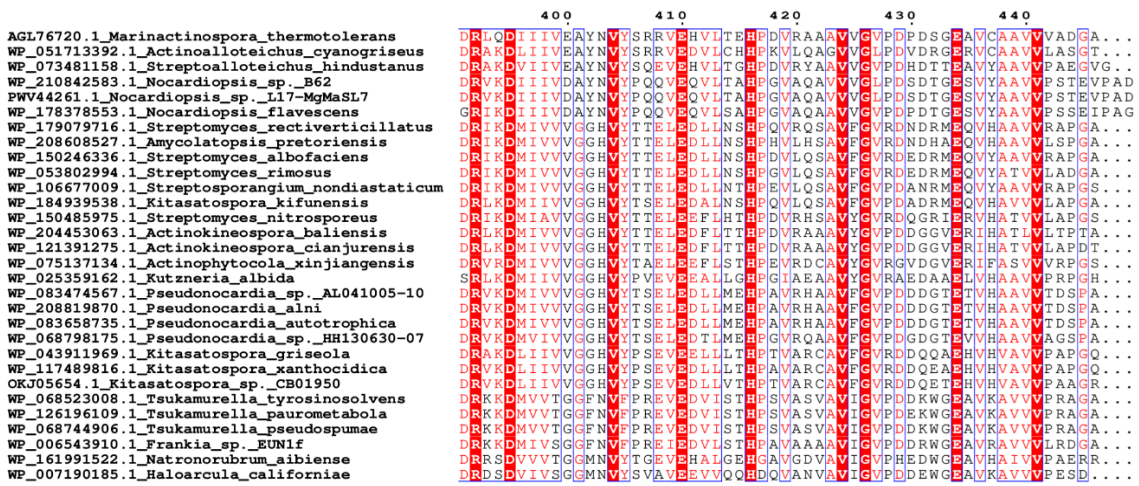

130
140
150
160
170
180
AGL76720.1_Marinactinospora_thermotolerans
WP_051713392.1_Actinoalloteichus_cyanogriseus
WP_073481158.1_Streptoalloteichus_hindustanus
WP_210842583.1_Nocardiopsis_sp._B62
PWV44261.1_Nocardiopsis_sp._L17-MgMaSL7
WP_178378553.1_Nocardiopsis_flavescens
WP_179079716.1_Streptomyces_rectiverticillatus
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_Actinokineospora_baliensis
WP_121391275.1_Actinokineospora_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_tyrosinosolvens
WP_126196109.1_Tsukamurella_paurometabola
WP_068744906.1_Tsukamurella_pseudospumae
WP_006543910.1_Frankia_sp._EUN1f
WP_161991522.1_Natronorubrum_aibiense
WP_007190185.1_Haloarcula_californiae


190
200
210
220
AGL76720.1_Marinactinospora_thermotolerans
WP_051713392.1_Actinoalloteichus_cyanogriseus
WP_073481158.1_Streptoalloteichus_hindustanus
WP_210842583.1_Nocardiopsis_sp._B62
PWV44261.1_Nocardiopsis_sp._L17-MgMaSL7
WP_178378553.1_Nocardiopsis_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_Actinokineospora_baliensis
WP_121391275.1_Actinokineospora_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_tyrosinosolvens
WP_126196109.1_Tsukamurella_paurometabola
WP_068744906.1_Tsukamurella_pseudospumae
WP_006543910.1_Frankia_sp._EUN1f
WP_161991522.1_Natronorubrum_aibiense
WP_007190185.1_Haloarcula_californiae


230
240
250
260
270
280
AGL76720.1_Marinactinospora_thermotolerans
WP_051713392.1_Actinoalloteichus_cyanogriseus
WP_073481158.1_Streptoalloteichus_hindustanus
WP_210842583.1_Nocardiopsis_sp._B62
PWV44261.1_Nocardiopsis_sp._L17-MgMaSL7
WP_178378553.1_Nocardiopsis_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_Actinokineospora_baliensis
WP_121391275.1_Actinokineospora_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_tyrosinosolvens
WP_126196109.1_Tsukamurella_paurometabola
WP_068744906.1_Tsukamurella_pseudospumae
WP_006543910.1_Frankia_sp._EUN1f
WP_161991522.1_Natronorubrum_aibiense
WP_007190185.1_Haloarcula_californiae


290
300
310
320
330
AGL76720.1_Marinactinospora_thermotolerans
WP_051713392.1_Actinoalloteichus_cyanogriseus
WP_073481158.1_Streptoalloteichus_hindustanus
WP_210842583.1_Nocardiopsis_sp._B62
PWV44261.1_Nocardiopsis_sp._L17-MgMaSL7
WP_178378553.1_Nocardiopsis_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_Actinokineospora_baliensis
WP_121391275.1_Actinokineospora_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_tyrosinosolvens
WP_126196109.1_Tsukamurella_paurometabola
WP_068744906.1_Tsukamurella_pseudospumae
WP_006543910.1_Frankia_sp._EUN1f
WP_161991522.1_Natronorubrum_aibiense
WP_007190185.1_Haloarcula_californiae


340
350
360
370
380
390
AGL76720.1_Marinactinospora_thermotolerans
WP_051713392.1_Actinoalloteichus_cyanogriseus
WP_073481158.1_Streptoalloteichus_hindustanus
WP_210842583.1_Nocardiopsis_sp._B62
PWV44261.1_Nocardiopsis_sp._L17-MgMaSL7
WP_178378553.1_Nocardiopsis_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_Actinokineospora_baliensis
WP_121391275.1_Actinokineospora_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_tyrosinosolvens
WP_126196109.1_Tsukamurella_paurometabola
WP_068744906.1_Tsukamurella_pseudospumae
WP_006543910.1_Frankia_sp._EUN1f
WP_161991522.1_Natronorubrum_aibiense
WP_007190185.1_Haloarcula_californiae


400
410
420
430
440
AGL76720.1_Marinactinospora_thermotolerans
WP_051713392.1_Actinoalloteichus_cyanogriseus
WP_073481158.1_Streptoalloteichus_hindustanus
WP_210842583.1_Nocardiopsis_sp._B62
PWV44261.1_Nocardiopsis_sp._L17-MgMaSL7
WP_178378553.1_Nocardiopsis_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_Actinokineospora_baliensis
WP_121391275.1_Actinokineospora_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_tyrosinosolvens
WP_126196109.1_Tsukamurella_paurometabola
WP_068744906.1_Tsukamurella_pseudospumae
WP_006543910.1_Frankia_sp._EUN1f
WP_161991522.1_Natronorubrum_aibiense
WP_007190185.1_Haloarcula_californiae

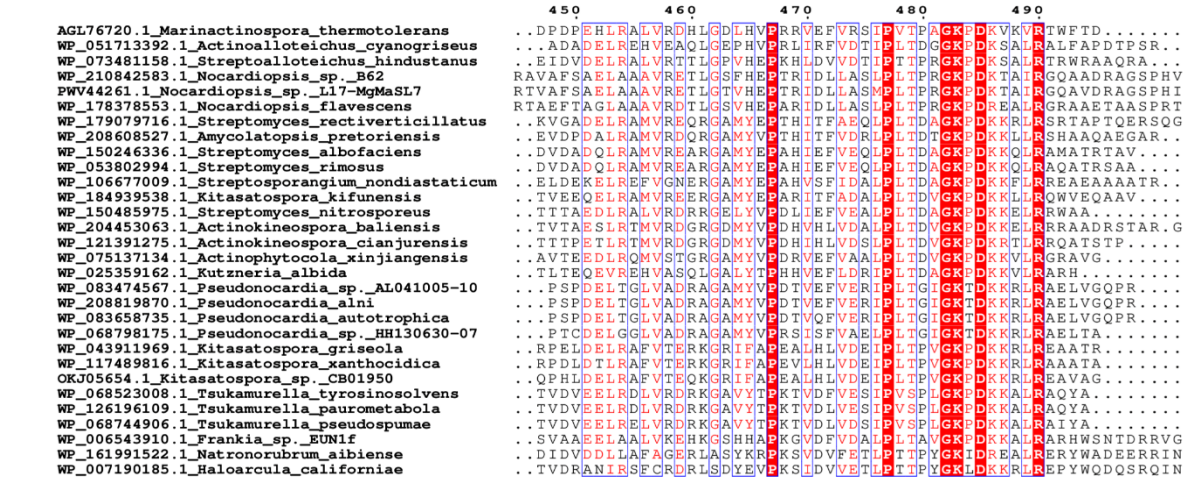


Figure S4. Alignment used for the phylogenetic tree construction. Alignment illustration made by using ESPript v 3.0, Robert, X. and Gouet, P. (2014) "Deciphering key features in protein structures with the new ENDscript server". Nucl. Acids Res. 42(W1), W320-W324 - doi: 10.1093/nar/gku316 (freely accessible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

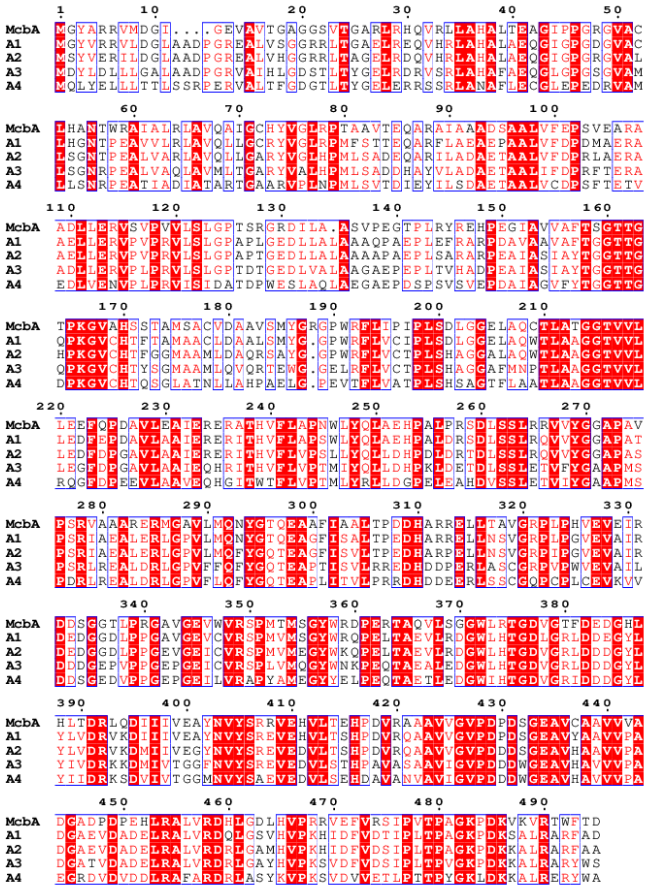


Figure S5. Alignment of McbA and ancestors A1-A4, derived from the phylogenetic tree in fig S1. Alignment illustration made by ESPript v 3.0, Robert, X. and Gouet, P. (2014) "Deciphering key features in protein structures with the new ENDscript server". Nucl. Acids Res. 42(W1), W320-W324 - doi: 10.1093/nar/gku316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

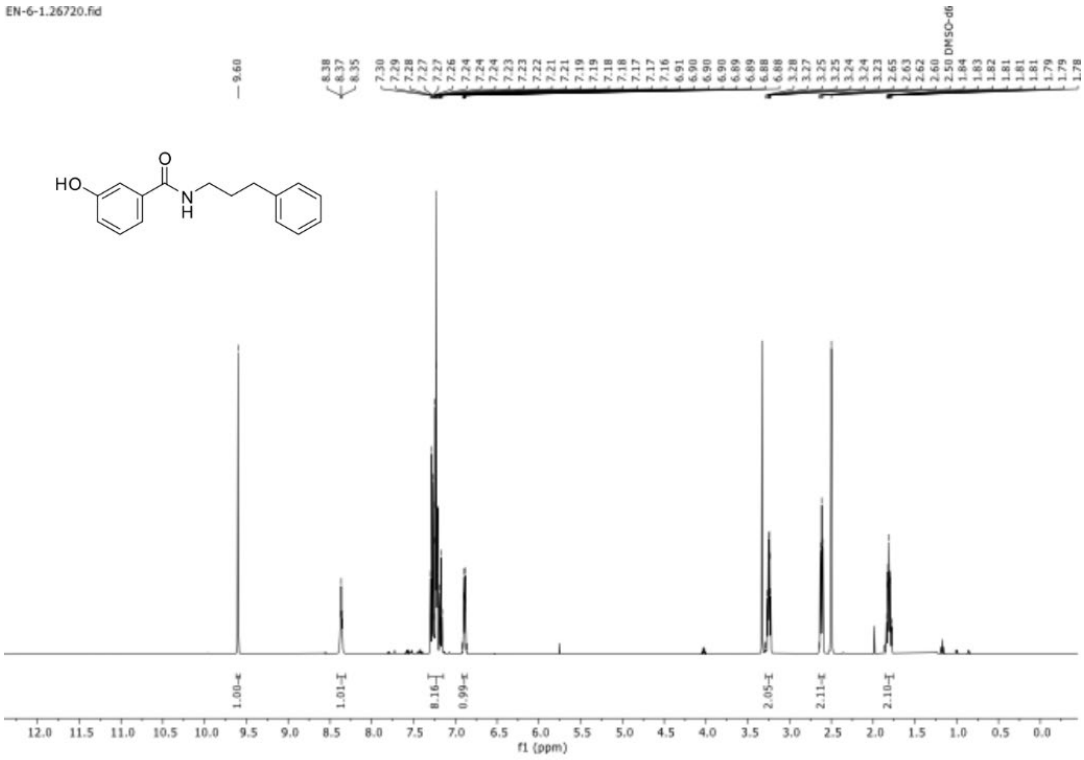


Figure S6. ^{1}H-NMR of synthesised 3-Hydroxy-N-(3-phenylpropyl)benzamide (standard).

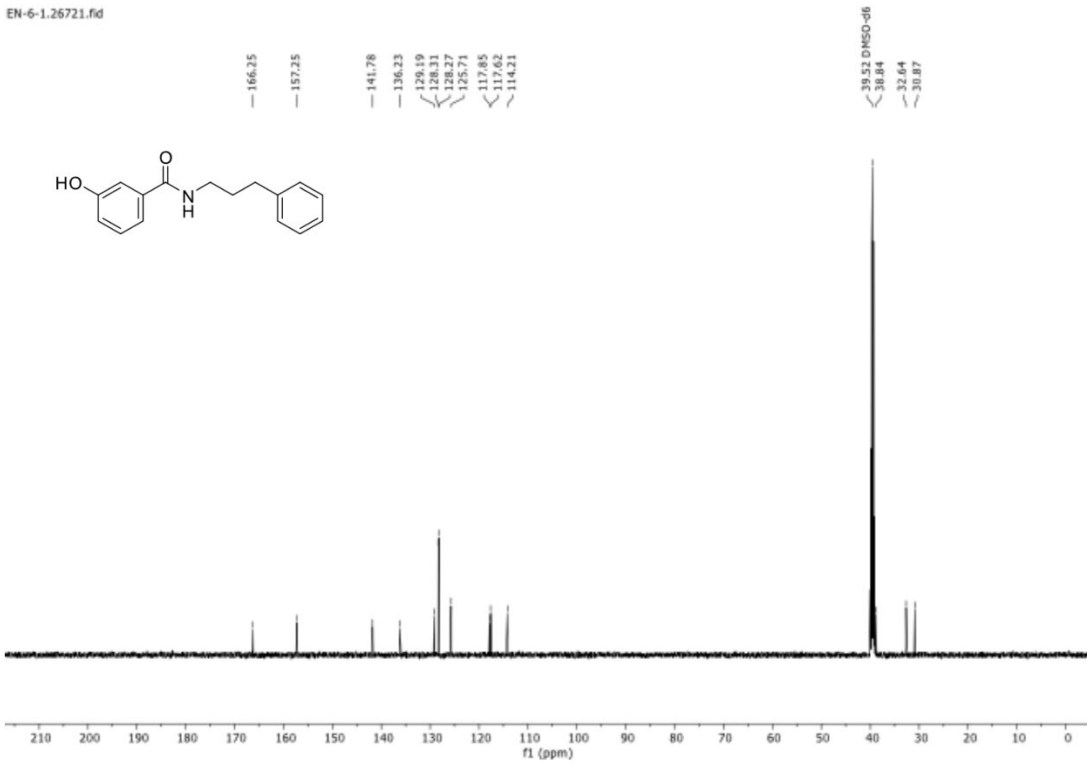


Figure S7. ^{13}C-NMR of synthesised 3-Hydroxy-N-(3-phenylpropyl)benzamide (standard).

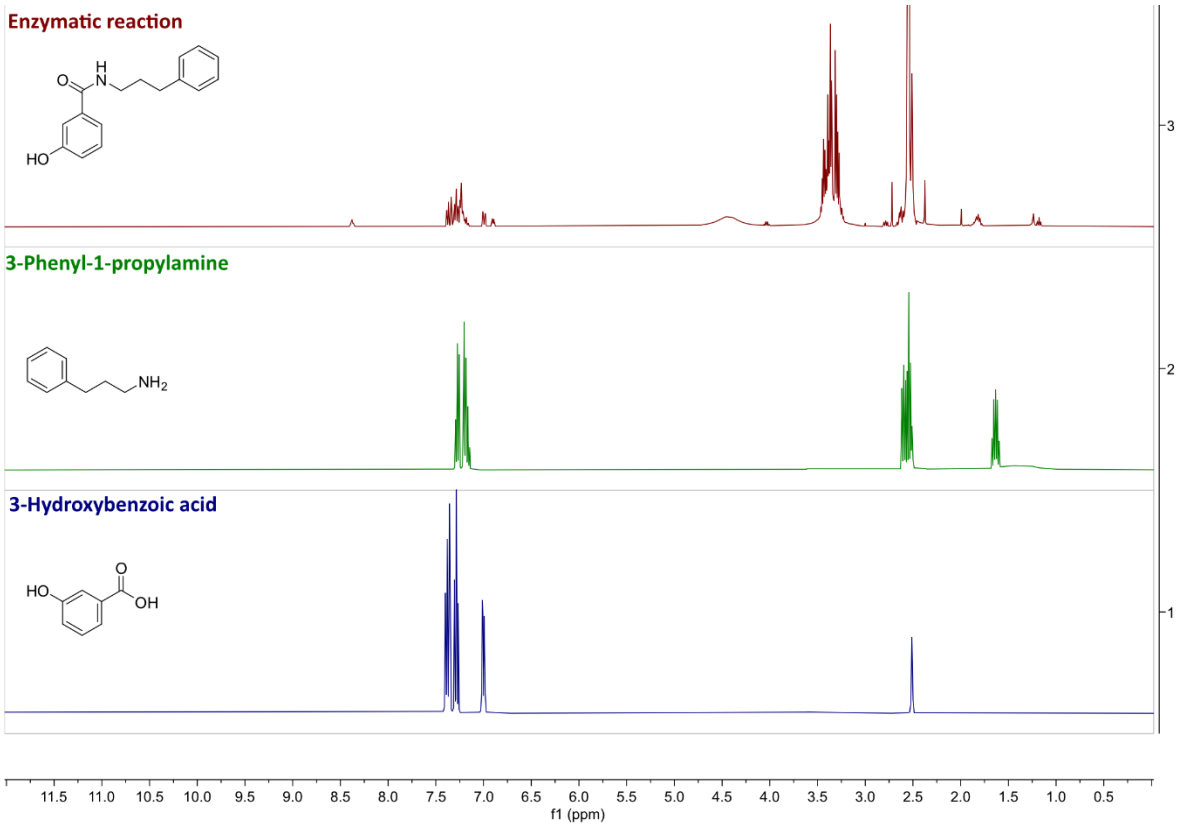


Figure S8. ^{1}H-NMR of the enzymatic synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide (crude NMR). As a control, ^{1}H-NMR of 3-phenyl-1-propylamine and 3-hydroxybenzoic acid are included.

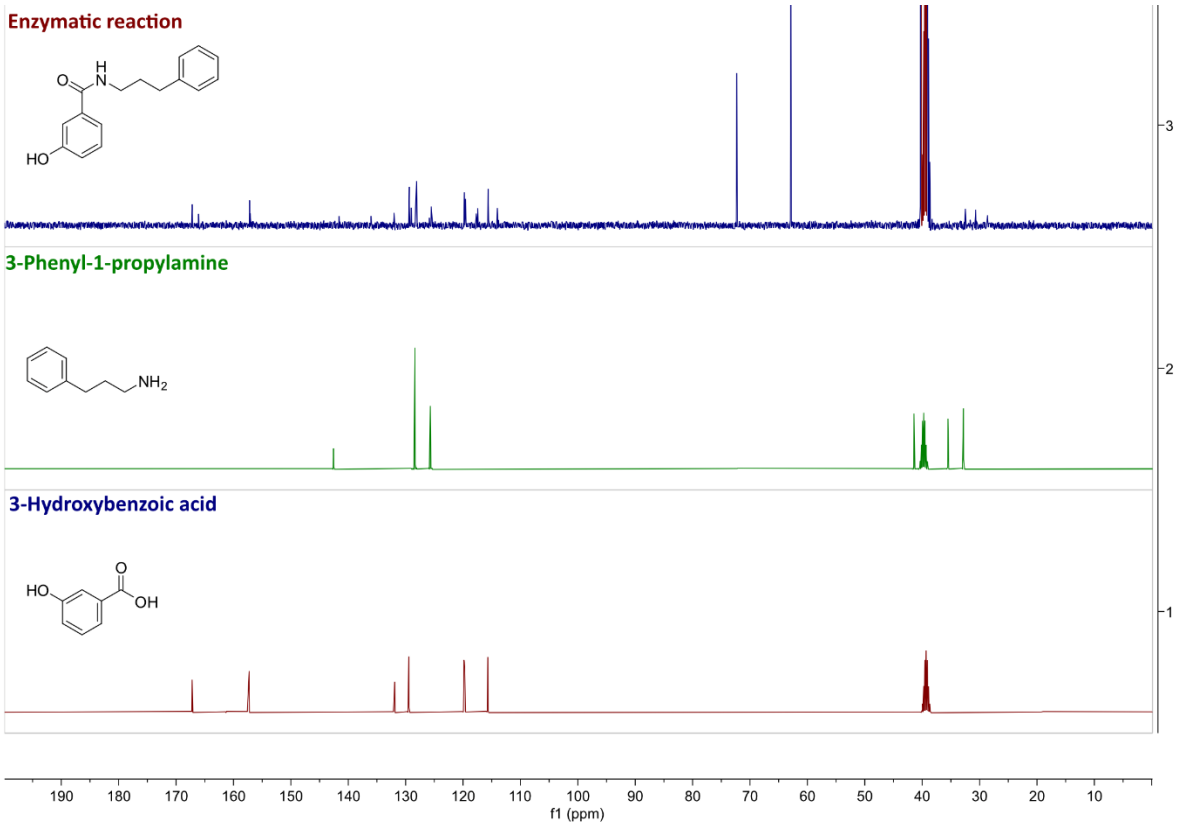


Figure S9. ^{13}C-NMR of the enzymatic synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide (crude NMR). As a control, ^{13}C-NMR of 3-phenyl-1-propylamine and 3-hydroxybenzoic acid are included.

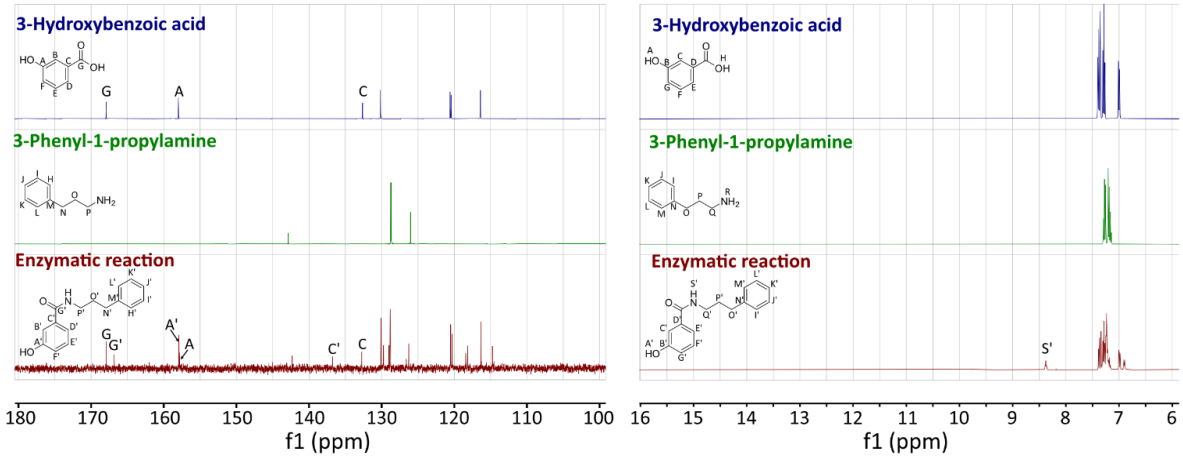


Figure S10 ^{13}C-NMR (left) and ^{1}H-NMR (right) of 3-hydroxybenzoic acid, 3-phenyl-1-propylamine, and the enzymatic synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide (crude NMR). In the left chromatogram, a shift of the carbonyl carbon (G'), the α-carbon (C') and the aromatic carbon (A') is consistent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

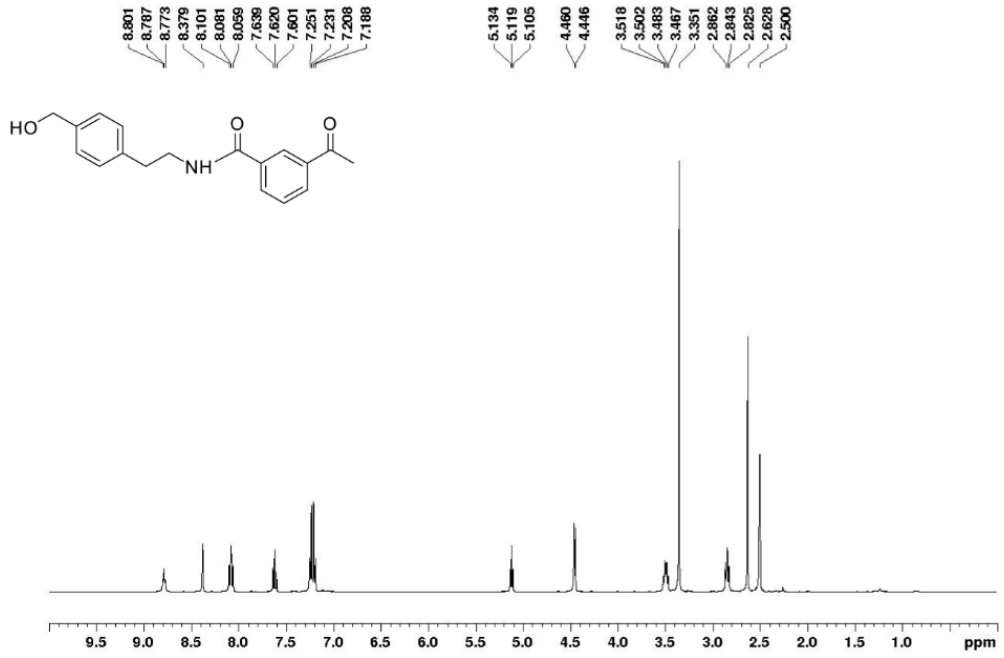


Figure S11. ^{1}H-NMR of synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide (standard).

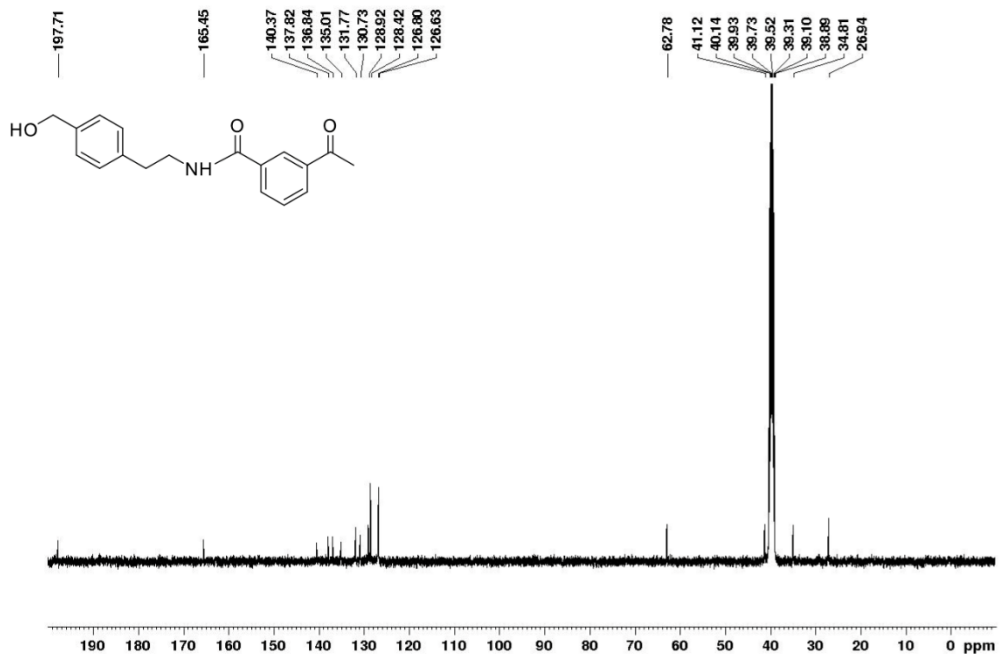


Figure S12. ^{13}C-NMR of synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide (standard).

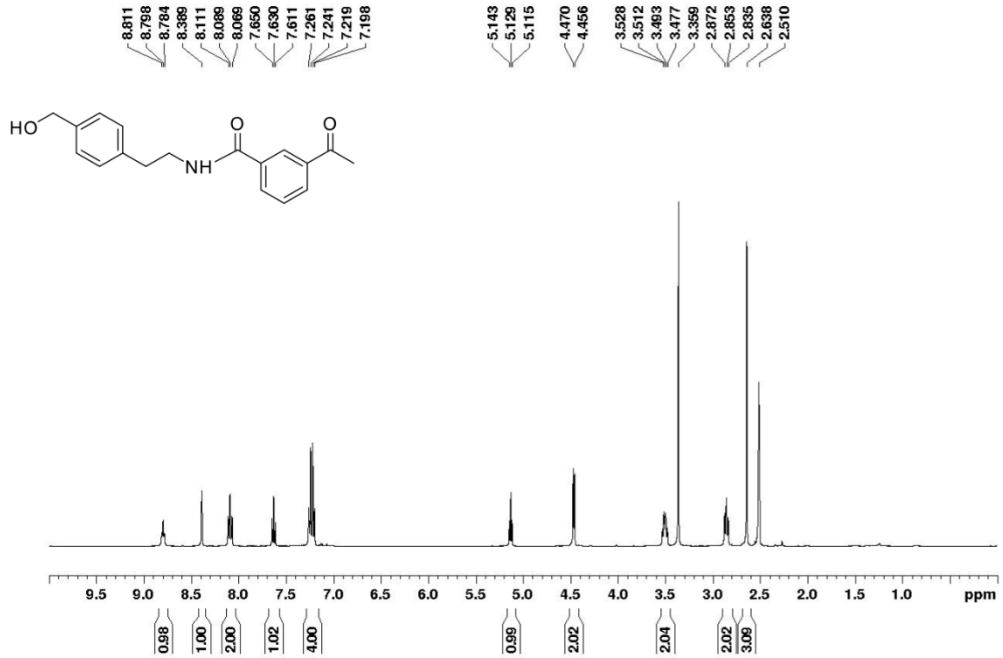


Figure S13. ^{1}H-NMR of biocatalytic synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide

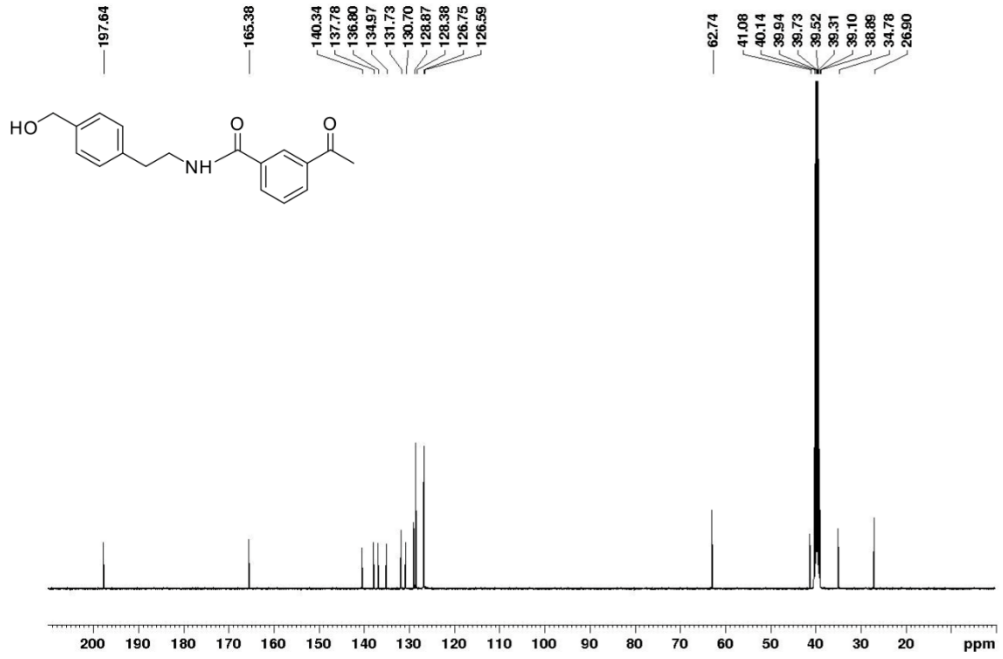


Figure S14. ^{13}C-NMR of biocatalytic synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide

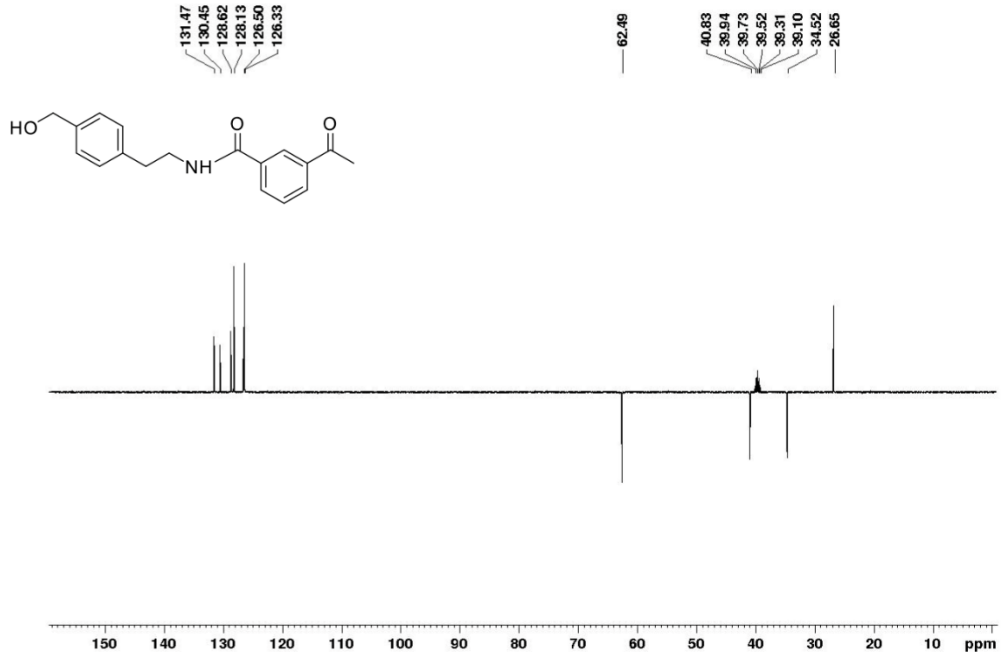


Figure S15. DEPT spectra of biocatalytic synthesised 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide

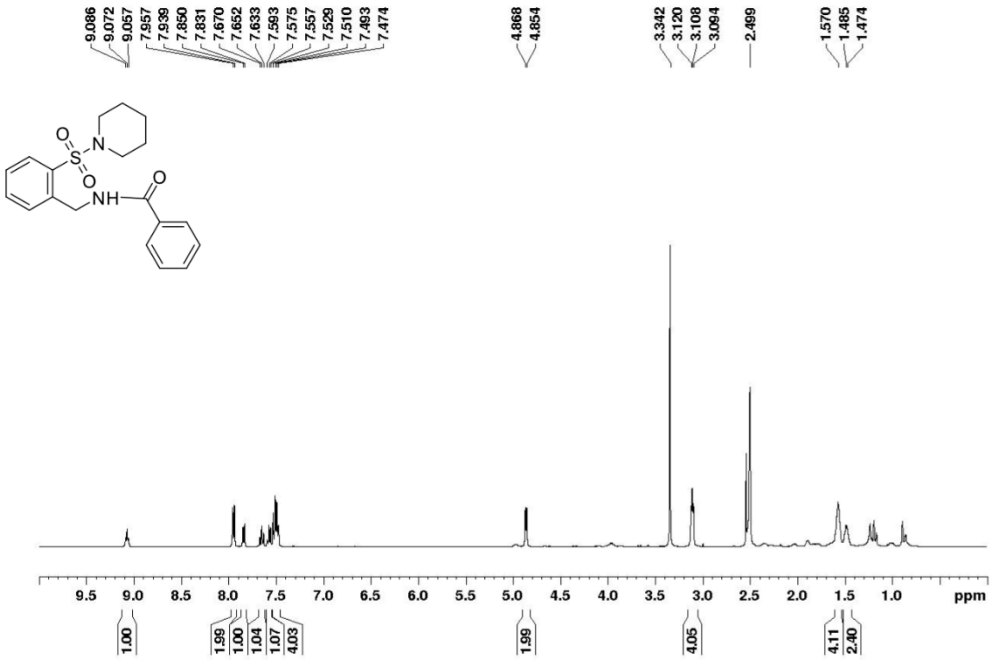


Figure S16. ^{1}H-NMR of synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide (standard).

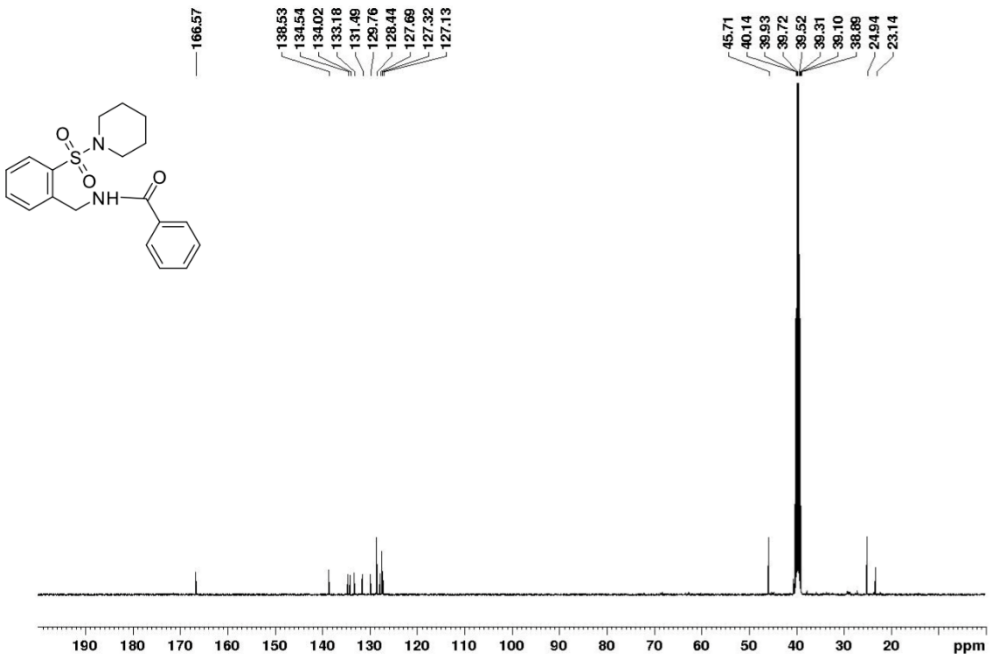


Figure S17. ^{13}C-NMR of synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide (standard).

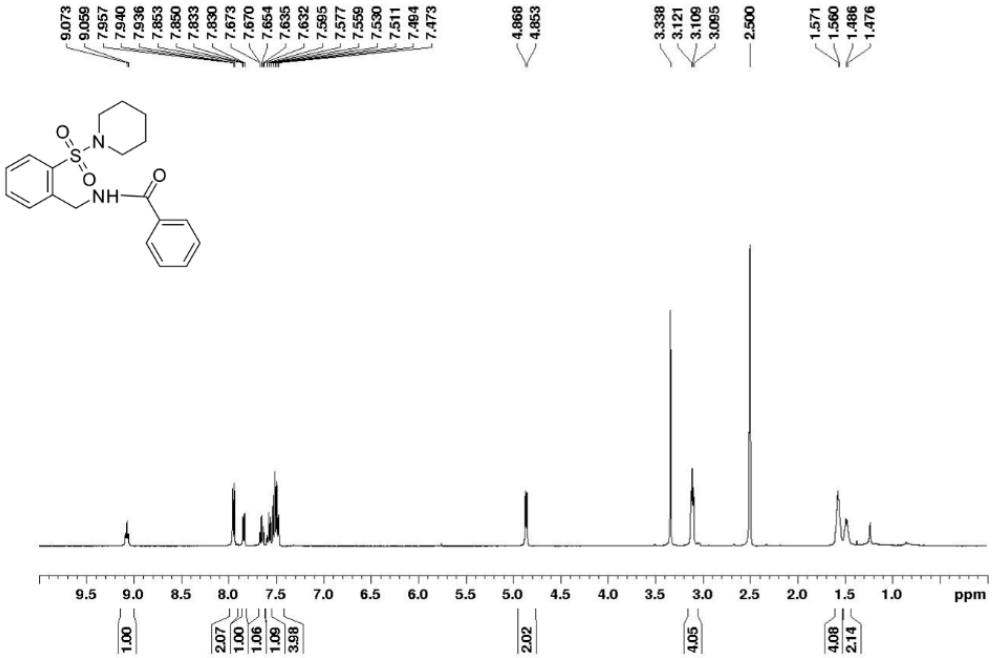


Figure S18 ^{1}H-NMR of biocatalytic synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide

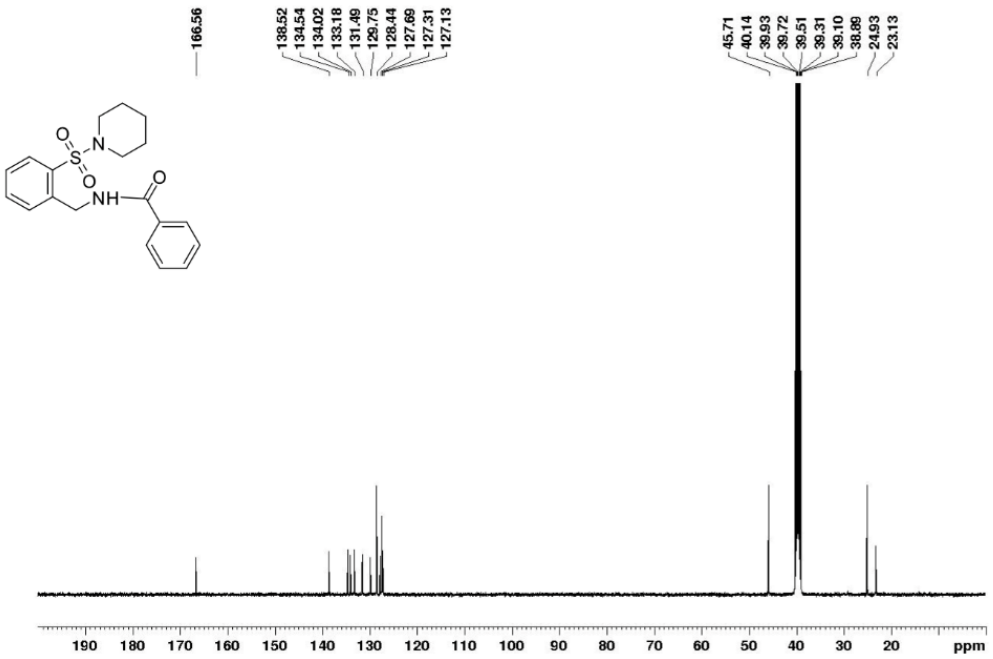


Figure S19 ^{13}C-NMR of biocatalytic synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide

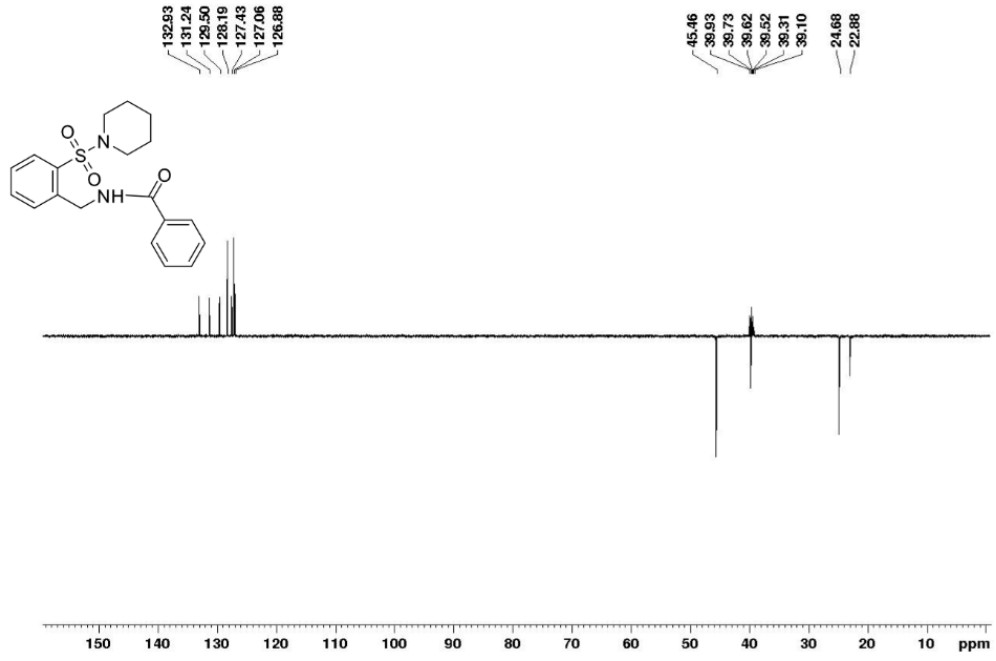


Figure S20. DEPT spectra of biocatalytic synthesised N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide

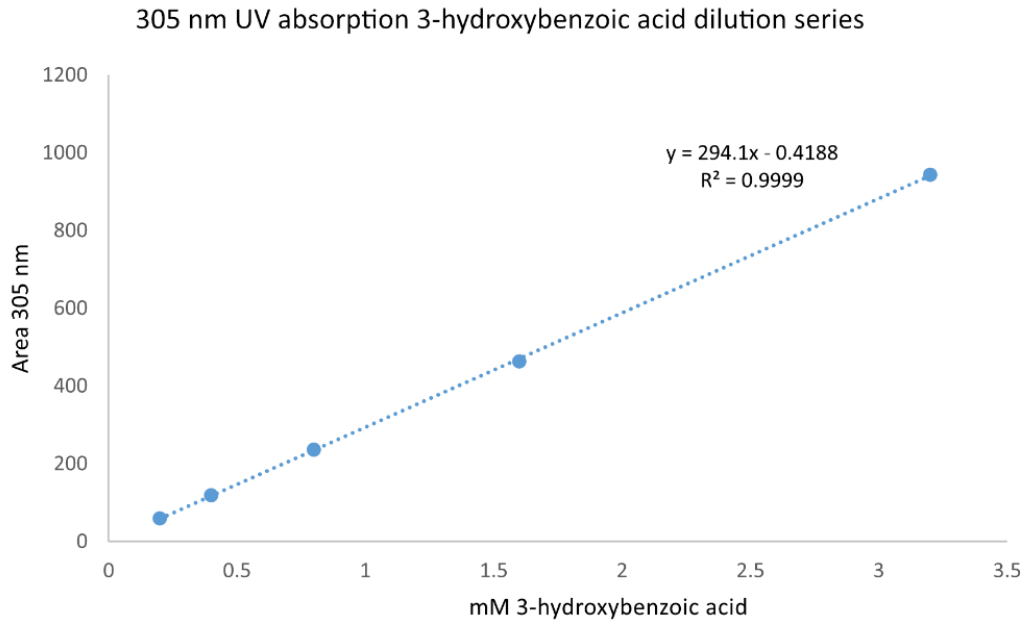


Figure S21. 305 nm UV absorption calibration curve of different 3-hydroxybenzoic acid concentrations. In the equation, y is the 305 nm peak area of the acid, and x is the mM of the acid in the HPLC sample.

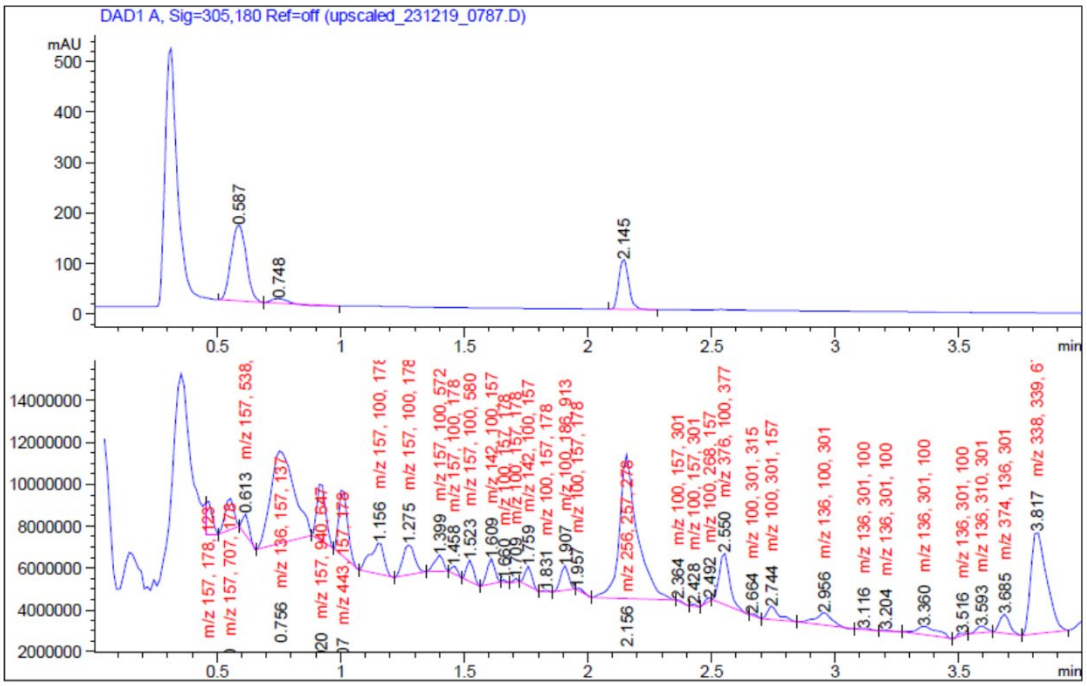


Figure S22. HPLC-MS (positive ionization mode, 305 nm) chromatogram of a sample taken from the upscaled biocatalytic synthesis of 3-Hydroxy-N-(3-phenylpropyl)benzamide 60 h after the start. At 0.587 min is 3-hydroxybenzoic acid, 0.748 min is 3-phenylpropylamine, and at 2.145 min is the 3-Hydroxy-N-(3-phenylpropyl)benzamide,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

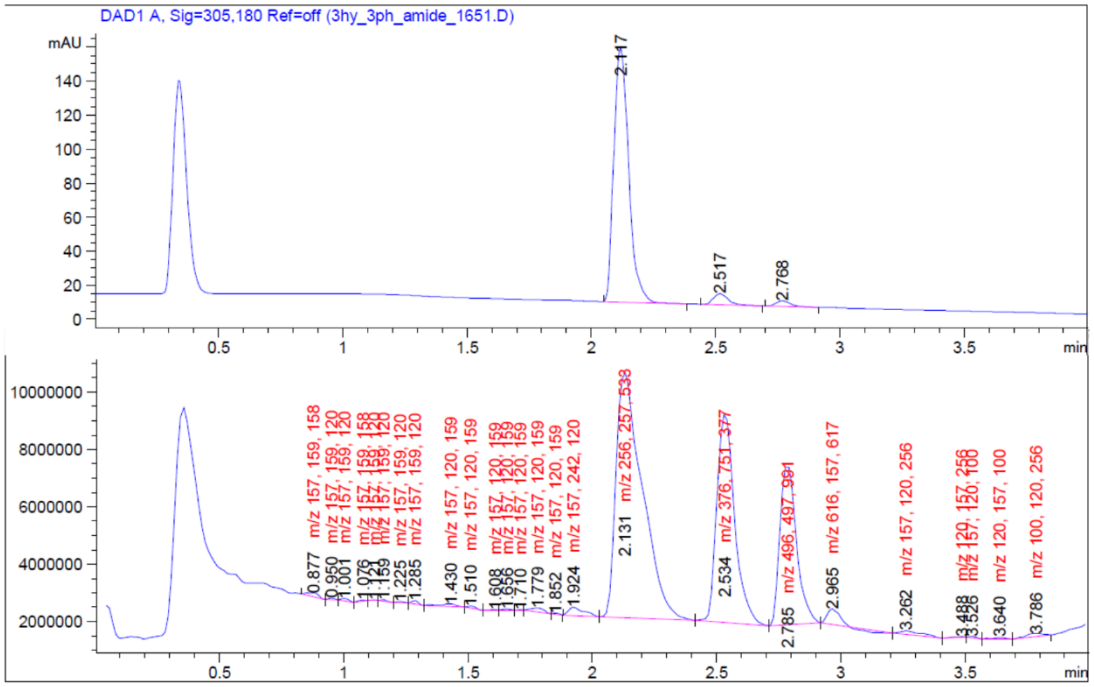


Figure S23. HPLC-MS (positive ionization mode, 305 nm) chromatogram of 3-Hydroxy-N-(3-phenylpropyl)benzamide standard. At 2.117 min is the amide, with its expected mass of 256 *m*/*z*. For reference, all mass peaks throughout the elution are shown.

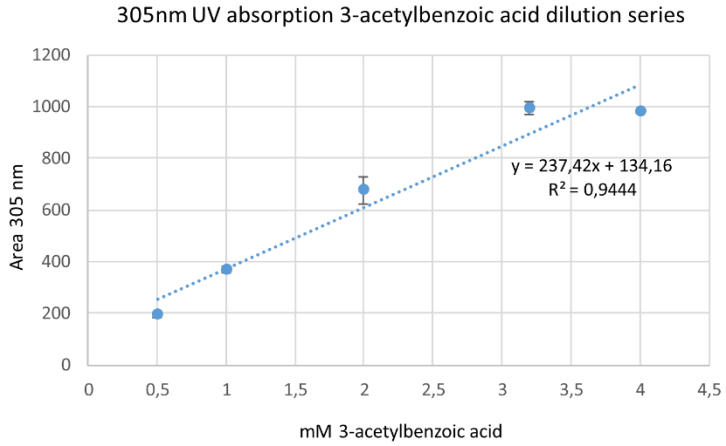


Figure S24. 305 nm UV absorption calibration curve of different 3-acetylbenzoic acid concentrations. In the equation, y is the 305 nm peak area of the acid, and x is the mM of the acid in the HPLC sample.

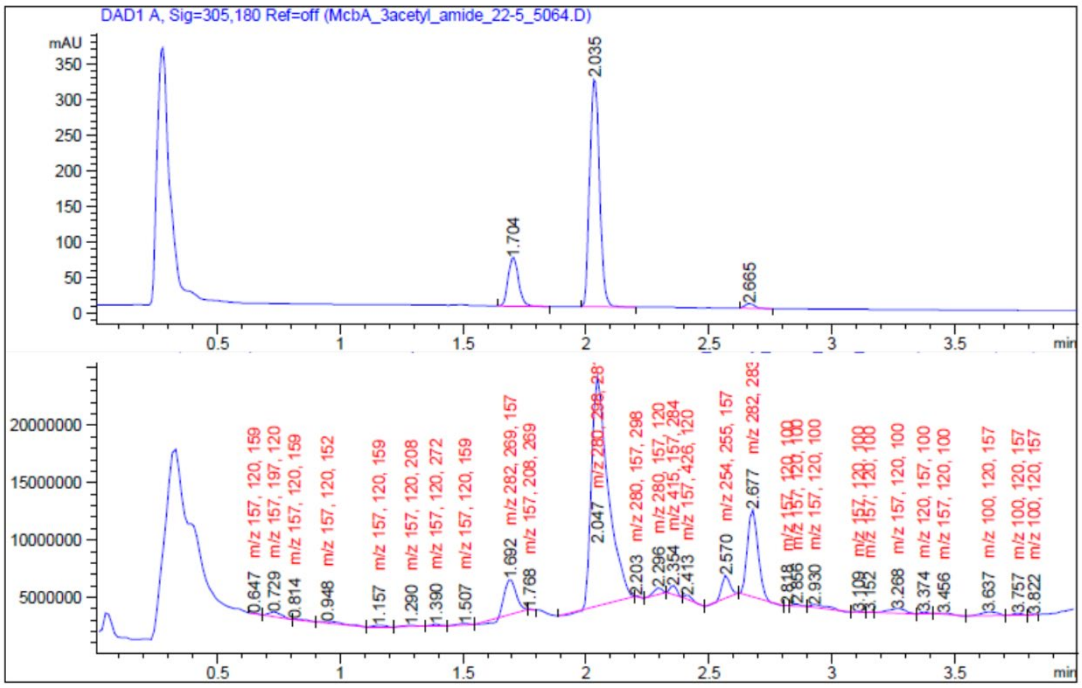


Figure S25. HPLC-MS (positive ionization mode, 305 nm) chromatogram of a sample taken from the upscaled biocatalytic synthesis of 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide 40 h after the start. At 1.704 min is 3-acetylbenzoic acid, and at 2.035 min is 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide,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

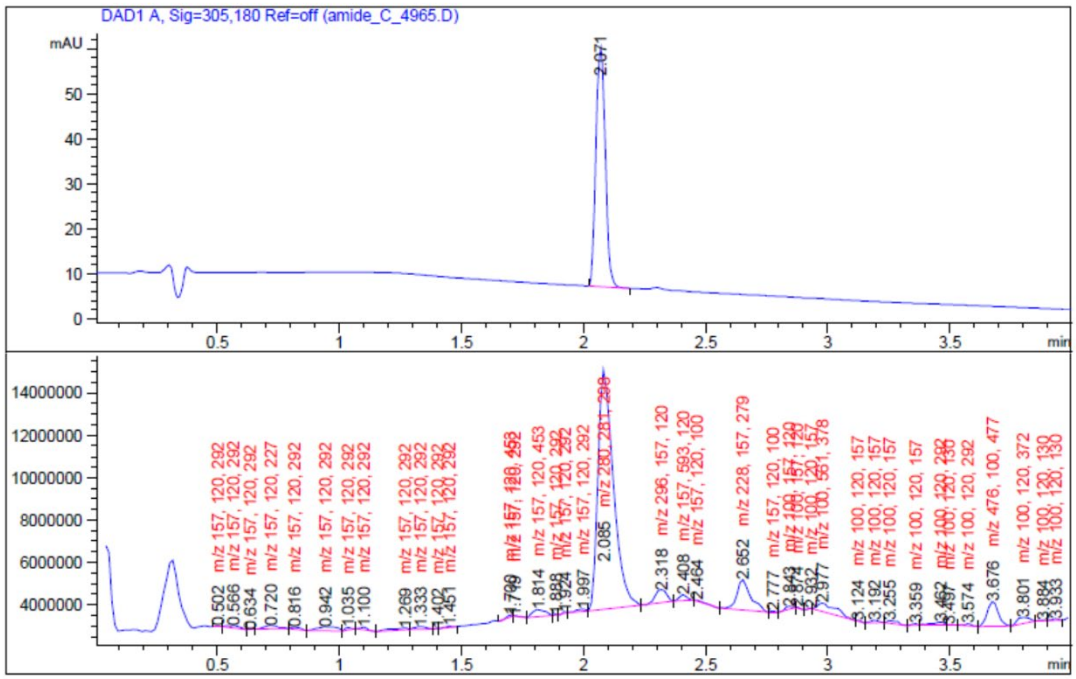


Figure S26. HPLC-MS (positive ionization mode, 305 nm) chromatogram of 3-acetyl-N-(4-(hydroxymethyl)phenethyl)benzamide standard. At 2.071 min is the amide, with its expected mass of 298 *m/z*. For reference, all mass peaks throughout the elution are shown.

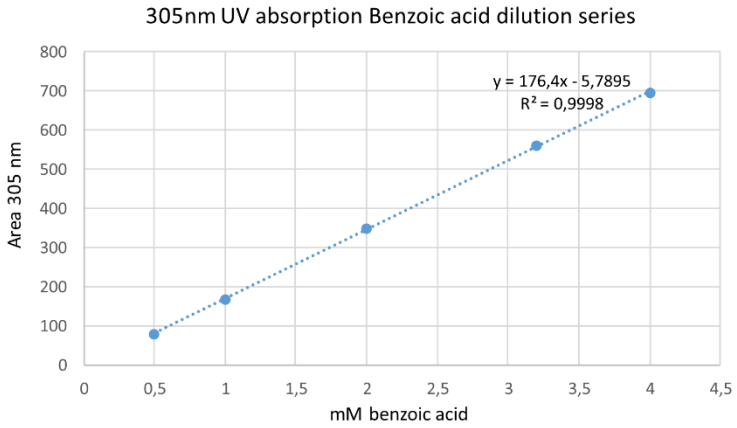


Figure S27. 305 nm UV absorption calibration curve of different benzoic acid concentrations. In the equation, y is the 305 nm peak area of the acid, and x is the mM of the acid in the HPLC sample.

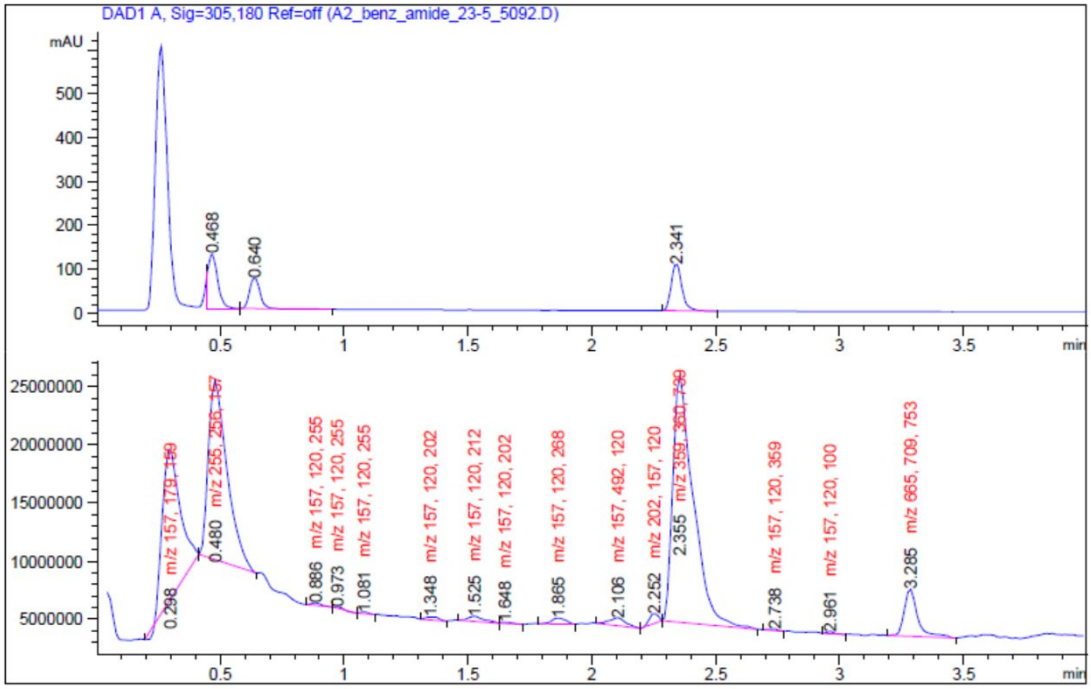


Figure S28. HPLC-MS (positive ionization mode, 305 nm) chromatogram of a taken from the upscaled biocatalytic synthesis of N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide 60 h after the start. At 0.640 min is benzoic acid, and at 2.341 min is N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide, with its expected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

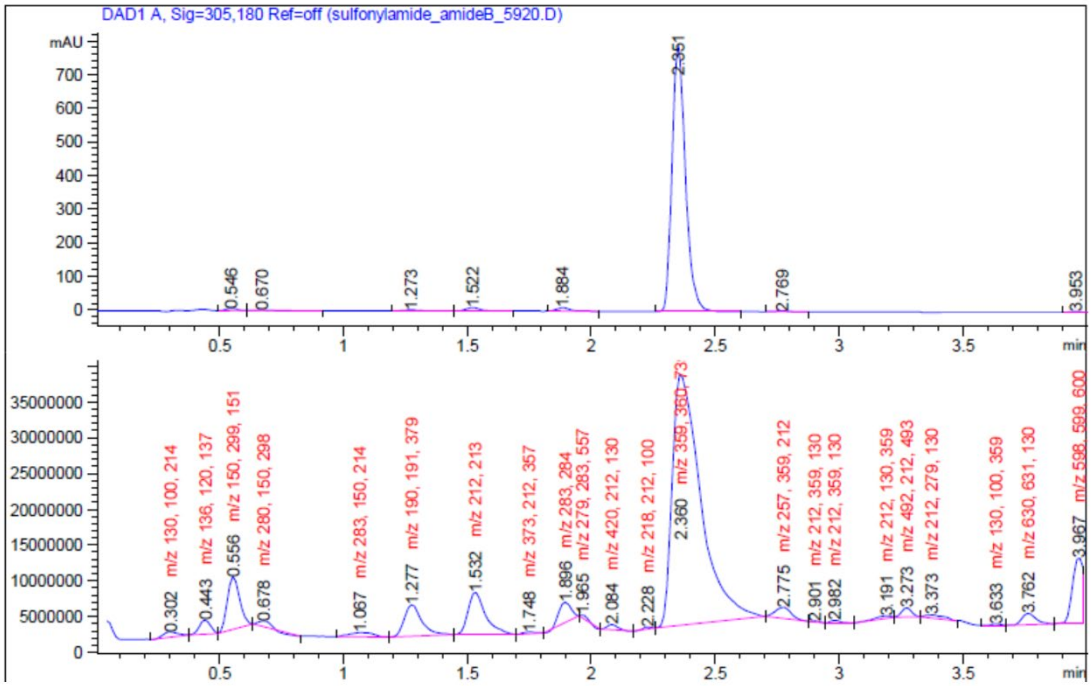


Figure S29. HPLC-MS (positive ionization mode, 305 nm) chromatogram of N-(2-(piperidin-1-ylsulfonyl)benzyl)benzamide standard. At 2.360 min is the amide, with its expected mass of 359 *m*/*z*. For reference, all mass peaks throughout the elution are shown.

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^{-1}
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^3/mol.s)]	$\frac{1.5e^6}{2} \cdot 10^{LogOH}$	1.14	with 1.5e^6 mol OH/cm^3
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 [a)], with uncertainty increased by a factor 1.5 for intermedia extrapolation
kdegSl [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 [a)], 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 = [1: fish, 2: daphnid, 3: algae]	1.4[b)]	with EC50 = ChV/0.4[c)] 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[d)], NOAEL-to-EC50 extrapolation factor = 9 [d)], human bodyweight = 70 kg, human lifetime = 70 years, 1 year = 365 days, 1 mg = 1e^{-6} 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 [d)], human bodyweight = 70 kg, human lifetime = 70 years, 1 year = 365 days, 1 mg = 1e^{-6} kg
ED50 cancer, inhalation [kg/lifetime]	50% chronic cancer effect dose via inhalation	CTV IUR [risk per mg/m^3]	$0.8 \cdot \frac{1}{IUR} \cdot 13 \cdot 70 \cdot 365 \cdot 1e^{-6}$	1.23	with 1/q*-to-ED50 conversion factor = 0.8 [d)], inhalation rate = 13 m^3/d, human lifetime = 70 years, 1 year = 365 days, 1 mg = 1e^{-6} 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⩾0.6): 1 Inside extended domain (0.4 ⩽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⩽1): 1 Inside extended domain (1 < AD ⩽ 3): 1.5 Outside domain (AD ⩾ 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

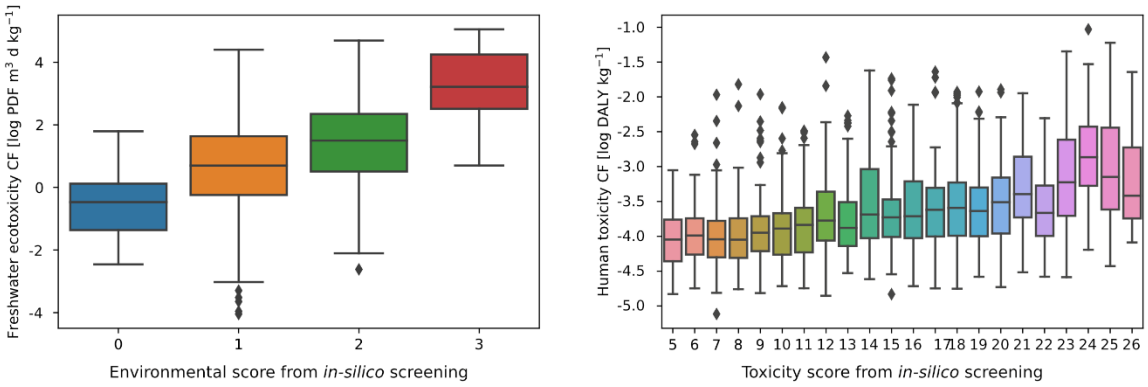


Figure S30.USEtox characterization factors vs *in-silico* screening scores of safechems and subset of filtered out aromatic acids and amines A) Median freshwater ecotoxicity USEtox characterization factor of the in-silico screening environmental score of safechems and sample of 408 aromatic amines and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

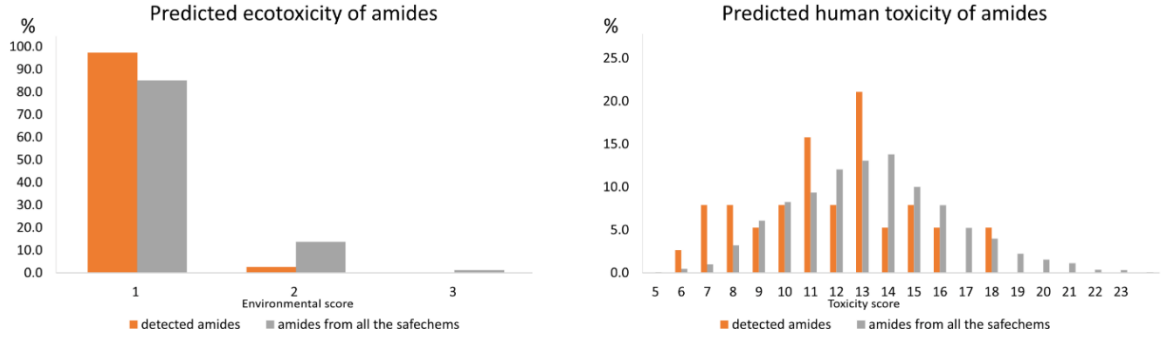


Figure S31. Predicted toxicity and environmental score of detected amides from the experimental coupling considering only a subset of possible safechems in comparison to all possible amides from the whole safechem panel.

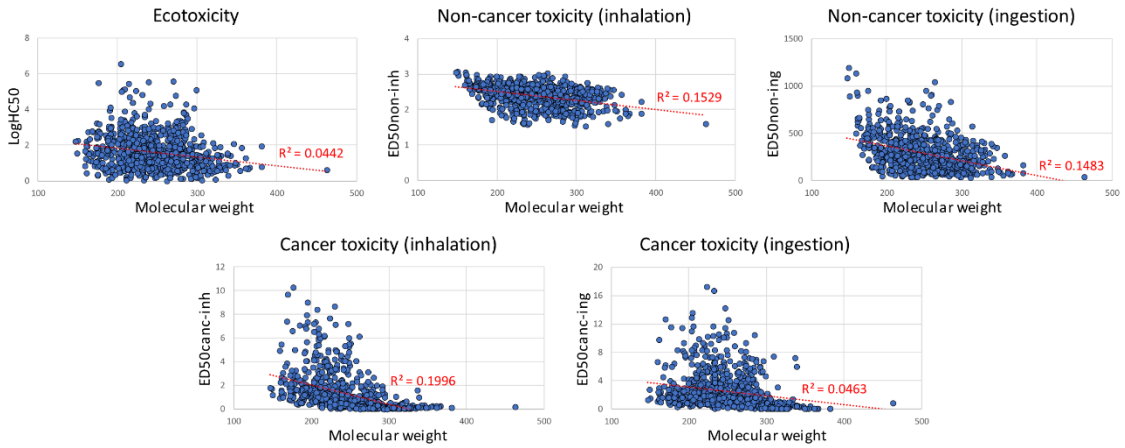


Figure S32. Toxic effects coupled to the molecular weight of amines and acids which were filtered out in the *in silico* filtering process (figure 3). The lower the value (y-axes), the more toxic. HC50 is the hazardous concentration of a chemical at which 50% of the species in an aquatic ecosystem are 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).

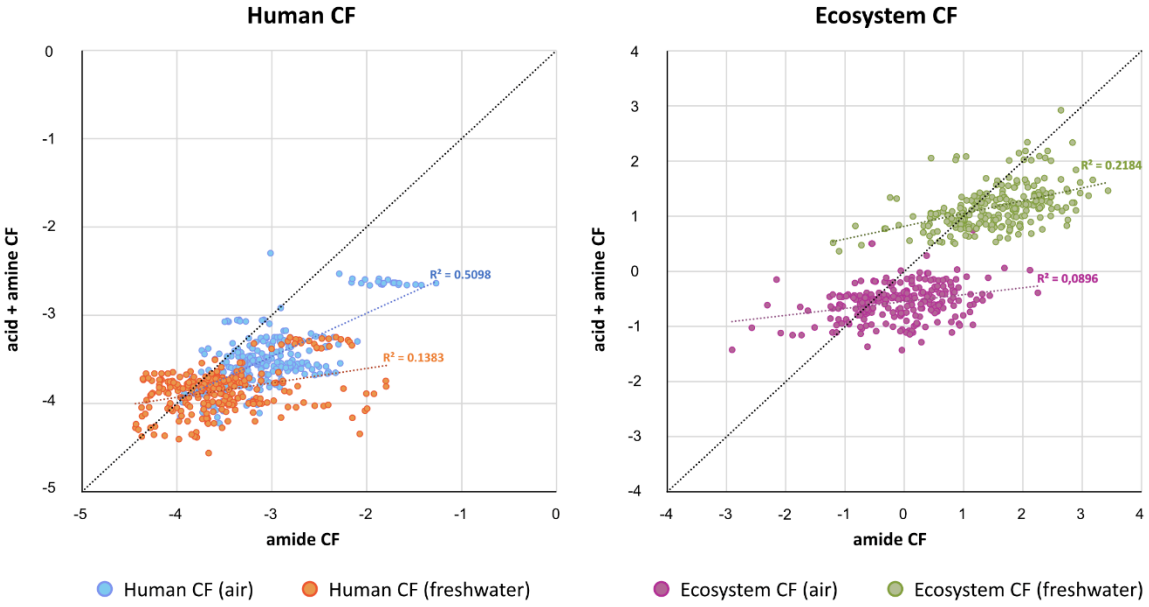


Figure S33. Scatter plot of the USEtox CFs of amides with *in silico* hazard score equal to or lower than ten (figure 3) compared to the sum of the CFs of their constituent acid and amine moieties. The results demonstrate a clear deviation between the amide CF and the combined CFs of the acid and amine. 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	NC1=CC=CC=C1	2	0	yes
amine8634	NCc1ccccc1S(=O)(=O)N1CCCCC1	3	1	yes
amine729	NC1(c2ccccc2)CCNC1	3	1	yes
amine6672	NC1CCN(S(=O)(=O)c2ccccc2)C1	3	1	yes
amine6328	NCCCc1ccccc1	1	1	yes
amine516	NCCCN1Cc2ccccc2C1	3	1	yes
amine3327	NCc1ccccc1CN1CCCCC1	2	1	yes
amine3289	NCCc1ccc(CCN)cc1	4	1	yes
amine2540	NCc1ccccc1CN1CCCC1CO	3	1	yes
amine2119	NCCc1ccc(CO)cc1	2	1	yes
amine14086	Cc1ccccc1CCCN	2	1	yes
amine13521	CN(CCN)Cc1ccccc1	1	1	yes
amine13034	NCCc1ccccc1O	3	0	yes
amine12828	NCc1ccc(S(=O)(=O)NCC2CCCO2)cc1	2	1	yes
amine1217	COCCNC(=O)C(N)Cc1ccccc1	4	1	yes
amine11946	CN(C)S(=O)(=O)Cc1ccc(CN)cc1	3	1	yes
2-aminopyridine	NC1=CC=CC=N1	2	1	yes
amine9676	NC1(c2ccccc2)CCC(=O)NC1	4	1	
amine8732	NC1CCCS(=O)(=O)c2ccccc21	3	1	
amine8721	COc1ccc(S(=O)(=O)N2CCC(N)CC2)cc1	2	1	
amine8666	NC1CCc2cc(C(=O)O)ccc2C1	4	1	
amine8369	NC1CCCN(c2ccc(C(=O)O)cc2)C1	3	1	
amine7654	CC(c1ccc(S(C)(=O)=O)cc1)N1CCC(N)CC1	3	1	
amine7539	NCCCCc1ccccc1	1	1	
amine6912	C=Cc1ccc(CCN)cc1	2	1	
amine6247	Cc1cc(C)cc(C(O)CN)c1	2	1	
amine5745	CN(CCN)S(=O)(=O)c1ccc(F)cc1	4	1	
amine5247	NCCCc1cccc(O)c1	2	1	

amine5086	COc1ccc(CN)cc1	3	1	
amine4400	COc1ccc(CCN)c(F)c1	3	1	
amine429	NCC(O)c1ccccc1	1	0	
amine4167	Cc1cccc(N2CCCC(N)C2=O)c1	2	1	
amine3884	NCCN(CC(=O)O)Cc1ccccc1	4	1	
amine3444	CC(N)c1ccc(N2CCCS2(=O)=O)cc1	3	1	
amine3280	CCN(CC)Cc1ccccc1CN	4	1	
amine3274	Cc1ccc(CCN)cc1	2	1	
amine3222	CCN(CC)Cc1ccc(CN)cc1	3	1	
amine3084	CCN(CCCN)c1ccccc1	4	1	
amine3076	CN(CCCN)Cc1ccccc1	1	1	
amine3005	COC(=O)C(N)c1cc(C)cc(C)c1	3	1	
amine2998	COC(=O)C(N)c1ccc2c(c1)CCO2	3	1	
amine1943	NC(c1ccccc1)C1CCOC1	4	1	
amine14452	NC1CCCN(C(=O)c2ccccc2)C1	4	1	
amine14127	CN(C)CCOc1ccccc1CN	4	1	
amine14103	CN(C)c1ccc(CCN)cc1	4	1	
amine14084	NCCc1cccc(O)c1	2	1	
amine13680	NCCN1Cc2ccccc2C1	3	1	
amine1335	NCCc1ccccc1F	2	1	
amine13327	COc1ccc(CCCN)cc1	3	1	
amine13246	NCCc1ccc(O)cc1	2	1	
amine13174	NCCCOCCc1ccccc1	1	1	
amine12952	NC(c1ccccc1)C1CCCO1	3	1	
amine12947	CCN(CC)Cc1cccc(CN)c1	3	1	
amine12928	CN(C)CCOc1cccc(CN)c1	4	1	
amine12902	CNS(=O)(=O)c1cc(F)c(Cl)c(CN)c1	4	2	
amine12829	Nc1cccc(C(N)CC(=O)O)c1	3	1	
amine1276	CCc1ccccc1CCN	1	1	
amine11623	CC(N)(CO)c1cccc(N)c1	3	1	

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

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

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

4-methoxybenzoic acid	COc1ccc(C(=O)O)cc3	4	0	yes
acid22013	NC1Cc2ccc(C(=O)O)cc2C1	2	1	
acid9525	O=C(O)c1c(F)cccc1F	2	0	
acid8330	O=C(O)c1ccc(C(=O)O)cc1	2	0	
acid7972	O=C(O)c1ccc2c(c1)CCNCC2	4	1	
acid7295	O=C(O)c1ccc2c(c1)CC2	3	0	
acid4638	O=C(O)c1ccc2c(c1)CNC2	4	1	
acid6413	CC(=O)C(NS(=O)(=O)c1ccc(C)c(C)c1)C(O)O	4	1	
acid7343	CCc1ccccc1CC(=O)O	3	0	
acid2256	CCc1ccccc1C(=O)O	2	0	
acid6407	CC(=O)C(NS(=O)(=O)c1ccc(C)cc1)C(O)O	4	1	
acid8776	Cc1cc(O)cc(C(O)O)c1C(O)O	4	1	
acid6280	COc1cc(C=O)ccc1C(O)O	4	0	
Nicotinic acid	O=C(O)c1cccnc3	2	0	
hydrocinnamic acid	O=C(O)CCc1ccccc3	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-benzoicacid	**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 A1 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 $_1$ 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, Piraee 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.