

Supporting Information (SI)

for

Rational design of flavonoid production routes using combinatorial and precursor-directed biosynthesis

Johann E. Kufs^{1,2*}, Sandra Hoefgen^{1*}, Julia Rautschek¹, Alexander U. Bissell^{1,2}, Carola Graf³,
Jonas Fiedler^{1,2}, Daniel Braga⁴, Lars Regestein⁵, Miriam A. Rosenbaum^{2,5}, Julian Thiele³,
and Vito Valiante^{1#}

¹Leibniz Research Cluster Group “Biobricks of Microbial Natural Product Syntheses”, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute, Jena, Germany; ²Faculty of Biological Sciences, Friedrich Schiller University, Jena, Germany; ³Leibniz Research Cluster Group “Polymer Micro(bio)reactors”, Leibniz Institute of Polymer Research, Dresden, Germany; ⁴Synthetic Microbiology Group, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute, Jena, Germany; ⁵Bio Pilot Plant, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute, Jena, Germany

*Both authors contributed equally to this work. #Corresponding author: vito.valiante@leibniz-hki.de.

Table of Contents

Figure S1: pMGE-T7 plasmid features	2
Figure S2: MS ² analyses of produced polyketides from <i>E. coli</i> pJK02	3
Figure S3: Quantification of naringenin heterologously produced by <i>E. coli</i>	4
Figure S4: Calibration curve of molybdate-based PP _i assay	5
Figure S5: Activity of the coumaroyl-CoA ligase Nt4CL2.....	6
Figure S6: LC-HRMS identification of CoA esters produced by Nt4CL2	7
Figure S7: LC-HRMS identification of flavanone derivatives produced in <i>E. coli</i>	9
Figure S8: Activity of the coumaroyl-CoA synthetase Nt4CL2 incubated with pyridine derivatives....	11
Figure S9: Scale-up of 2,3-(methylenedioxy)pinocembrin into a stirred tank bioreactor	12
Figure S10: NMR data of isolated 2,3-(methylenedioxy)pinocembrin	13
Figure S11: Purification of AtCHS.....	19
Figure S12: <i>In vitro</i> production of pinocembrin derivatives using Nt4CL2 and AtCHS	20
Figure S13: LC-HRMS measurements of chemically synthesized CoA esters.....	21
Figure S14: <i>In vitro</i> production of pinocembrin derivatives using synthetic CoA esters	22
Figure S15: <i>In vivo</i> production of glycosylated and hydroxylated flavonoids.....	23
Table S1: Oligonucleotides used in this study	24
Table S2: Plasmids used in this study	24
Supplementing Materials and Methods	25

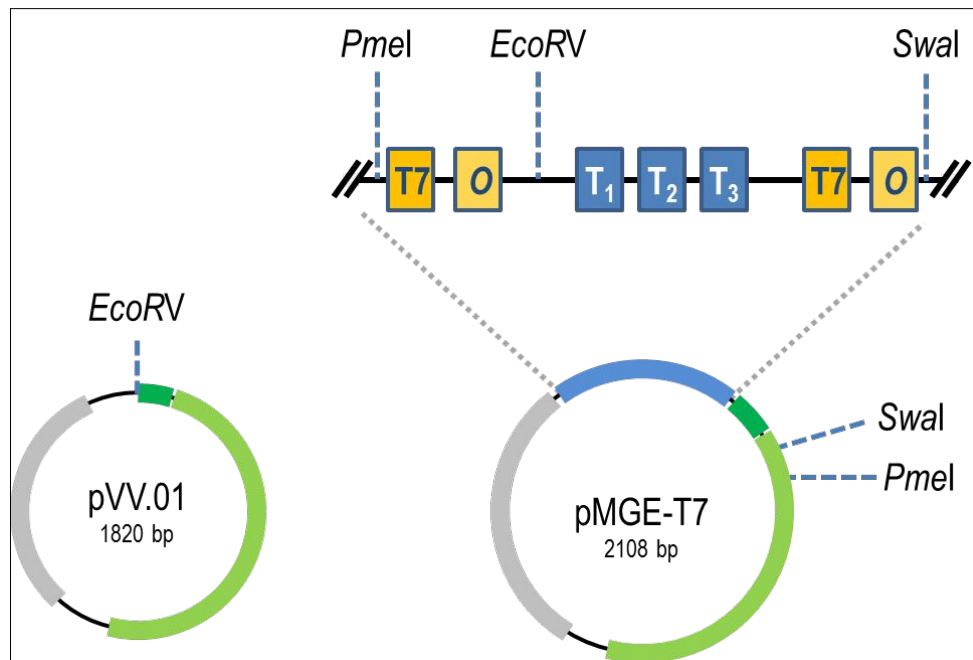


Figure S1 | pMGE-T7 plasmid features. The DNA fragment containing the T7/lacO repetitions, an *EcoRV* restriction site and the three terminators, was synthetically produced and inserted in the pVV.01 plasmid¹. The kanamycin resistance cassette is reported in green while the pUC *ORI* sequence in gray.

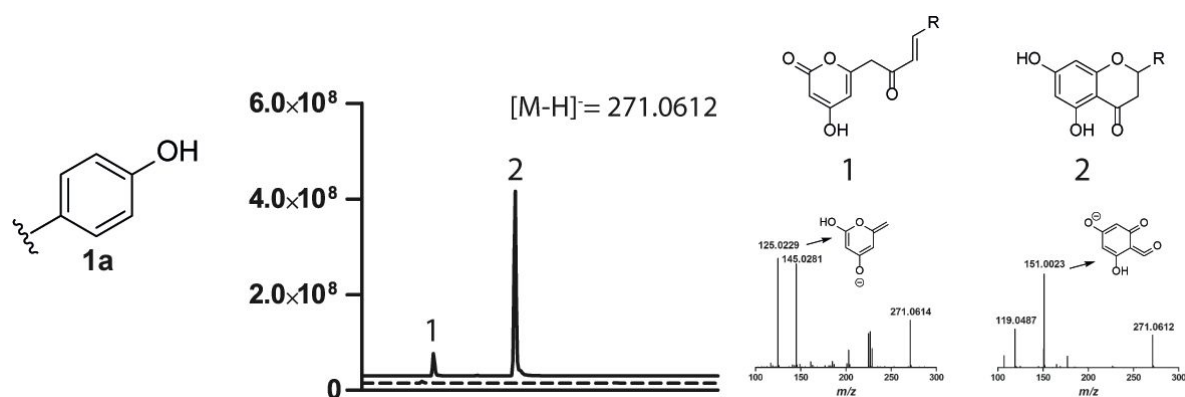


Figure S2 | MS^2 analyses of produced polyketides from *E. coli* pJK02. Extracted ion chromatogram for m/z 271.0612 $[M-H]^-$ in the negative mode and ESI- MS^2 spectra for the peaks 1 and 2. Characteristic fragments were detected for *p*-coumaroyltriacetic acid lactone (**1**) and naringenin (**2**) with m/z 125.0244 $[M-H]^-$ and 151.0037 $[M-H]^-$, respectively².

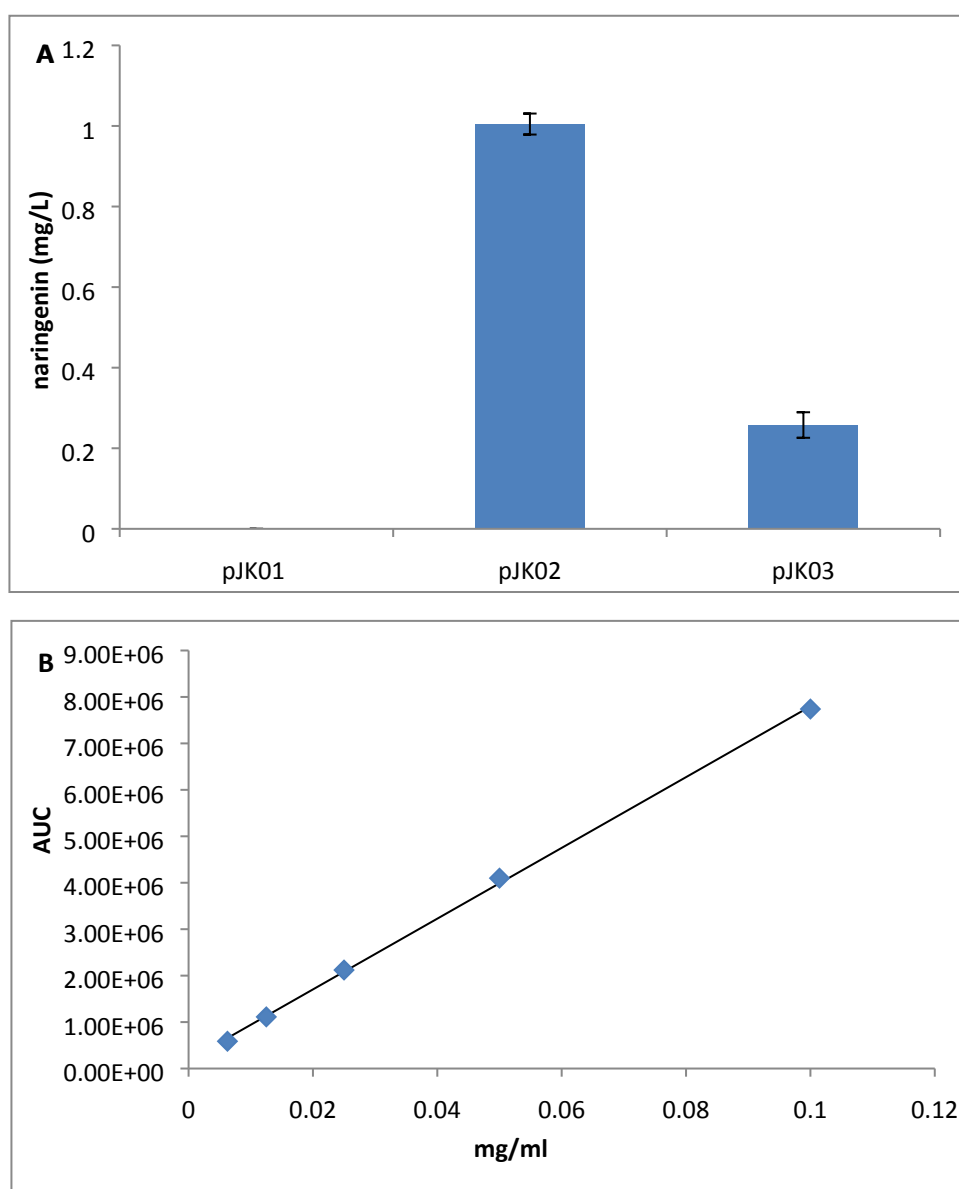


Figure S3 | Quantification of naringenin heterologously produced by *E. coli*. A) Naringenin production of *E. coli* strains harboring pJK01, pJK02 and pJK03 after supplementation with 3 mM *p*-coumaric acid for 48 h showing a titer of 0, 1, and 0.26 mg/L naringenin, respectively. Data are shown from three independent biological replicates. Error bars represent the calculated standard deviations. B) Calibration curve of naringenin reference based on the area under the curve (AUC) determined by peak integration upon HPLC analysis.

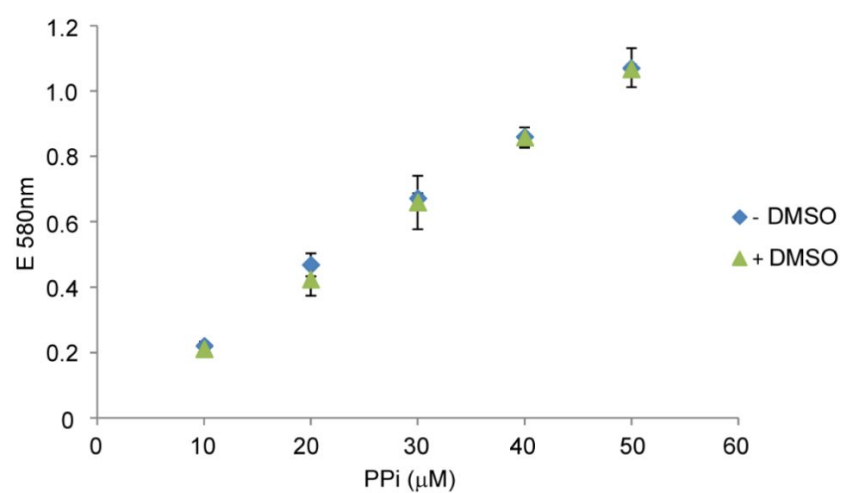


Figure S4 | Calibration curve of molybdate-based PP_i assay. Calibration curve w/o 10% DMSO showing that the addition of DMSO does not influence PP_i determination. Error bars represent standard deviations.

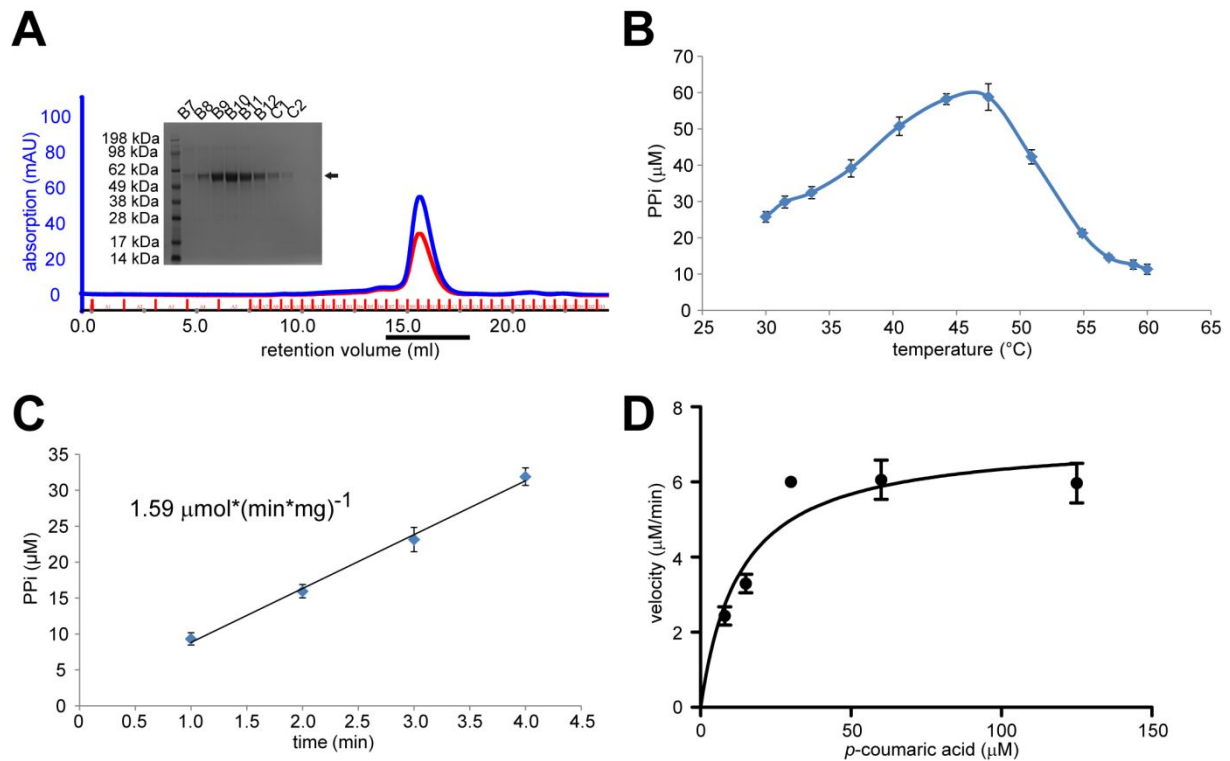


Figure S5 | Activity of the coumaroyl-CoA ligase Nt4CL2. A) SD200 10/300 Increase size exclusion column profile of Nt4CL2 (theoretical MW: 62 kDa) purification. The blue and red curves represent the absorption at 280 nm, 260 nm, respectively. Protein-containing fractions were analyzed *via* Coomassie-stained SDS-PAGE gels. Applied fractions are marked with a black bar. B) Nt4CL2 optimal reaction temperature measurements determined by using the molybdate-based assay and *p*-coumaric acid as substrate. C) Starting activity of Nt4CL2 at 47 $^{\circ}\text{C}$ using *p*-coumaric acid as substrate. D) For the calculation of kinetic parameters, starting activities of Nt4CL2 with increasing concentrations of *p*-coumaric acid were determined. $K_m = 13.12 \pm 3.71 \mu\text{M}$ and $K_{cat} = 3.73 \pm 0.30 \text{ s}^{-1}$. Error bars represent the standard error from three independent biological and technical replicates.

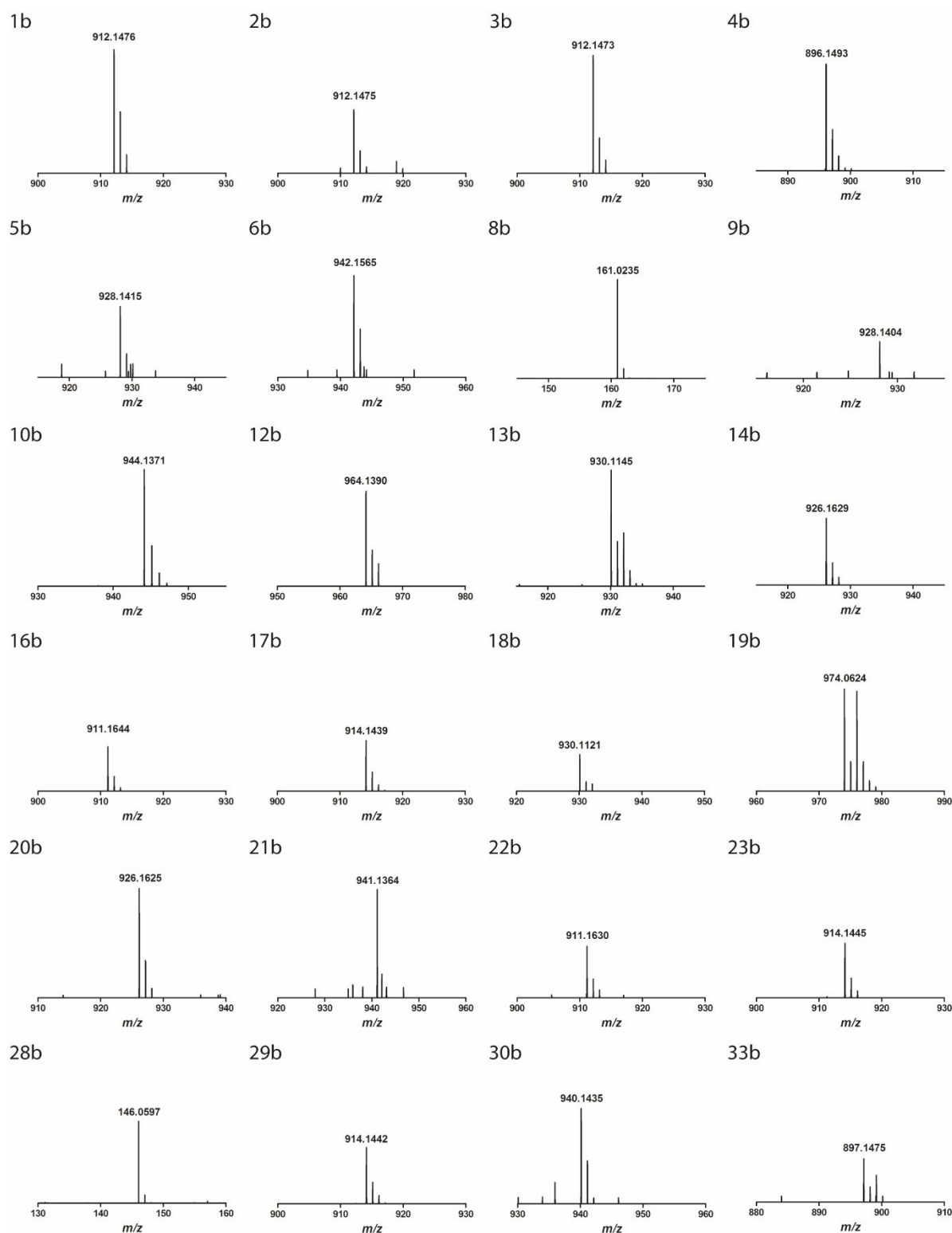


Figure S6 | LC-HRMS identification of CoA esters produced by Nt4CL2. Cinnamoyl-CoA ester formation was confirmed based on the calculated m/z values: *p*-coumaroyl-CoA (**1b**: 912.1447 [M-H]⁻); *m*-coumaroyl-CoA (**2b**: 912.1447 [M-H]⁻); *o*-coumaroyl-CoA (**3b**: 912.1447 [M-H]⁻); cinnamoyl-CoA (**4b**: 896.1498 [M-H]⁻); caffeoyl-CoA (**5b**: 928.1396 [M-H]⁻); feruoyl-CoA (**6b**: 942.1553 [M-H]⁻); 3,5-dihydroxycinnamoyl-CoA (**9b**: 928.1396 [M-H]⁻); 3,4,5-

trihydroxycinnamoyl-CoA (**10b**: 944.1345 [M-H]⁻); 4-(trifluoromethyl)cinnamoyl-CoA (**12b**: 964.1372 [M-H]⁻); 4-chlorocinnamoyl-CoA (**13b**: 930.1108 [M-H]⁻); 4-methoxycinnamoyl-CoA (**14b**: 926.1604 [M-H]⁻); 4-aminocinnamoyl-CoA (**16b**: 911.1607 [M-H]⁻); 4-fluorocinnamoyl-CoA (**17b**: 914.1404 [M-H]⁻); 3-chlorocinnamoyl-CoA (**18b**: 930.1108 [M-H]⁻); 3-bromocinnamoyl-CoA (**19b**: 974.0603 [M-H]⁻); 3-methoxycinnamoyl-CoA (**20b**: 926.1604 [M-H]⁻); 3-nitrocinnamoyl-CoA (**21b**: 942.1427 [M-H]⁻); 3-aminocinnamoyl-CoA (**22b**: 911.1607 [M-H]⁻); 3-fluorocinnamoyl-CoA (**23b**: 914.1404 [M-H]⁻); 2-fluorocinnamoyl-CoA (**29b**: 914.1404 [M-H]⁻); 2,3-(methylenedioxy)cinnamoyl-CoA (**30b**: 940.1396 [M-H]⁻) and 3-(2-pyridyl)acryloyl-CoA (**33b**: 897.1450 [M-H]⁻). For 2,4-dihydroxycinnamoyl-CoA (**8b**) and 2-aminocinnamoyl-CoA (**28b**), no CoA ester was detected due to intramolecular lactonization that is represented by *m/z* 161.0235 [M-H]⁻ and 146.0597 [M-H]⁻, respectively.

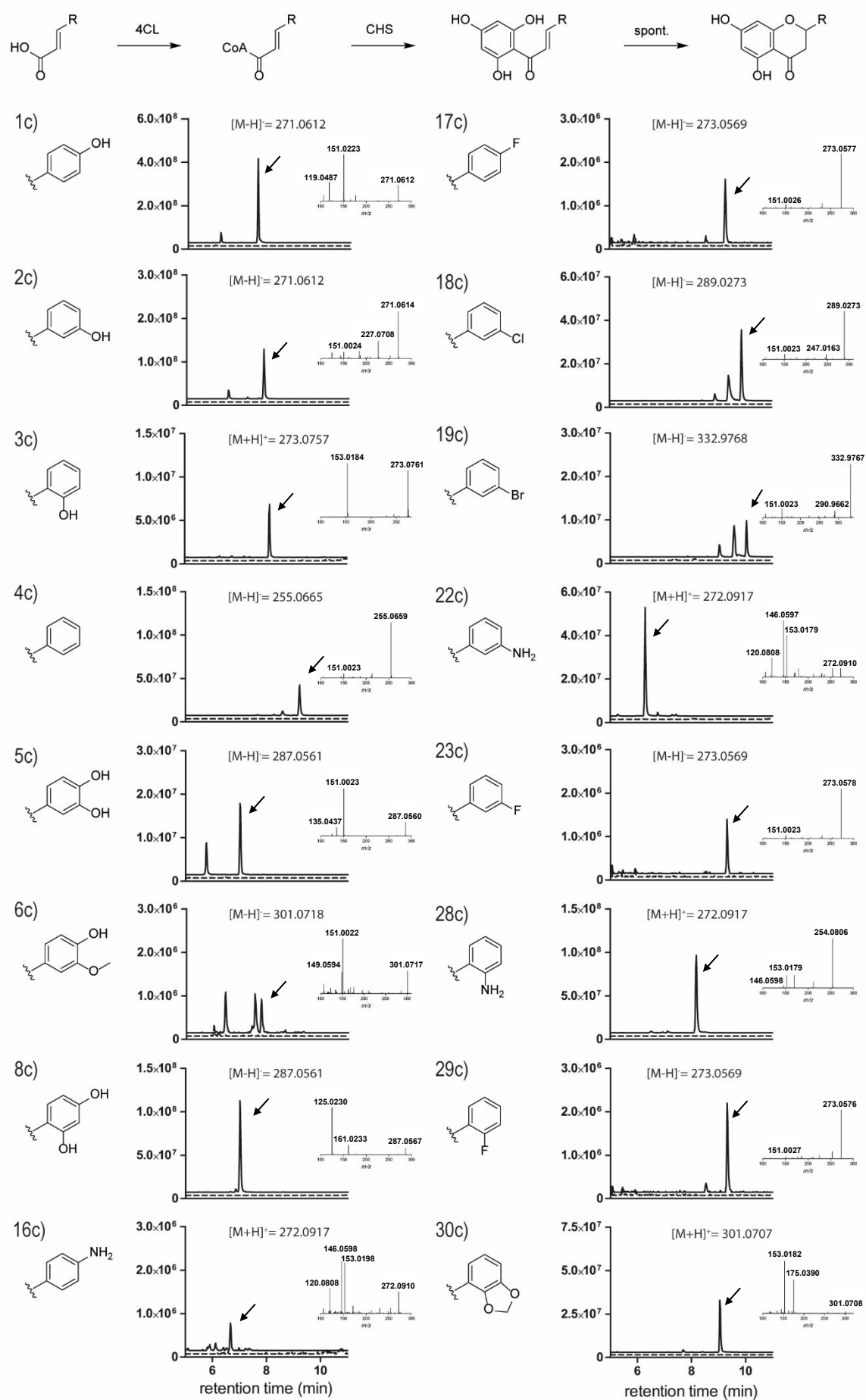


Figure S7 | LC-HRMS identification of flavanone derivatives produced in *E. coli*.

Extracted ion chromatograms of engineered *E. coli* strain pJK02 (solid line), expressing Nt4CL2 and AtCHS, grown in the presence of **1a, 2a, 3a, 4a, 5a, 6a, 8a, 16a, 17a, 18a, 19a, 22a, 23a, 28a, 29a and 30a** and thereby producing naringenin (**1c**); 3-hydroxypinocembrin (**2c**); 2-hydroxypinocembrin (**3c**); pinocembrin (**4c**); eriodictyol (**5c**); homoeriodictyol (**6c**); 2,4-dihydroxypinocembrin (**8c**); 4-aminopinocembrin (**16c**); 4-fluoropinocembrin (**17c**); 3-chloropinocembrin (**18c**); 3-bromopinocembrin (**19c**); 3-aminopinocembrin (**22c**); 3-fluoropinocembrin (**23c**); 2-aminopinocembrin (**28c**), 2-fluoropinocembrin (**29c**) and 2,3-(methylenedioxy)pinocembrin (**30c**), respectively. Presence of the corresponding flavanones was confirmed by MS² analysis showing the characteristic fragments at m/z 151.0023 [M-H]⁻ in the negative mode and 153.0198 [M+H]⁺ in the positive mode^{2, 3}. The *E. coli* strain pJK01 served as negative control (dashed line).

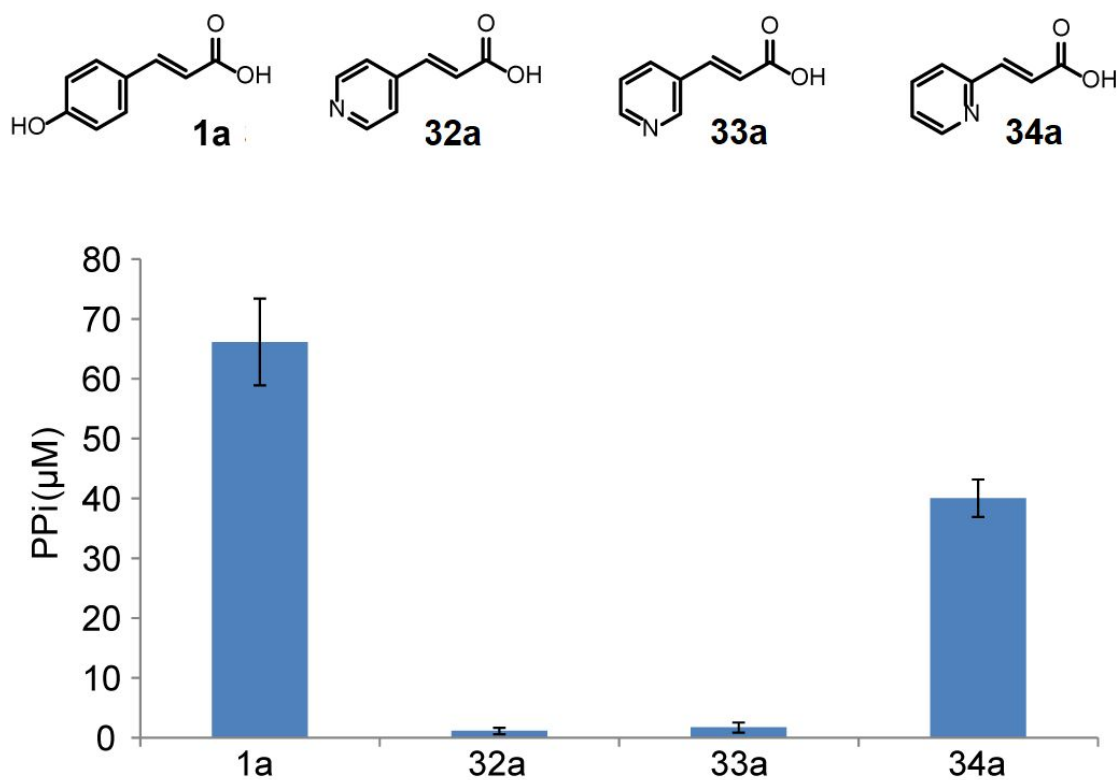


Figure S8 | Activity of the coumaroyl-CoA ligase Nt4CL2 incubated with pyridine derivatives. The purified Nt4CL2 was incubated with 3-(4-pyridyl)acrylic acid (**32a**), 3-(3-pyridyl)acrylic acid (**33a**) and 3-(2-pyridyl)acrylic acid (**34a**). The ester formation was detected only for **34a**. *p*-coumaric acid (**1a**) was used as positive control in the reaction.

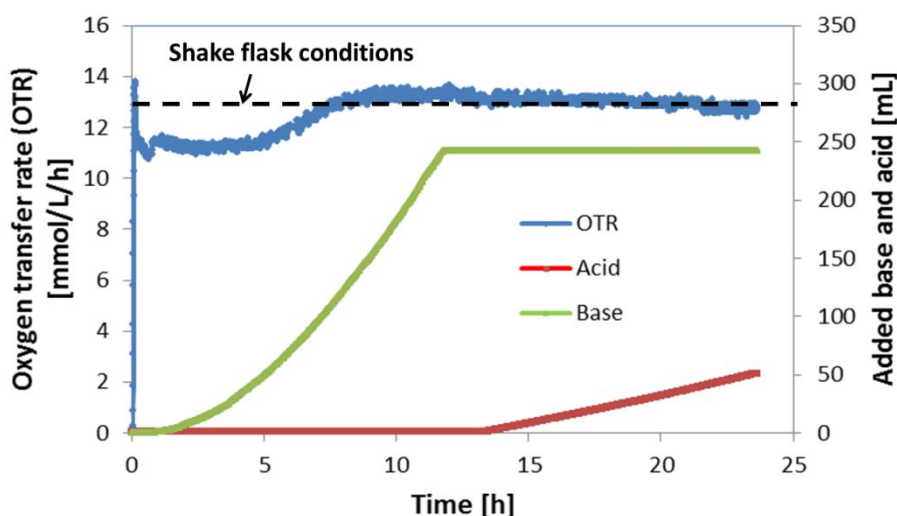
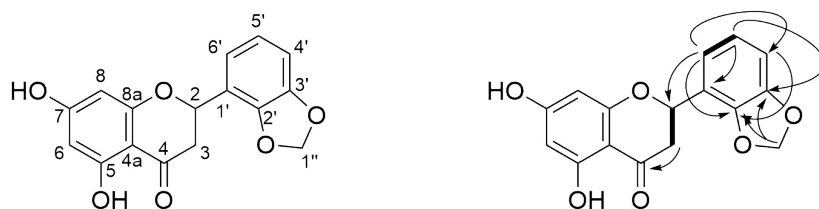


Figure S9 | Scale-up of 2,3-(methylenedioxy)pinocembrin into a stirred tank bioreactor.

Oxygen transfer rate of a batch culture of *E. coli* BL21 (DE3) in a 7 L stirred tank reactor. Experimental conditions: gas flow rate = 1.25 L/min; stirring rate = 470 rpm; temperature = 30 °C; pH controlled at 7 ± 0.05 by addition of NH_4OH and sulfuric acid. The scale-up of previous shake flasks experiments was performed on the criterion of a similar oxygen transfer rate. For this reason, the maximum oxygen transfer capacity of the shake flask experiments was calculated on an equation published by Meier *et al.* (2016)⁴. The calculated maximum oxygen transfer capacity under shake flask conditions was 13 mmol/L/h. To realize a similar value under the conditions of a stirred tank reactor, the maximum stirring speed was limited to 470 rpm and a gas flow rate of 1.25 L/min, which reflects the oxygen diffusion through the cotton plug in the shake flask experiments. As depicted here, the maximum measured oxygen transfer rate (blue) was 13 mmol/L/h, which is in agreement with the previously calculated value. Therefore, the scale-up from small-scale shake flask conditions into lab-scale stirred tank reactor was successfully conducted. With respect to the kinetic of the product, it could be demonstrated that oxygen limited conditions support the formation of the desired flavanone.

Figure S10 | NMR data of isolated 2,3-(methylenedioxy)pinocembrin. NMR spectra were recorded on a Bruker Avance III spectrometer at 600 MHz for ^1H NMR and 150 MHz for ^{13}C NMR using CD_3OD as a solvent. Chemical shifts are shown with reference to CD_3OD as 3.31 ppm for ^1H NMR and 49.05 ppm for ^{13}C NMR. The following abbreviations are used: d, doublet; dd, doublet of doublets.

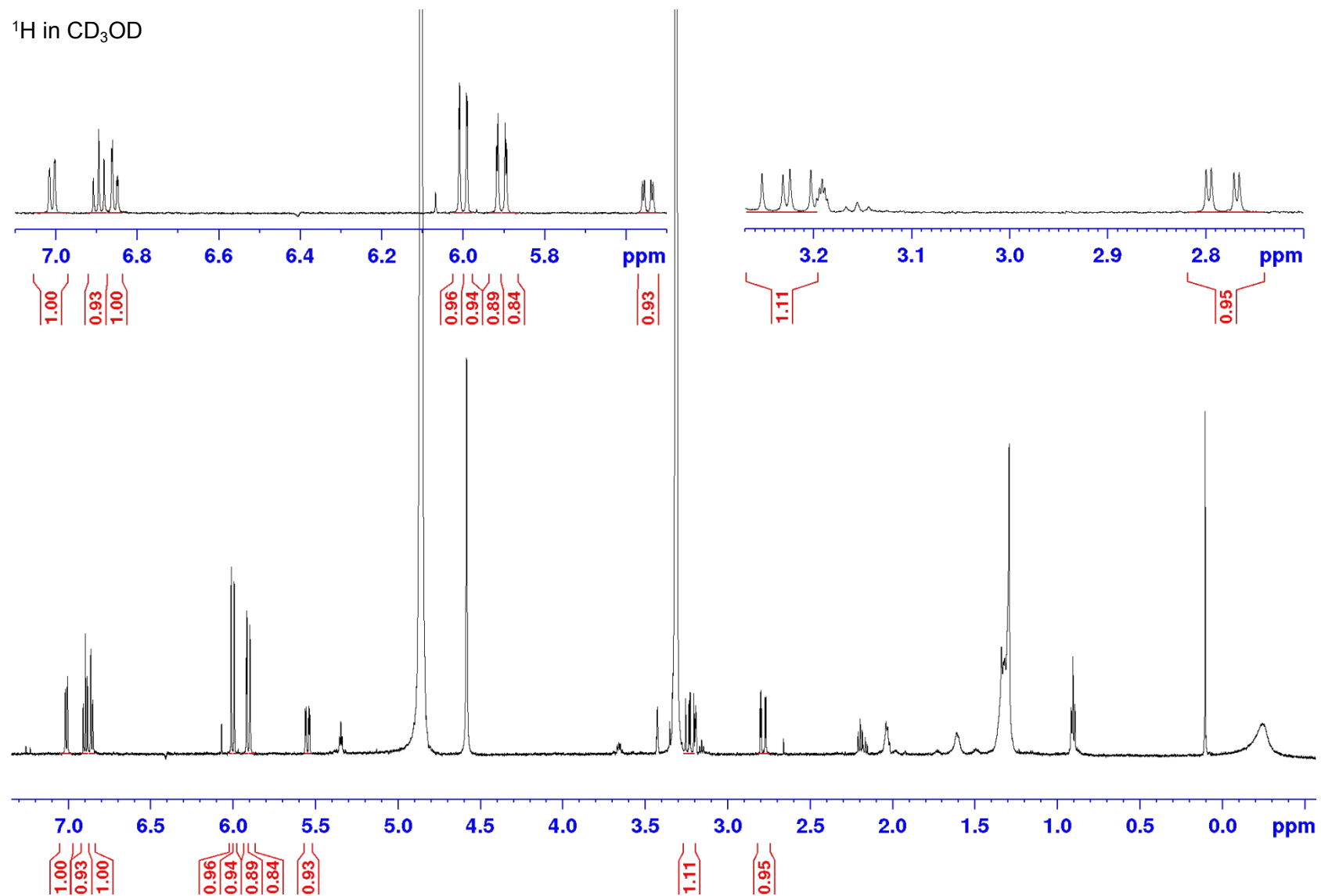


Position	δ_{C} [ppm]	δ_{H} [ppm]	Multiplicity	J_{HH} [Hz]
2	76.13	5.55	dd	12.8, 3.1
3	42.28	2.78	dd	17.1, 3.1
		3.23	dd	17.1, 12.8
4	$\sim 197.1^a$	-	-	-
4a	$\sim 97.8^a$	-	-	-
5	$-^b$	-	-	-
6	$\sim 96.1^{c,d}$	5.90^e	d	2.2
7	$-^b$	-	-	-
8	$\sim 96.7^{c,d}$	5.92^e	d	2.2
8a	$-^b$	-	-	-
1'	~ 123.8	-	-	-
2'	$\sim 146.1^a$	-	-	-
3'	$\sim 149.3^a$	-	-	-
4'	109.83	6.86	dd	7.8, 1.2
5'	123.06	6.89	d	7.8
6'	120.83	7.01	dd	7.8, 1.2
1''	102.73	5.99	d	1.1
		6.01	d	1.1

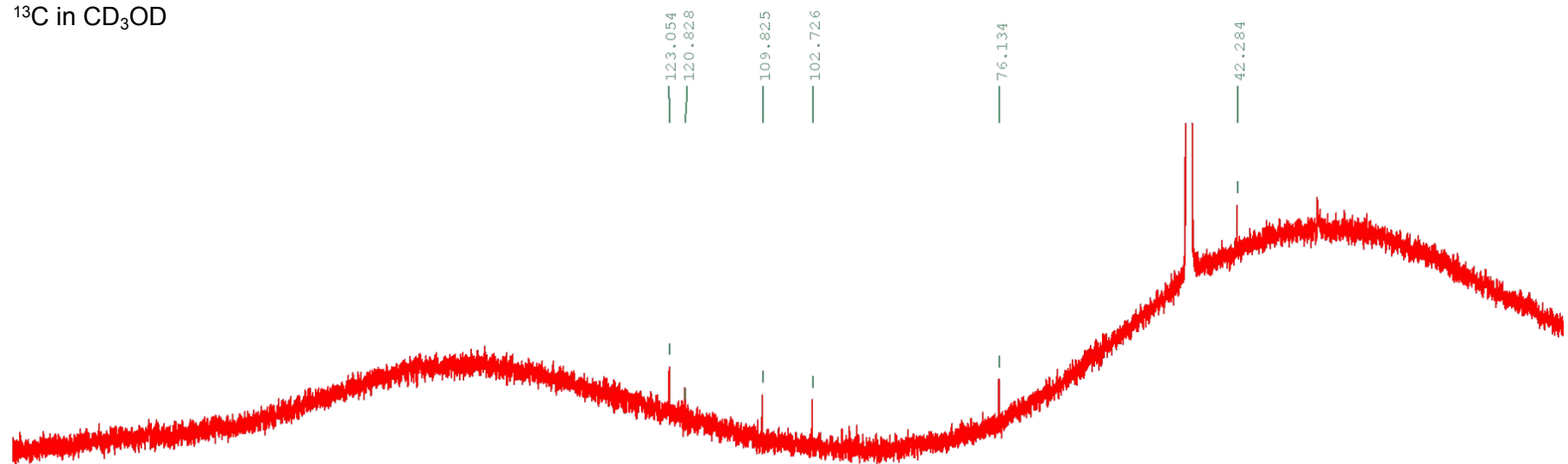
^aChemical shifts determined by HMBC correlation. ^bCould not be determined. An HMBC correlation $\delta_{\text{H}} = 5.90$ ppm (H-6/8)/ $\delta_{\text{C}} = 164.9$ ppm was observed, so that C-5, 7 or 9 (or multiple) have $\delta_{\text{C}} = 164.9$.

^cChemical shifts are exchangeable. ^dChemical shifts determined by HMQC correlation. ^eChemical shifts are exchangeable.

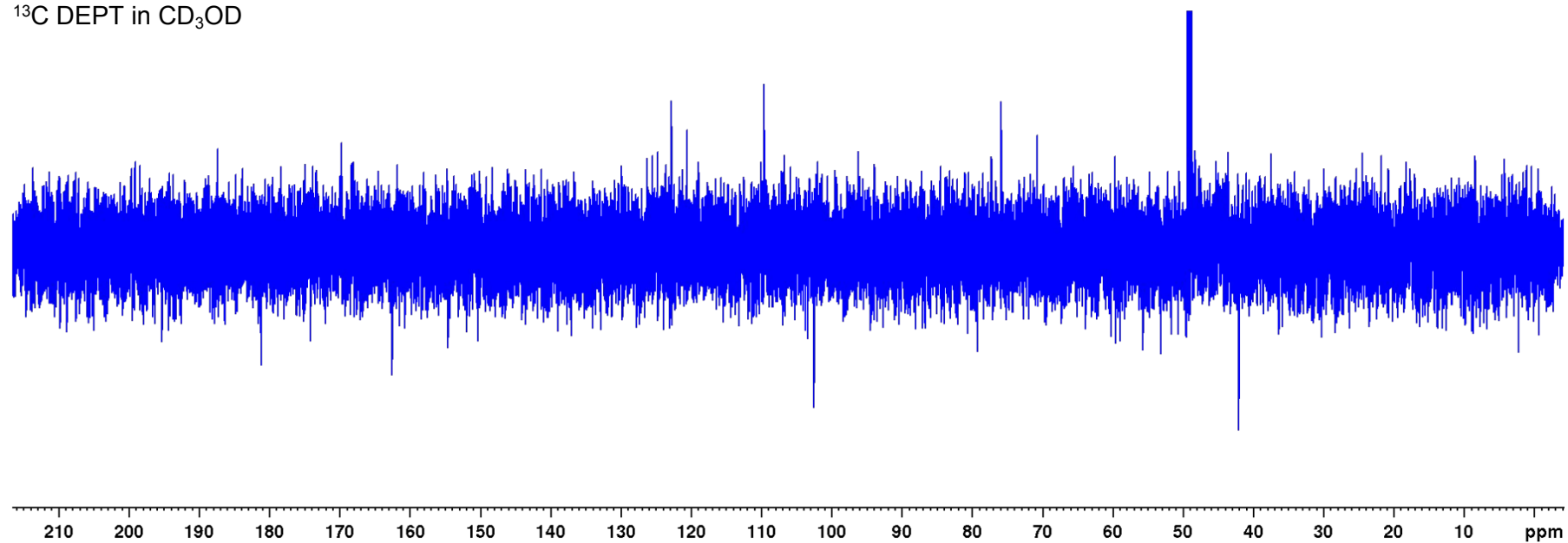
^1H in CD_3OD



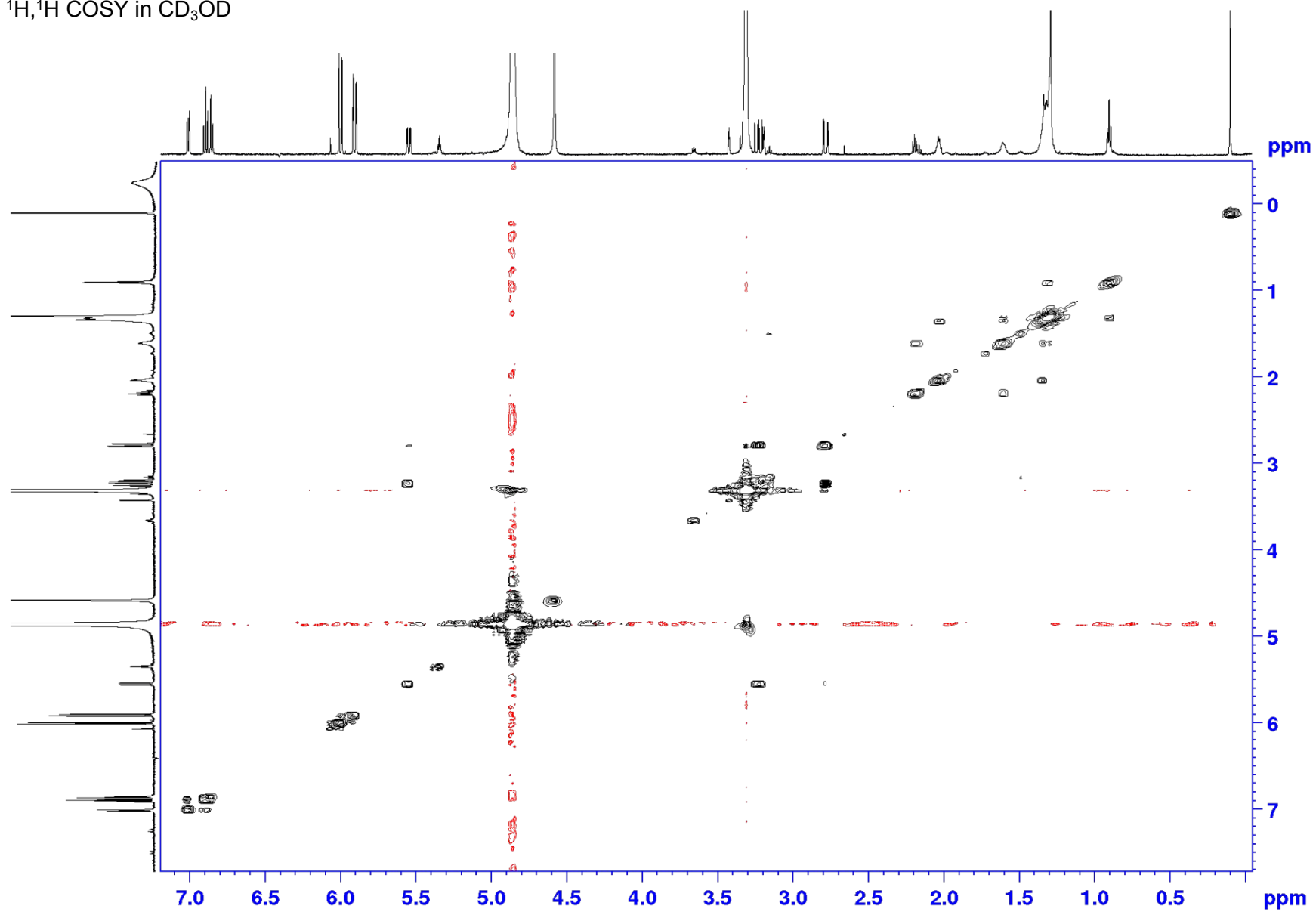
^{13}C in CD_3OD



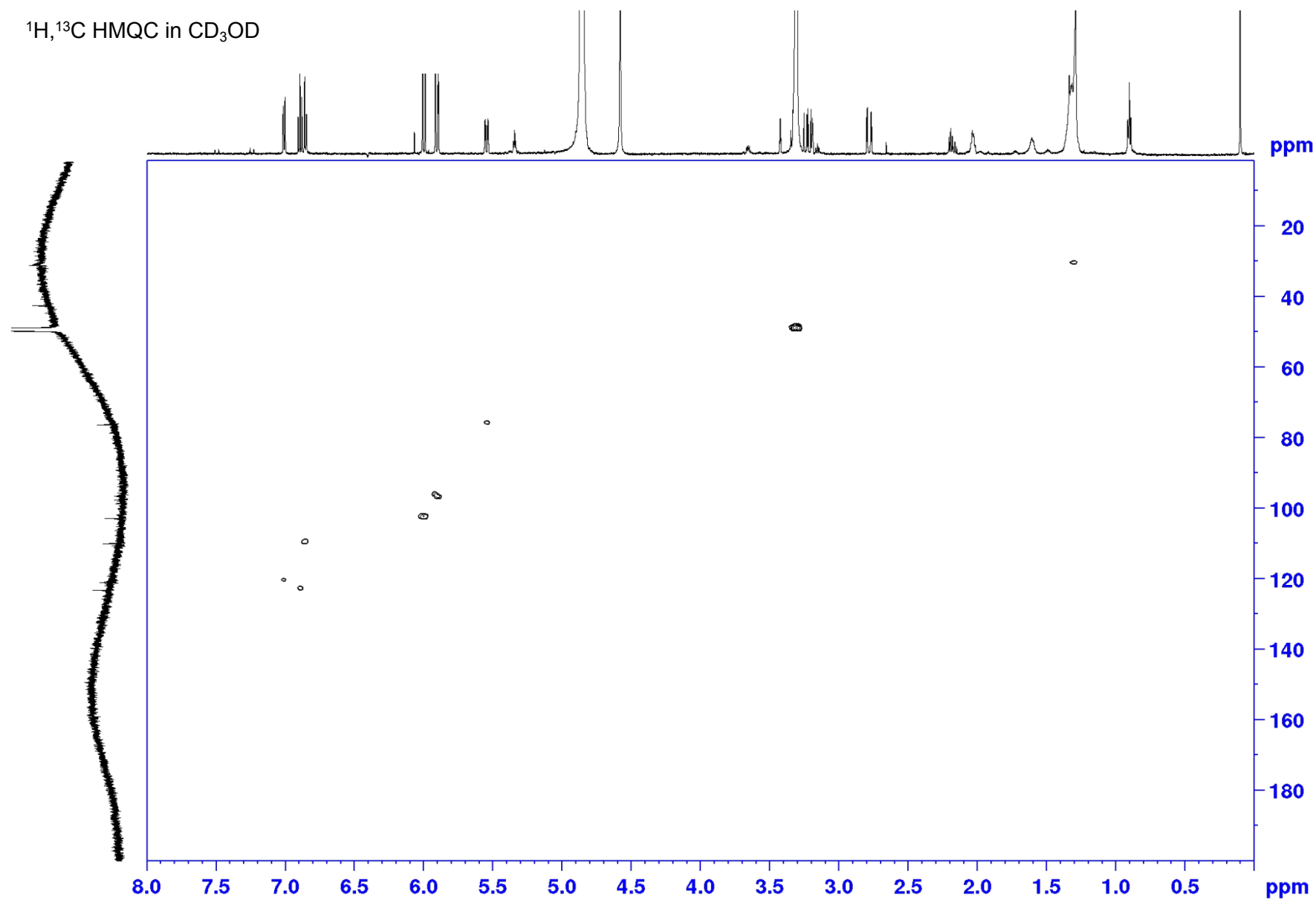
^{13}C DEPT in CD_3OD



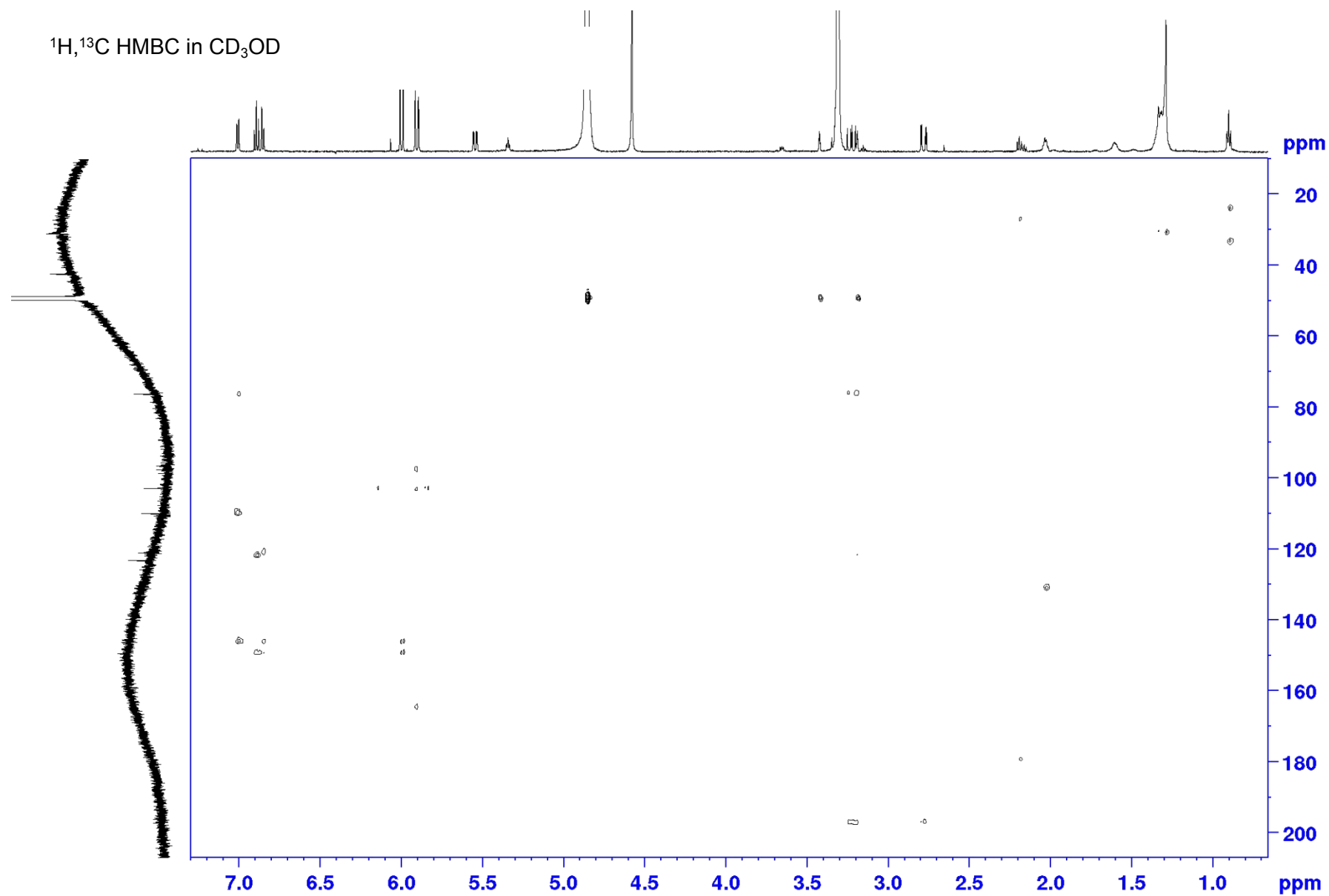
$^1\text{H}, ^1\text{H}$ COSY in CD_3OD



$^1\text{H}, ^{13}\text{C}$ HMQC in CD_3OD



$^1\text{H}, ^{13}\text{C}$ HMBC in CD_3OD



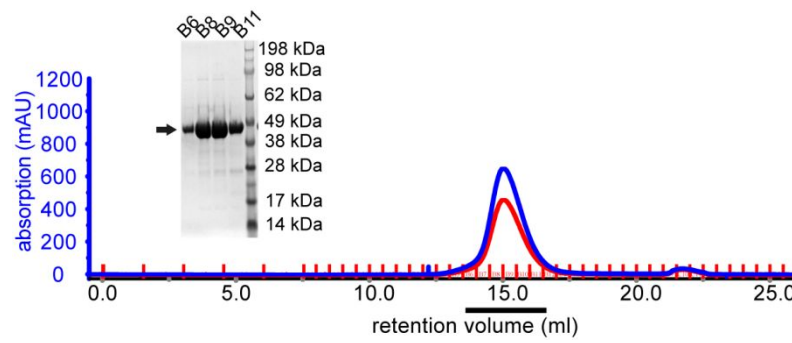


Figure S11 | Purification of AtCHS. Size exclusion profile of AtCHS (theoretical MW: 46.1 kDa) purification. The blue and red curves represent absorption at 280 nm and 260 nm, respectively. Protein-containing fractions were analyzed *via* Coomassie-stained SDS-PAGE gels and applied fractions are marked with a black bar.

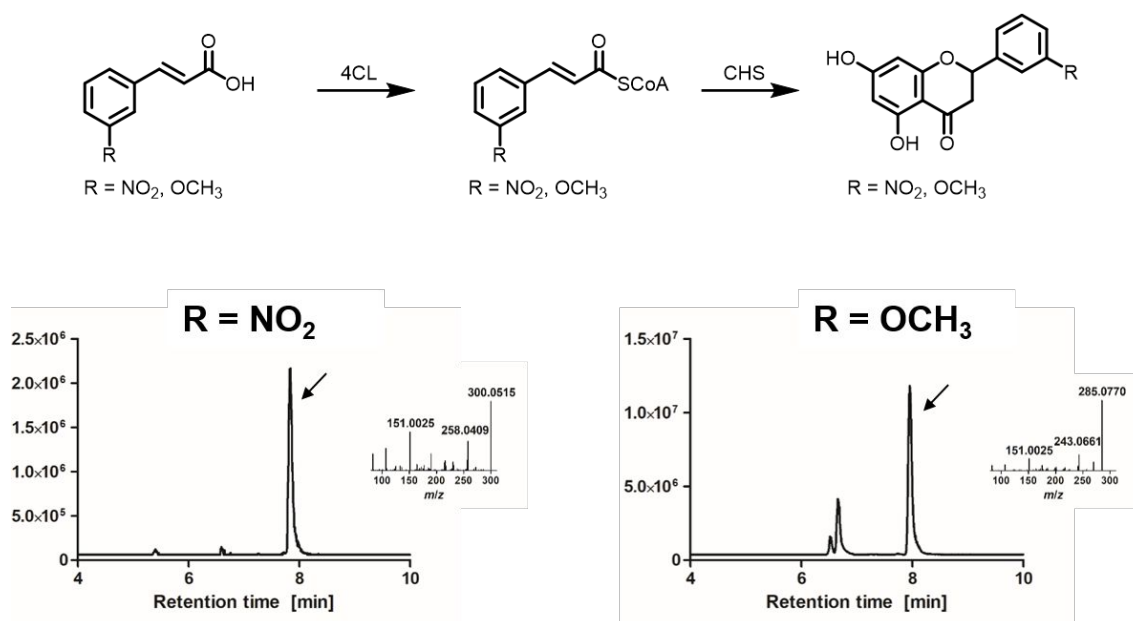


Figure S12 | *In vitro* production of pinocembrin derivatives using Nt4CL2 and AtCHS.

The cinnamic acid analogues 3-methoxycinnamic acid (**20a**) and 3-nitrocinnamic acid (**21a**) were esterified by the Nt4CL2, but not incorporated into the corresponding flavanones *in vivo*. To exclude uptake issues, these substrates were used together with CoA and malonyl-CoA for an *in vitro* assay with purified Nt4CL2 and AtCHS. LC-HRMS and MS² analyses confirmed the incorporation of both analogues into 3-methoxypinocembrin (**20c**) and 3-nitropinocembrin (**21c**).

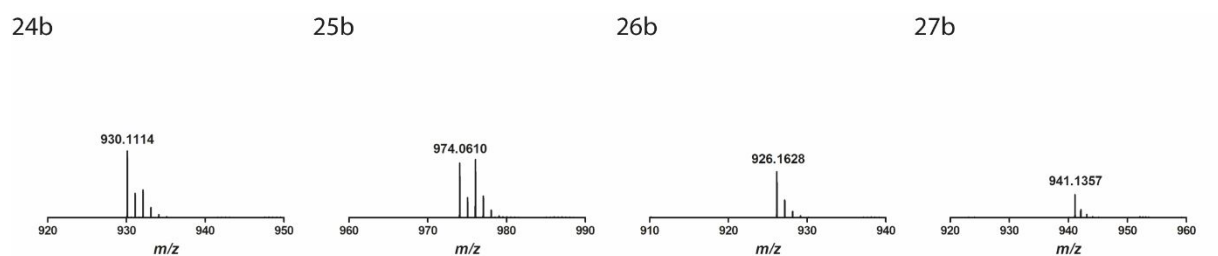


Figure S13 | LC-HRMS measurements of chemically synthesized CoA esters. Cinnamoyl-CoA esters harboring *ortho* substitutions that could not be obtained enzymatically were synthesized as described before. HRMS data confirmed the presence of 2-chlorocinnamoyl-CoA (**24b**: 930.1108 [M-H]⁻); 2-bromocinnamoyl-CoA (**25b**: 974.0603 [M-H]⁻); 2-methoxycinnamoyl-CoA (**26b**: 926.1604 [M-H]⁻) and 2-nitrocinnamoyl-CoA (**27b**: 942.1357 [M-H]⁻) based on the calculated m/z values.

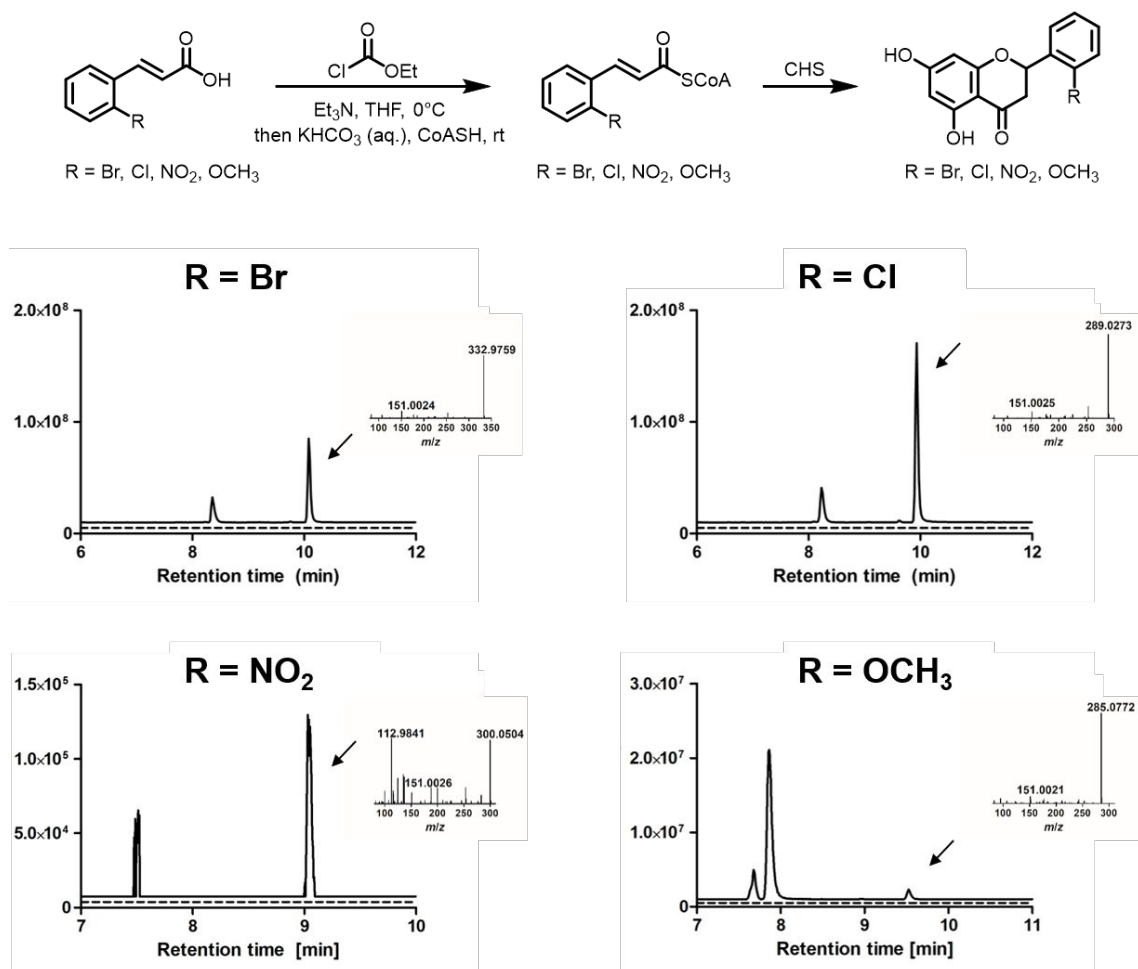


Figure S14 | *In vitro* production of pinocembrin derivatives using synthetic CoA esters.

CoA esters that could not be produced by the Nt4CL2, namely 2-chlorocinnamoyl-CoA (**24b**), 2-bromocinnamoyl-CoA (**25b**), 2-methoxycinnamoyl-CoA (**26b**) and 2-nitrocinnamoyl-CoA (**27b**), were chemically synthesized and used together with malonyl-CoA and the purified AtCHS enzyme (solid line). The HRMS and MS² analyses confirmed the production of the corresponding flavanone derivatives 2-chloropinocembrin (**24c**), 2-bromopinocembrin (**25c**), 2-methoxypinocembrin (**26c**) and 2-nitropinocembrin (**27c**). Separate reactions were performed without the addition of AtCHS as negative control (dashed line).

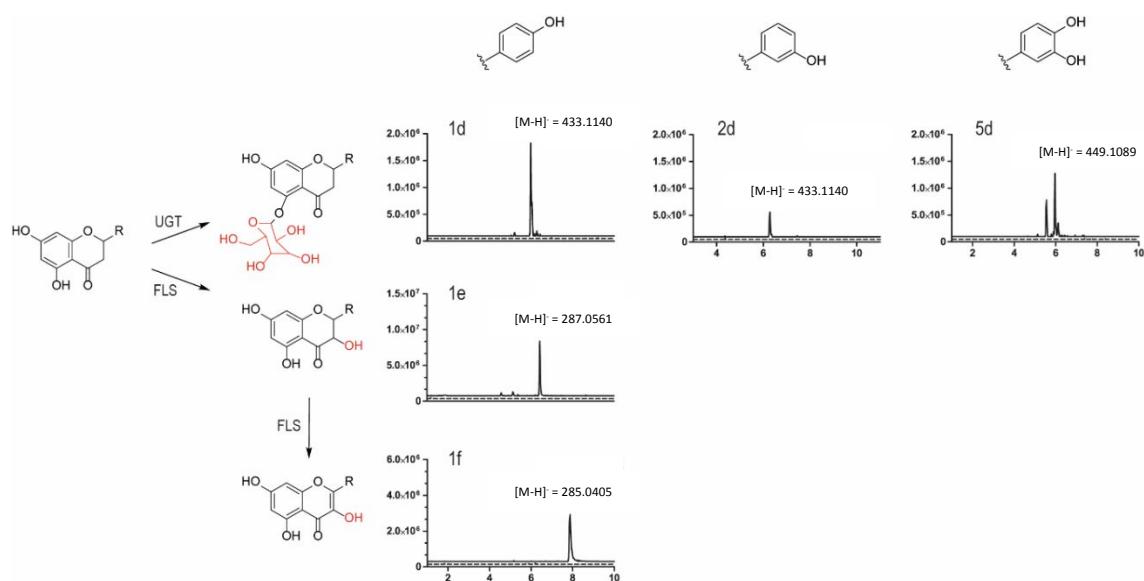


Figure S15 | *In vivo* production of glycosylated and hydroxylated flavonoids. The *E. coli* strain pJK04 expressing *Nt4CL2*, *AtCHS* and *UGT* was grown in the presence of different cinnamic acid analogues. EIC revealed the production of naringenin-7-O-glucoside (**1d**), 3-hydroxypinocembrin-7-O-glucoside (**2d**) and eriodictyol-7-O-glucoside (**5d**) after supplementation with the corresponding substrates **1a**, **2a** and **5a**, respectively. *FLS1* together with *Nt4CL2* and *AtCHS* (pJK05) resulted in the production of dihydrokaempferol (**1e**) and kaempferol (**1f**) in the presence of **1a**. The *E. coli* strain pJK01 served as negative control (dashed line).

Table S1 | Oligonucleotides used in this study.

No	Oligonucleotide	Sequence (5' - 3')
1	Nt4CL2_1	GATAACAATTTGCGCAAGAAGGAGATAGATATGTCGTACTACCATCACCA
2	Nt4CL2_2	CCCCGCGGGGTTGCTTCACTCAGTGCAGATTTAATTTGGAAGCCCAGCAG
3	CHS_1	GATAACAATTTGCGCAAGAAGGAGATAGATATGGTGATGGCTGGTGCTTC
4	CHS_2	CCCCGCGGGGTTGCTTCACTCAGTGCAGATTTAGAGAGGAACGCTGTGCA
5	CI_1	GATAACAATTTGCGCAAGAAGGAGATAGATATGTCTTCATCCAACGCCTG
6	CI_2	CCCCGCGGGGTTGCTTCACTCAGTGCAGATTTAGTTCTCTTTGGCTAGTT
7	FLS1_1	GATAACAATTTGCGCAAGAAGGAGATAGATATGGAGGTCGAAAGAGTCCA
8	FLS1_2	CCCCGCGGGGTTGCTTCACTCAGTGCAGATTCAATCCAGAGGAAGTTTAT
9	UGT_1	GATAACAATTTGCGCAAGAAGGAGATAGATATGGGTCAGCTCCATATTTT
10	UGT_2	CCCCGCGGGGTTGCTTCACTCAGTGCAGATTTAATGACCAGTGGAACAT
11	Nt4CL2_3	ACGACCGAAAACCTGTATTTTCAGGGATCCGAGAAAGATACAAAACAG
12	Nt4CL2_4	GTGGTGCTCGAGTGCGGCCGCAAGCTTTTATTAATTTGGAAGCCCAGCAG
13	CHS_3	ACGACCGAAAACCTGTATTTTCAGGGATCCGTGATGGCTGGTGCTTCTTC
14	CHS_4	GTGGTGCTCGAGTGCGGCCGCAAGCTTTTATTAGAGAGGAACGCTGTGC

Table S2 | Plasmids used in this study.

Plasmid name	Relevant features	Reference
pVV.01	pUC-ORI, <i>kan</i> ^R	1
pMGE-T7	pVV.01, T7 promoter, terminators	This study
pMGE_CHS	pMGE-T7, <i>CHS Arabidopsis thaliana</i>	This study
pMGE_CI	pMGE-T7, <i>CI Arabidopsis thaliana</i>	This study
pMGE_FLS1	pMGE-T7, <i>FLS1 Arabidopsis thaliana</i>	This study
pMGE_UGT	pMGE-T7, <i>UGT Withania somnifera</i>	This study
pJK01	pMGE-T7, <i>4CL2 Nicotiana tabacum</i>	This study
pJK02	pMGE_CHS, <i>4CL2</i>	This study
pJK03	pJK02, <i>CI</i>	This study
pJK04	pJK02, <i>UGT</i>	This study
pJK05	pJK02, <i>FLS1</i>	This study
pJK06	pJK04, <i>FLS1</i>	This study
pET28a_H6TEV	T7 promoter, <i>amp</i> ^R	5
pET28a_Nt4CL2	pET28a_H6TEV, <i>4CL2</i>	This study
pET28a_CHS	pET28a_H6TEV, <i>CHS</i>	This study

SI Materials and Methods

Cloning of expression plasmids. All PCR amplifications were performed using the Phusion Flash High-Fidelity DNA polymerase (Thermo Fisher) and oligonucleotides are shown in Table S1. The DNA sequence of Nt4CL2 from *N. tabacum* (UniProt: O24146)⁶ was synthesized by Thermo Fisher, amplified using primer pair [11]/[12], digested with *Bam*HI and *Hind*III, and cloned into the pET28a_H6TEV expression vector⁵, in frame with an N-terminal 6xHis-tag and a cleavage site for the *Tobacco etch virus* (TEV) protease, resulting in the plasmid pET28a_Nt4CL2. The DNA sequence of CHS from *A. thaliana* (UniProt: P13114)⁷ was synthesized by Biomatik, amplified using primer pair [13]/[14], and cloned into the *Bam*HI and *Hind*III digested pET28a_H6TEV plasmid using the NEBuilder Gibson Master Mix (New England Biolabs) resulting in the pET28a_CHS expression plasmid.

Expression of enzymes for purification. For expression of the two enzymes Nt4CL2 and AtCHS, *E. coli* BL21 (DE3) cells containing either pET28a_Nt4CL2 or pET28a_CHS plasmids were used. A pre-culture with LB medium and 30 µg/mL kanamycin was inoculated with the corresponding cells and grown overnight at 37 °C while shaking. A main culture containing auto-induction medium⁸ and 30 µg/mL kanamycin was inoculated with pre-culture in a 1:100 ratio using baffled shaker flasks. The main culture was incubated at 37 °C while shaking until OD₆₀₀ reached 1.0 and afterwards temperature was lowered to 18 °C for overnight incubation. Cells were harvested by centrifugation and stored at -20 °C.

Purification of enzymes. Cells were thawed, re-suspended in buffer A (0.1 M TRIS, 0.5 M NaCl, pH 8.0) and lysed by sonication (Sonopuls 2070, Bandelin, cycle 6, 75% intensity, 2x 2 min) on ice. After centrifugation (16,000 x *g*, 4 °C, 20 min), supernatants were applied to a HisTrap FF crude column connected to an Aekta Explorer system (both GE Healthcare). After washing with 25 mM imidazole, Nt4CL2 was eluted with 500 mM imidazole using buffer B (0.1 M TRIS, 0.5 M NaCl, 0.5 M imidazole, pH 8.0), while CHS was eluted with 125 mM imidazole. Protein-containing fractions were analyzed using Coomassie-stained SDS-PAGE gels and pooled. Afterwards, the proteins were concentrated (Amicon Ultra-4, cut-off 10 kDa, Merck KGaA), filtered (0.22 µm centrifugal filter, Millipore) and applied to an SD200 10/300 Increase size exclusion column connected to an Aekta Explorer system (both GE Healthcare) using buffer C (20 mM TRIS, 150 mM NaCl, pH 8.0). Once again, protein-containing fractions were analyzed using Coomassie-stained SDS-PAGE gels (Figure S4A and S12), pooled, and protein concentration was determined using Bradford assay.

Determination of temperature optimum. Stock solutions of all chemicals were stored at -20 °C prior to usage. To determine the temperature optimum of Nt4CL2, a FlexCycler (Analytik Jena) with gradients from 30 to 60 °C was used. Samples of 100 µL volume containing 500

μM of each substrate (ATP, CoA, *p*-coumaric acid), 2.5 mM MgCl_2 , 0.5 mM TCEP, 20 mM HEPES at pH 7.5, and 0.2 μg enzyme were incubated at different temperatures in the cycler for 20 min, before the reaction was stopped by adding ammonium molybdate (cf. PP_i -assay), and the PP_i detection was performed. Blank samples contained all chemicals except CoA, which was replaced by water.

Standard curve for molybdate-based activity assay. To set up a standard curve, 100 μL of different PP_i concentrations (10-50 μM) in 20 mM HEPES at pH 7.5 and in case 10% DMSO were used and PP_i detection was performed as described in the main text. All samples were measured in triplicates.

***In vitro* biosynthesis of pinocembrin derivatives using Nt4CL2 and AtCHS.** Samples of 300 μL containing 1 mM of ATP, CoA, malonyl-CoA, cinnamic acid derivatives (**20a** and **21a** as 100 mM stock solutions in DMSO), 20 mM HEPES pH 7.5, 0.5 mM TCEP, 2.5 mM MgCl_2 , 30 μg Nt4CL2 and 30 μg CHS enzyme were incubated at 30°C overnight. Afterwards samples were extracted twice with equal volume of ethyl acetate, the organic phase was separated, evaporated and the extracts were dissolved in 200 μL methanol. After filtration the samples were analyzed by LC-HRMS.

LC-HRMS for the *in vitro* experiment using both Nt4CL2 and AtCHS. LC-HRMS was performed on a Q-Exactive Hybrid Quadrupole Orbitrap mass spectrometer using electrospray ionization and an Accela HPLC system (Thermo Fisher) equipped with a Accucore C18 column (2.1 \times 100 mm, 2.6 μm , Phenomenex). HPLC was performed with an injection volume of 10 μL and a gradient elution of solvents A (water, 0.1% formic acid) and B (acetonitrile, 0.1 % formic acid) at a flow rate of 0.3 mL/min: a linear gradient from 5% to 98% B for 10 min, then 98% B for 4 min and 5% B for 6 min.

Kinetic measurements. For kinetic measurements, 100 μL samples containing 500 μM of each substrate (ATP, CoA, *p*-coumaric acid), 2.5 mM MgCl_2 , 0.5 mM TCEP, 20 mM HEPES at pH 7.5, and 0.5 μg enzyme were incubated at 47 °C in a thermo-block for 1 to 4 min. The reactions were terminated by addition of ammonium molybdate (cf. PP_i -assay), and PP_i detection was performed. Blank samples contained all chemicals except CoA, being replaced by water. To determine K_m and K_{cat} values, 100 μL samples were incubated containing 0.5 mM ATP and CoA, different concentrations of *p*-coumaric acid (8-125 μM), 2.5 mM MgCl_2 , 0.5 mM TCEP, 20 mM HEPES at pH 7.5, and 0.2 μg enzyme. Samples were incubated at 47 °C for 1-5 min before the reaction was stopped by adding ammonium molybdate (cf. PP_i -assay) followed by PP_i detection. Afterwards, obtained starting activities were plotted using GraphPad Prism 5.03 and analyzed employing Michaelis-Menten kinetics.

Scale-up of 2,3-(methylenedioxy)pinocembrin production into a stirred tank bioreactor.

For a bioreactor-based flavanone production, a pre-culture of *E. coli* BL21 (DE3) carrying pJK02 was grown in a 300 mL shake flask with a filling volume of 100 mL LB medium overnight. The main culture was performed in two subsequent steps, to realize separate process conditions for biomass formation and production of 2,3-(methylenedioxy)pinocembrin. The first batch culture was performed in a 7 L-stirred tank reactor (Sartorius) with a filling volume of 5 L LB medium that was inoculated with a starting OD₆₀₀ of 0.1. The major aim of the first step was biomass formation. Bacteria were grown to an OD₆₀₀ of 0.6 and induced with 1 mM IPTG for 3 h at 30°C. After centrifugation for 20 min at 7000x *g* and 4°C, the pellet was resuspended and transferred to the second reactor (Sartorius) containing 5 L of M9 medium supplemented with kanamycin, 1 mM IPTG and 3 mM of 2,3-(methylenedioxy)cinnamic acid. The second process step was performed at 30°C for 16 h. The process was stopped after depletion of the main carbon source. Both reactor steps were pH-controlled at 7, by addition of NH₄OH and sulfuric acid if necessary. The aeration was kept constant at 1.25 L/min (=0.25 vvm) as well as the stirring rate at 470 rpm. The dissolved oxygen tension was measured by a Clark-electrode (Mettler-Toledo) as well as the pH value (7 ± 0.05) in both stirred tank reactors. The consumed oxygen was measured by a paramagnetic sensor and the produced CO₂ by infrared measurement (Emerson).

Isolation and purification of 2,3-(methylenedioxy)pinocembrin. In order to achieve a decent amount for structure confirmation by NMR, a 5 L large-scale fermentation was performed as described. After centrifugation for 20 min at 12,000x *g* and 4°C, the supernatant was extracted twice with an equal volume of ethyl acetate. The organic phase was concentrated to dryness under reduced pressure. The pellet was lyophilized and extracted with 800 mL methanol for 1 h. Then, the methanol extract was filtrated and concentrated under reduced pressure. For pre-purification, both dried extracts were dissolved in 50 mL of 50% methanol and loaded onto a Chromabond C18 SPE column (Macherey-Nagel) that was pre-equilibrated with 2 column volumes (CV) of methanol followed by 2 CV water. After washing the column with 2 CV of water and 2 CV of 30% methanol, a step-wise elution was performed using 2 CV of 50%, 2 CV of 60%, 2 CV of 70%, 2 CV of 80%, 2 CV of 90%, 2 CV of 100% methanol. The elutions were analyzed by LC-HRMS for the presence of the desired compound 2,3-methylenedioxy pinocembrin that was detected in the 80% methanol elution. This fraction was then concentrated to dryness, re-dissolved in 2 mL methanol and separated on a semi-preparative HPLC using a Shimadzu LC-20AD instrument, equipped with a Kinetex C18 column (10 × 250 mm, 5 µm, 100 Å, Phenomenex) maintained at 30°C. Fractionation was performed using linear gradient of solvents A (water, 0.1% formic acid) and B (acetonitrile) at a flow rate of 2 mL/min under the following conditions: 10% B for 1 min, a linear increase to 60% B for 14 min, held at 60% B

for 15 min, increase to 97% for 1 min, followed by column rinsing with 97% B for 10 min, decrease to 10% in 2 min and re-equilibration with 10% B for 10 min. Organic compounds were detected by monitoring the absorbance at 254 and 280 nm using a PDA detector (Shimadzu). The compound of interest was eluting at 22.4 min and concentration of the corresponding fraction using the Genevac EZ-2 sample concentrator (SP Scientific) yielded in 0.7 mg of **30c**.

NMR analysis. NMR measurements (^1H , ^1H COSY, HMQC, HMBC) were performed in deuterated methanol (VWR) on a Bruker Avance III 600 MHz spectrometer. The chemical shifts are reported in ppm relative to the solvent residual peak (δ (CD_3OD) = 3.31 (49.05) ppm) (Fig. S9).

Bibliography

1. Hoefgen, S., Lin, J., Fricke, J., Stroe, M. C., Mattern, D. J., Kufs, J. E., Hortschansky, P., Brakhage, A. A., Hoffmeister, D., and Valiante, V. (2018) Facile assembly and fluorescence-based screening method for heterologous expression of biosynthetic pathways in fungi, *Metab. Eng.* 48, 44-51.
2. Schmidt, J. (2016) Negative ion electrospray high-resolution tandem mass spectrometry of polyphenols, *J. Mass. Spectrom.* 51, 33-43.
3. Pluskal, T., Torrens-Spence, M. P., Fallon, T. R., De Abreu, A., Shi, C. H., and Weng, J. K. (2019) The biosynthetic origin of psychoactive kavalactones in kava, *Nat. Plants* 5, 867-878.
4. Meier, K., Klöckner, W., Bonhage, B., Antonov, E., Regestein, L., and Büchs, J. (2016) Correlation for the maximum oxygen transfer capacity in shake flasks for a wide range of operating conditions and for different culture media, *Biochem. Eng. J.* 109, 228-235.
5. Huber, E. M., Scharf, D. H., Hortschansky, P., Groll, M., and Brakhage, A. A. (2012) DNA minor groove sensing and widening by the CCAAT-binding complex, *Structure* 20, 1757-1768.
6. Li, Z., and Nair, S. K. (2015) Structural basis for specificity and flexibility in a plant 4-coumarate:CoA ligase, *Structure* 23, 2032-2042.
7. Burbulis, I. E., and Winkel-Shirley, B. (1999) Interactions among enzymes of the *Arabidopsis* flavonoid biosynthetic pathway, *Proc. Natl. Acad. Sci. USA* 96, 12929-12934.
8. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures, *Protein Expr. Purif.* 41, 207-234.