Supporting Information

Engineering Regioselectivity of a P450 Monooxygenase Enables the Synthesis of Ursodeoxycholic Acid via 7β-Hydroxylation of Lithocholic Acid

Sascha Grobe, Christoffel P. S. Badenhorst, Thomas Bayer, Emil Hamnevik, Shuke Wu, Christoph W. Grathwol, Andreas Link, Sven Koban, Henrike Brundiek, Beatrice Großjohann, and Uwe T. Bornscheuer*

anie_202012675_sm_miscellaneous_information.pdf
Table of Contents

Experimental procedures ........................................................................................................... 3
Materials ........................................................................................................................................ 3
Bacterial strains ............................................................................................................................ 3
Plasmids ......................................................................................................................................... 3
Site-directed mutagenesis ............................................................................................................ 3
Library creation ............................................................................................................................ 3
Expression and purification of recombinant Collinsella aerofaciens 7β-HSDH ................................... 4
UDCA assay using the recombinant Collinsella aerofaciens 7β-HSDH ........................................... 4
Whole-cell biocatalysis .................................................................................................................. 4
HPLC ............................................................................................................................................... 5
Determinaton of the P450 concentration ....................................................................................... 5
Preparative-scale biotransformation .............................................................................................. 5
NMR spectroscopy and MS analysis .............................................................................................. 5
Molecular docking ........................................................................................................................ 5
3DM analysis ................................................................................................................................ 6
Additional results .......................................................................................................................... 7
Calibration curves ......................................................................................................................... 13
Sequences ...................................................................................................................................... 15
Chemoenzymatic synthesis of UDCA .......................................................................................... 16
7β HSDH assay ............................................................................................................................ 17
NMR spectra .................................................................................................................................. 18
References ...................................................................................................................................... 20
Author contributions ...................................................................................................................... 20
Experimental procedures

Materials

All chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA) and used without further purification. Oligonucleotides were purchased from Thermo Fisher Scientific (Hennigsdorf, Germany).

Bacterial strains

*Escherichia coli* C43(DE3), purchased from Lucigen (Middleton, WI, USA), was used for expression and whole-cell biotransformations. For plasmid amplification after site-directed mutagenesis, *E. coli* TOP10 cells purchased from Thermo Fisher Scientific (Hennigsdorf, Germany) or NEB5α cells from New England Biolabs (Ipswich, MA, USA) were used.

Plasmids

The plasmid for expression of the redox partners PdR and PdX was a gift from Prof. Anett Schallmey (Technical University Braunschweig, Braunschweig, Germany). A codon-optimized synthetic gene encoding CYP107D1 (UniProtKB entry: Q59819) was purchased from GenScript and subcloned into pET-28a using the NcoI and XhoI restriction sites. A three-domain low-diversity combinatorial library of OleP (randomized at F84, S240, and V291) was purchased from Twist Bioscience (San Francisco, CA, USA) and cloned into pET-28a using the restriction sites NcoI and XhoI.

Site-directed mutagenesis

Site-directed mutagenesis was performed using the QuickChange™ method. PCRs contained 2.5 µL 10x Pfu* buffer, 1.25 µL (5%) DMSO, 0.25 µM forward and reverse primer (each 1.25 µL), dNTPs (0.25 mM each), 50 ng of template (pET-28a-oleP), 0.4 µL Pfu* polymerase and 17.35 µL ddH2O. After initial denaturation for 30 s at 95 °C, 19 cycles of denaturation at 95 °C for 30 s, annealing at 63 °C (all primers were designed to anneal at this temperature) for 30 s, and extension at 72 °C for 7 min were followed by a final elongation step at 72 °C for 10 min. After DpnI digest, the PCR mixtures (2 µL) were used for heat shock transformation of chemocompetent *E. coli* TOP10 or NEB5α cells, which were plated on Luria Bertani broth (LB, Miller) agar containing 25 µg/mL kanamycin and incubated at 37 °C overnight. Three colonies for each mutagenesis reaction were picked, cultured in LB supplemented with 25 µg/mL kanamycin at 37 °C and 220 rpm for 10 h, followed by plasmid isolation (innuPREP Plasmid Mini Kit 2.0, Analytik Jena, Jena, Germany) and Sanger sequencing (Eurofins Genomics Germany GmbH, Ebersberg, Germany). Confirmed mutants were co-transformed with pACYCDuet-1-pdR/pdX into *E. coli* C43(DE3) for further studies and biotransformations.

Library creation

The synthetic *oleP* library was ordered as codon-optimized gene strings from Twist Bioscience (San Francisco, CA, US). Residue F84 was fully saturated while S240 (A, D, E, F, G, H, I, L, M, N, Q, S, T, and V) and V291 (A, D, E, F, G, I, L, M, N, P, Q, R, S, T, V, and W) were varied to residues naturally occurring at these positions (see 3DM analysis in the Supporting Figures 1-4). The flanking sequence prior to the start codon was CTGTTAGTGTTATCAGTAGAGGATATACC. After the stop codon, the sequence TTGACTCGAGACCACAGTGCATCTGGAGCCTAGACCTTTGGCTTATCAGTAGAAGGAGATATACC was added. The flanking regions included restriction sites for NcoI and XhoI. The DNA sequence is given in the Additional results section. The library was cloned into pET-28a using these restriction sites. The vector was digested in a reaction containing 1x CutSmart® Buffer (New England Biolabs, Ipswich, MA, USA), 579 ng pET-28a (concentration of 57.9 ng/µL), 1 µL Xhol (20000 U/mL), 1 µL Ncol-HF (10000 U/mL), 23 µL ddH2O in a final volume of 25 µL. For the digestion of the insert, 5 µL CutSmart® buffer were used, 1 µL Xhol (20000 U/mL), 1 µL Ncol-HF (10000 U/mL), 7.5 µL oleP library (150 ng), and 35.5 µL ddH2O. The digestion reactions were incubated for 16 h at 37 °C, followed by heat inactivation of restriction enzymes at 80 °C for 20 min. The digested vector was separated on a 1% (w/v) agarose gel and isolated using a gel purification kit (NucleoSpin™ Gel and PCR Clean-up Kit, Macherey-Nagel, Düren, Germany). The library was isolated after the double digestion by PCR clean up (NucleoSpin™ Gel and PCR Clean-up Kit, Macherey-Nagel, Düren, Germany).

The ligation of vector and insert was done in a 30 µL reaction containing 10 µL of digested pET-28a vector (100 ng) and 8.8 µL insert (75.38 ng), 3 µL of 10x ligase buffer, 1.5 µL T4-ligase, and 6.7 µL ddH2O. The amounts of DNA required for a 1:3 vector to insert ratio was calculated using NEBioCalculator. The ligase was heat-inactivated at 70 °C for 15 min and dialyzed against ddH2O. The pET-28a-oleP library was co-transformed with pACYCDuet-1-pdR/pdX by electroporation into *E. coli* C43(DE3) using a 1:8 molar ratio[1]. A QPix 420 Bench-top Colony Picker (Molecular Devices, LLC. San Jose, CA, US) was used to transfer 10,560 clones into 96-well microtiter
Expression and purification of recombinant *Collinsella aerofaciens* 7β-HSDH

We initially considered the 7β-hydroxysteroid dehydrogenases (7β-HSDH) from *Clostridium absonum* and *Collinsella aerofaciens*. The *C. absonum* enzyme[2] has a high specific activity (89 U/mg) and about 24 mg can be expressed per liter of culture. The *C. aerofaciens*[3] enzyme has a lower specific activity (17 U/mg) but about 300 mg can be expressed per liter of culture. Most importantly, the enzyme can be stored at -20 °C for several months, not losing activity after more than 100 freeze-thaw cycles[2,3,4].

A synthetic gene encoding the *C. aerofaciens* 7β-HSDH (GenBank accession number: ZP_01773061.1) was ordered from BioCat GmbH (Heidelberg, Germany) and subcloned into pET-28a(+). DNA and protein sequences are provided below in the section Sequences. *E. coli* BL21(DE3) was transformed using approximately 50 ng of the expression vector and plated on LB agar containing 50 µg/mL kanamycin and 1% (w/v) glucose. After overnight incubation at 37 °C, single colonies were used to inoculate three 2 L flasks, each containing 200 mL of autoinduction medium (terrific broth supplemented with 0.05% glucose, 0.2% lactose, and 100 µg/mL kanamycin). Cultures were incubated at 30 °C and 200 rpm for 24 h. The three 200 mL cultures were pooled, and the cells harvested by centrifugation at 4500 g and 4 °C for 30 min. Cells were resuspended in 20 mL LEW buffer (300 mM sodium chloride, 50 mM sodium phosphate, pH 8.0) and lysed by a single pass through a French press. Crude lysate was clarified by centrifugation at 10000 g and 4 °C for 1 h. The clarified lysate was applied to a 2 mL column of Protino Ni-3,4-TED resin (Macherey-Nagel, Düren, Germany) preequilibrated with LEW buffer. After washing the column with 20 mL of LEW buffer, 10 mL of elution buffer (LEW containing 250 mM imidazole, pH 8.0) was applied to the column and the eluate was collected in ~1 mL fractions. Fractions containing the most protein were identified by measuring absorbance at 280 nm and pooled. An equal volume of 80% (v/v) glycerol was added before storage at -20 °C.

UDCA assay using the recombinant *Collinsella aerofaciens* 7β-HSDH

The UDCA produced by OleP variants was quantified using a spectrophotometric assay. The NADP+‐dependent 7β-hydroxysteroid dehydrogenase (7β-HSDH) from *C. aerofaciens* was used to convert UDCA and NADPH, respectively. NADPH formation was coupled to the reduction of the tetrazolium salt WST-1 using 1-methoxyphenazine methosulfate (mPMS). Reduction of WST-1 results in formation of a water-soluble yellow formazan dye (ε_{450nm} = 34,860 M⁻¹ cm⁻¹).

Samples (200 µL) of biotransformations using the OleP variants were transferred to 96-well microtiter plates. Because this crude supernatant gives significant background absorbance at 450 nm, all assay components except the 7β-HSDH were added and the 7β-HSDH-independent absorbance was measured. The 7β-HSDH was then added and a second absorbance measurement was performed. To each 200 µL sample, 40 µL of a 5x master mix was added, giving final concentrations of 40 mM Tris-HCl (pH 7.5), 100 µM NADP+, 10 µM mPMS, and 100 µM WST-1. Plates were gently shaken for 5–10 min before absorbance at 450 nm was measured using a Varioskan™ LUX plate reader (Thermo Fisher Scientific, Hennigsdorf, Germany). Next, 10 µL (25 µg) of 7β-HSDH (2.5 mg/mL in 40% v/v glycerol) was added. Plates were gently shaken for about 10 min before absorbance at 450 nm was measured again. A 7β-HSDH-dependent increase in absorbance between the two measurements reflected the presence of UDCA in the sample. To allow UDCA quantification, a standard curve was set up. A series of UDCA concentrations (0, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 µM) was prepared in control biotransformation supernatant harboring only the reductase system PdR/PdX on a pACYCDuet-1 vector, closely mimicking the screening conditions. The 40 µL assay master mix was added and plate was shaken for about 5 min before the first absorbance measurement was taken. The 7β-HSDH was then added and the plate gently shaken for about 10 min before taking the second absorbance measurement. The difference in absorbance (450 nm) was calculated and plotted against UDCA concentration. A linear relationship (R² = 0.99) between absorbance at 450 nm and UDCA concentration was observed, with a detection limit of 6 µM (Supplementary Figure 10). This method allowed accurate quantification of UDCA formation. The 7β-HSDH from *Collinsella aerofaciens*, like that from *Clostridium absonum*, is highly selective for oxidation of 7β-OH groups and has no measurable activity towards 3α-, 7α-, or 12α-OH groups[2,5]. Therefore, it can be used for the selective quantification of ursodeoxycholic acid in samples like human and bear bile, which contain complex mixtures of bile acids[6].

Whole-cell biocatalysis

The whole-cell biocatalysis was performed as described before[5]. Starter cultures were prepared by inoculating 96-well microtiter plates containing 240 µL LB (Miller) medium per well (supplemented with 25 µg/mL kanamycin and 25 µg/mL chloramphenicol) using the library glycerol stocks. Plates were sealed using air permeable membranes and grown for 12–16 h at 37 °C and 250 rpm (Infors Multitron Standard, Bottmingen, Switzerland). Prewarmed TB medium (2 mL per well of a 24-deepwell pate, supplemented with 25 µg/mL kanamycin and 25 µg/mL chloramphenicol) was inoculated with 1% (v/v) of the starter culture and grown at 37 °C for 4–5 h until an OD₆₅₀ of 1 was reached. The plates were sealed with air permeable membranes. The cultures were cooled to 28 °C and supplemented with 0.64 mM 5-aminolevulinic acid and 0.3 mM FeSO₄, followed by induction of protein expression using 0.4 mM IPTG.

SUPPORTING INFORMATION

plates containing 150 µL LB medium per well (supplemented with 25 µg/mL kanamycin and 25 µg/mL chloramphenicol). Cultures were grown at 37 °C and 250 rpm for 16 h and glycerol stocks were prepared by adding 150 µL of 60% (v/v) glycerol to each well, resulting in a final concentration of glycerol of 30% (v/v).
Determination of P450 concentrations

The concentration of P450-monoxygenase was determined as described by Omura and Sato(7). To a 2 mL lysate sample, 2 µM Safranin T (in 50 mM phosphate buffer pH 7.5) was added, followed by the addition of a spatula tip of sodium dithionite(8). The tube was inverted and split into two 1 mL samples. Carbon monoxide (CO) was bubbled through one sample at a rate of one bubble per second for one minute. A CO-differential spectrum (Jasco V-550, Pfungstadt, Germany) was recorded between 400 and 500 nm and the concentration of the P450 monoxygenase was calculated based on the given extinction at 450 nm.

Preparative-scale biotransformation

Whole-cell biocatalysis was performed with the P450 monoxygenase CYP107D1 F84Q/S240A/V291G as described before, but on 500 mL reaction scale. After 24 h, the reaction was stopped, and cells and supernatant were separated by centrifugation at 4000 g for 20 min. The supernatant was extracted twice with ethyl acetate, the organic phases combined, and dried with anhydrous sodium sulfate. The solvent was evaporated, the residue dissolved in 10 mL ethanol and applied to preparative and analytical HPLC measurements. Preparative and analytical HPLC were performed using Shimadzu devices CBM-20A, LC-20A P, SIL-20A, and FRC-10A with an ELSD-LTII detector. A semi-preparative post-column adjustable-flow splitter (Analytical Scientific Instruments) was utilized for evaporative light scattering detection in preparative mode. A LiChroCART® (250×4 mm, Merck) and a Hibar® RT (250×25 mm, Merck) column, both containing LiChrospher® 100 RP-18e (5 µm), were used in analytical and preparative mode, respectively. The UDCA product was separated from impurities and remaining substrate by preparative HPLC. The fractions containing UDCA were pooled, the solvents evaporated, and the residue dried in vacuo. The dried residue was dissolved in 500 µL deuterated methanol and analyzed by NMR-spectroscopy.

NMR spectroscopy and MS analysis

NMR spectra were recorded on a Bruker Avance III instrument (1H NMR: 400 MHz, 13C NMR: 100.6 MHz). Chemical shifts were referenced to tetramethylsilane (TMS) as internal standard in deuterated methanol and reported in parts per million (ppm). Proton signals are described using the abbreviations: br = broad, s = singlet, m = multiplet, and combinations thereof. To assign proton and carbon atoms, 2D NMR techniques (H,H-COSY, HSCP, HMBC) were used. High accuracy mass spectra were recorded on a Shimadzu LCMS-IT-TOF using ESI. Purity of final compounds was determined by HPLC using a DAD detector at 220 and 254 nm.

Molecular docking

Molecular docking experiments were performed with YASARA (Vienna, Austria). The structures of OleP in complex with clotrimazole (PDB ID: 4XE3) and OleP in complex with 6-deoxyerythronolide B (PDB ID: 5MNS) were used as template structures. Visualization and structural analysis was performed using UCSF Chimera®. Prior to the docking experiments, water and ligands were removed from the PDB files and the simulation cell moved to the heme cofactor. The compounds to dock were LCA (PDB-ligand: 4OA) and testosterone. Afterwards, the dock_run and dockrunensemble macros from YASARA were used with its standard settings for all docking
experiments. The resulting clusters were viewed with Chimera and compared to proper positioning of the B-ring of the steroid in 5 Å distance towards the heme center and a horizontal orientation in binding towards the heme cofactor. The results for LCA were compared with the docking experiments for testosterone and residues for the alanine scanning picked based on a zone of 5–11 Å around the heme iron.

**3DM analysis**

The distribution of the amino acids at the positions F84, V93, L94, S240, and V291 was determined using 3DM (Bio-Prodict, Nijmegen, The Netherlands) based on a P450 monooxygenase superfamily database. For comparison with the overall dataset, two subsets originating from the P450 monooxygenase superfamily were built and used. One of these corresponds to OleP (PDB ID: 4XE3A) and the other to the human P450 monooxygenase CYP3A4. The resulting distributions at each position are displayed in the Supplementary Figures S1–S4. Based on this analysis, the amino acids substitutions for the library were chosen for site-directed mutagenesis.
Additional results

Alanine scanning
The residues picked for the alanine scanning were: F84, E89, V93, L94, L179, S240, I243, H246, E247, T248, S249, L290, V291, S292, S295, F296, F321, M395, and I397. The residues G82, A95, A244, and G294 were also within the 14 Å zone around the heme but were not mutated to alanine.

Table S1. Oligonucleotides used for site-directed mutagenesis.

| Oligonucleotide sequence | Position of the exchanged amino acid |
|--------------------------|-------------------------------------|
| CAACCCCGCTATGGCGCCGACCCCCC | F84A_Fwd |
| CCGCCGGGCTGGCCGCCCATACCGGGGTTG | F84A_Rev |
| CGACCCCGCCCAGATGGTTC | E89A_Fwd |
| GAACACATCCGGGCGGGCTCG | E89A_Rev |
| GAACCGATGGTGGCTGCCCAGGATCG | V93A_Fwd |
| GATCTCTGGCCAGCGACCATCCGGTTC | V93A_Rev |
| CCGGATGGTGGTGGCCCAGGATCC | L94A_Fwd |
| GATCTCTGGCCAGCACCACCATCCGG | L94A_Rev |
| CAGCGATCGATGGCGAGCTTACCCG | L179A_Fwd |
| CGGTTAGAGCTGCATCGCATCGT | L179A_Rev |
| GTGAATATGGGTTGCTGCTGATCGCCGG | S240A_Fwd |
| CCGTCGACACGCGCAACCCCATATTCAC | S240A_Rev |
| GTGGTTAGCTGCTGCCCAGTATGAAAC | I243A_Fwd |
| GTTTACATGACCGCCGCAACGCTAACCAC | I243A_Rev |
| GCCGGTCAATGACGTCGTAACC | T248A_Fwd |
| GTTCACGGACCGCTTCATGACC | T248A_Rev |
| CGCTACACGCCGGGCCATTTCAGCTG | L290A_Fwd |
| CCGCTGAACCCGCCGGATGCTAGC | L290A_Rev |
| CACGCCGCTGGGGATCGCGCTGC | V291A_Fwd |
| CGACGGCCTGACACCGCGGCTG | V291A_Rev |
| GATGCGATGCGATGGTTGCTG | S295A_Fwd |
| CAGAAACAAGCGCCCGCTGAAC | S295A_Rev |
| CAGCGGCTGGCGGTTCCGTCGCC | F296A_Fwd |
| GGCGCACACGCGCGACCCGCTG | F296A_Rev |
| GTGCGTGTTCGCGCCGCTGGATGCAAACC | F321A_Fwd |
| GTGATCCGATGGCGGCGGACCATCCGG | F321A_Rev |
| GAAACGGTATGGCGGATTCGGCTG | L396A_Fwd |
| CCCAGCCGGAATCCGATCCGGTTTTC | L396A_Rev |
| GAAACGGGATATGGCGGCACCGCTGGAACTG | I397A_Fwd |
| GACGTTCACAGACCGCGCCGACAACACCGTTTTC | I397A_Rev |
| GACCCGGCCTAGCAGCAAAACCCCATATTC | S240V_Rev |
| GATATGGGTTATTCGCTGATCGCCCGG | S240I_FWD |
| CCGCGATCAGCAGAATAACACCCATATTC | S240I_Rev |
F84 was not part of the 3DM core alignment and could, therefore, not be analyzed. CYP3A4 as prominent member of P450s with a broad substrate scope was used as second reference with its derived subset besides the overall dataset.

Residue V93

![Graph of V93 amino acid distribution](image1)

**Figure S1.** 3DM analysis of residue V93 which is assigned as core position 56 in the 3DM database for P450 monooxygenases showing detailed occurrence of each of the amino acids at this position. Three different subsets are shown: the full dataset to the left covering 77,699 sequences of which 13,306 have gaps followed by the subset used for comparisons based on CYP3A4 harboring 4,774 sequences of which 1,900 have gaps and to the right the subset of OleP with 306 sequences of which 45 have gaps.

Residue L94

![Graph of L94 amino acid distribution](image2)

**Figure S2.** 3DM analysis of residue L94 which is assigned as core position 57 in the 3DM database for P450 monooxygenases showing detailed occurrence of each of the amino acids at this position. Three different subsets are shown: the full dataset to the left covering 77,776 sequences of which 13,321 have gaps followed by the subset used for comparisons based on CYP3A4 harboring 4,775 sequences of which 1,899 have gaps and to the right the subset of OleP with 306 sequences of which 45 have gaps.
Residue S240

Figure S3. 3DM analysis of residue S240 which is assigned as core position 211 in the 3DM database for P450 monoxygenases with detailed occurrence of each of the amino acids at this position. Three different subsets are shown: the full dataset to the left covering 86,622 sequences of which 4,385 have gaps followed by subset used for comparisons based on CYP3A4 harboring 6,452 sequences of which 222 have gaps and to the right the subset of OleP with 332 sequences of which 19 have gaps.

Residue V291

Figure S4. 3DM analysis of residue V291 which is assigned as core position 281 in the 3DM database for P450 monoxygenases with detailed occurrence of each of the amino acids at this position. Three different subsets are shown: the full dataset to the left covering 89,903 sequences of which 1,104 have gaps followed by subset used for comparisons based on CYP3A4 harboring 6,595 sequences of which 79 have gaps and to the right the subset of OleP with 351 sequences of which 0 have gaps.
**Library design**

Table S2. Showing the library design used for synthetic library

| Variant | 1 | 2 | 3 |
|---------|---|---|---|
| Amino acid (position)- ORF | 104 | 260 | 311 |
| Amino acid (F) | (S) | (V) |
| Base pair (start)-ORF | 310 | 778 | 931 |
| Codon | TTC | AGT | GTT |

DNA sequence of the OleP library (blue) with attached linker regions (black), mutation sites (red)

```
AGAAGGAGATATACC atgggcagcagccatcatcatcatcatcatcatcacagcagcagccatcatcatcatcatcagcagcagccatcatcacagcagcagccatcatcatcatcatcatcagcagcagcagccatcatcatcatcatcatcacagcagcagcagccatcatcatcatcatcacagcagcagcagccatcatcatcatcatcatcacagcagcagcagccatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat
Results from prescreening the library

**Figure S5.** Screening results using the colorimetric UDCA assay. OleP variants showing up to 18-fold increased absorbance compared to the positive control OleP S240A were identified. S240A was the 35th best variant found in the screening by the colorimetric assay and is indicated with a red bar. The level of relative absorbance of S240A is indicated with a red line.

**Sequencing results of screening hits after HPLC confirmation**

**Figure S6.** Amino acid distribution at residues F84, S240, and V291 of variants confirmed to produce UDCA by HPLC.
SUPPORTING INFORMATION

UDCA formation of selected variants

Table S3. UDCA formation after 24 h whole-cell reactions of selected single, double, and triple mutants of OleP.

| Mutant      | UDCA formation [µM] | Variant         | UDCA formation [µM] |
|-------------|---------------------|-----------------|---------------------|
| S240A       | 8.7 ± 3.1           | F84Q V291A      | 8.5 ± 3.4           |
| F84A        | 0.1                 | F84Q V291G      | 4.2                 |
| F84Q        | 0.1                 | F84Q V291D      | 9.5                 |
| F84C        | 0.1                 | F84C S240A      | 0.1                 |
| F84M        | 3.7                 | F84C V291A      | 7.7                 |
| V291A       | 0.7                 | F84C V291G      | 0.1                 |
| V291G       | 5.4                 | F84C V291D      | 2.4                 |
| V291D       | 3.7                 | F84M S240A      | 0.1                 |
| F84A S240A  | 5.7                 | F84M V291A      | 0.1                 |
| F84A V291A  | 65.2 ± 3.2          | F84M V291G      | 1.2                 |
| F84A V291G  | 0.1                 | F84M V291D      | 2.9                 |
| F84A V291D  | 14 ± 2              | S240A V291A     | 18.2 ± 3.8          |
| F84Q S240A  | 2.4                 | S240A V291D     | 2.8                 |
| S240A V291G | 6.4                 | F84A S240A V291G| 4.2                 |
| F84A S240A V291A | 58.5 ± 20.2 | F84A S240A V291D | 3.2             |
| F84Q S240A V291A | 9.7 ± 3.1 | F84Q S240A V291G | 67.1 ± 5.1         |
| F84Q S240A V291D | 10.4 ± 3.1 | F84C S240A V291A | 7.2 ± 2.8          |
| F84C S240A V291G | 9.2 ± 2.6 | F84C S240A V291D | 5.6 ± 1.8          |
| F84M S240A V291A | 10.2 ± 2.6 | -                | -                   |

Calibration curves

Calibration curve for lithocholic acid (LCA)

Figure S7. Calibration curve for HPLC-RI of the substrate LCA.
Calibration curve for murideoxycholic acid (MDCA)

Figure S8. Calibration curve for HPLC-RI of the 6β-hydroxylation product MDCA.

Calibration curve for ursodeoxycholic acid (UDCA)

Figure S9. Calibration curve for HPLC-RI of the 7β-hydroxylation product UDCA.
Calibration curve for the 7β-HSDH assay

Figure S10. Calibration curve for the 7β-HSDH assay.

**Sequences**

**DNA sequence of the 7β HSDH from *Collinsella aerofaciens***

ATGGGCACGCAGCCATCATCATCATCATTAGACGAGCAGCGCTGTGCGCGCGCGCCAGCCATTAGAAGGAGATACG
TGAGTGGGCTGCTCGGAGCCGACCGGAGGCGCTCGGAGGATGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTCGCGCGACTTATAGCCACCGCGCCCGTGCCGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
GTGGAGGCGCCGAGGCGCTCCGCACGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
AGCTGCGCTCGCAGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC
ACGGACTCCGCCGCGACCGGAGGCGCTCGGAGGACCTACCCGTGGAGACCAAGGTC

**Amino acid sequence of the 7β HSDH from *Collinsella aerofaciens***

MGSSHSHHHHHSSLVPGRSHMNLEKRYGWGLILGATEGVSOGKAFCEKIAAGGMNVVMGRREEKLNVLAGREIREETYGVETKVRA
DFPSQPAGAAETVFAATEGLMDGMFSYVACLHSFGKIQDTPWKEHAIAMINNVNTFLKCFHHYMRFAAQRDGAIVNVSMTGIISSSP
WNGQYGAQKAFILKMETAVACCEGETGVDEVTITLTGLTTPSLLSNLPGGGPQGEAVMKIAALTPEECVDEAFEKLGKELSVIAGQRN
DSVHDWKANHTEDEYIRYMGSFYRD
Chemoenzymatic synthesis of UDCA

Scheme S1. Chemical and chemoenzymatic synthesis routes towards UDCA starting from CA or CDCA.
Scheme S2. Reaction scheme for the detection of UDCA formation by monitoring the NADP⁺ dependent conversion of UDCA to 7-oxo LCA using a coupled assay. The electron coupling reagent 1-methoxy-5-methylphenazinium methyl sulfate (mPMS) reduces 2-(4-iodophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2H-tetrazolium (WST-1) to form the yellow formazan dye\textsuperscript{111}. 

\[ \text{OleP F64Q/S240V/V291G} \text{ } \begin{array}{c} \text{NADH} \\ \text{O}_2 \end{array} \Rightarrow \begin{array}{c} \text{UDCA} \\ \text{NAD}^+ \end{array} \Rightarrow \begin{array}{c} \text{7H-SDH} \\ \text{H}_2\text{O} \end{array} \Rightarrow \begin{array}{c} \text{7-oxo LCA} \\ \text{NADP}^+ \end{array} \Rightarrow \begin{array}{c} \text{NADPH} \\ \text{NAPOH} \end{array} \Rightarrow \begin{array}{c} \text{oxidised mPMS} \\ \text{reduced mPMS} \end{array} \Rightarrow \begin{array}{c} \text{WST-1} \\ \text{WST-1 formazan} \end{array} \]
NMR analysis of the isolated fraction UDCA

Figure S11. $^1$H-NMR spectroscopy of the isolated product confirmed as UDCA.

Figure S12. $^{13}$C-NMR spectroscopy of the isolated product confirmed as UDCA.
The signal for the carboxylic acid function at 178 ppm in the $^{13}$C-NMR spectrum is not clearly observed, which is expected when using deuterated methanol as solvent. Peak broadening, even diminishing, of carboxylic acid functional groups can be attributed to the Nuclear Overhauser effect in NMR-spectroscopy.

The $^1$H-NMR spectrum together with the $^{13}$C-NMR spectrum supported by the 2D spectra and HRMS clearly identified the isolated product as UDCA.

$^1$H NMR, H,H-COSY (400 MHz, CD$_3$OD): $\delta$ (ppm) = 3.55–3.46 (m, 2H), 2.41–2.17 (m, br, 2H), 2.07–2.04 (m, 2H), 1.96–1.75 (m, 5H), 1.66–0.98 (m, 19H), 0.98 (s, 3H), 0.97 (s, 3H), 0.73 (s, 3H); $^{13}$C NMR, DEPT135, HSQC, HMBC (101 MHz, CD$_3$OD): $\delta$ (ppm) = 72.1 (7), 72.0 (3), 57.5 (17), 56.6 (14), 44.8 (13), 44.5 (8), 44.1 (5), 41.6 (12), 40.7 (9), 38.6 (4), 38.0 (6), 36.7 (20), 36.1 (1), 35.2 (10), 32.6 – 32.5 (22+23), 31.0 (2), 29.6 (16), 27.9 (15), 23.9 (19), 22.4 (11), 18.9 (21), 12.6 (18); ESI-HRMS: calcd. for [C$_{24}$H$_{40}$O$_4$-H]$^+$ 391.2854, found 391.2870

NMR analysis MDCA produced in trace amounts

Figure S13. $^1$H-NMR spectroscopy of the isolated MDCA.
Figure S1. 13C-NMR spectroscopy of the isolated MDCA.

1H NMR, H,H-COSY (400 MHz, CD3OD): δ (ppm) = 3.68 (s, br, 1H, 6-H), 3.55–3.46 (m, 1H, 3-H), 2.40–2.14 (m, 2H), 2.04–2.00 (m, 1H), 1.95–1.73 (m, 4H), 1.66–0.92 (m, 28H), 0.72 (s, 3H, 18-H). 13C NMR, DEPT135, HSQC, HMBC (101 MHz, CD3OD): δ (ppm) = 74.0 (6), 72.1 (3), 57.8, 57.5, 44.0, 42.1, 37.3, 37.1, 36.8, 35.5, 35.4, 32.2, 29.3, 26.1 (19), 25.3, 21.8, 18.8, 12.5 (18).

References

[1] M. Goldsmith, C. Kiss, A.R.M. Bradbury, D. S. Tawfik, Protein Eng. Des. Sel. 2007, 20, 315-318.
[2] E. E. Ferrandi, G. M. Bertolesi, F. Polentini, A. Negri, S. Riva, D. Monti, Appl. Microbiol. Biotechnol. 2012, 96, 1221-1233.
[3] L. Liu, A. Aigner, R. D. Schmid, Appl. Microbiol. Biotechnol. 2011, 90, 127-135.
[4] I. A. MacDonald, C. N. Williams, J. D. Sutherland, A. C. MacDonald, Anal. Biochem. 1983, 135, 349-354.
[5] S. Hirano, N. Masuda, Appl. Environ. Microbiol. 1982, 43, 1057-1063.
[6] S. Grobe, A. Wszołek, H. Brundiek, M. Fekete, U. T. Bornscheuer, Biotechnol. Lett. 2020, 42, 819-824.
[7] T. Omura, R. Sato, J. Biol. Chem. 1964, 239, 2370-2378.
[8] F. P. Guengerich, M. V. Martin, C. D. Soh, Q. Cheng, Nature protocols 2009, 4, 1245-1251.
[9] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, J. Comput. Chem. 2004, 25, 1605-1612.
[10] F. Gielen, R. Hours, S. Emond, M. Fischlechner, U. Schell, F. Hollfelder, P. Nat. Acad. Sci. USA 2016, 113, E7383-E7389.

Author contributions

U.T.B initiated the study and directed the project. S.G., T.B., and E.H. designed the mutagenesis strategy. S.G., and C.P.S.B. developed and established the screening assay. S.G. performed screening and analysis. S.G. and C.W.G. supported by A.L. isolated and analyzed the products via NMR and HPLC-MS. S.G. wrote the manuscript with input from U.T.B., C.P.S.B., T.B., E.H., S.W., C.W.G., A.L., S.K., H.B., and B.G. All authors read and approved the final manuscript.