Supplemental Information for "Overcoming heterologous protein interdependency to optimize P450-mediated Taxol precursor synthesis in *Escherichia coli*"

Bradley Walters Biggs[1,2,†], Chin Giaw Lim[1,†], , Kristen Sagliani[1], Smriti Shankar[1], Gregory Stephanopoulos[3], Marjan De Mey[1,3,4] & Parayil Kumaran Ajikumar[1]

Author affiliations:
[1]Manus Biosynthesis, 1030 Massachusetts Avenue, Suite 300, Cambridge, MA 02138, USA
[2]Department of Chemical & Biological Engineering (Masters in Biotechnology Program), Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA
[3]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[4]Laboratory for Industrial Biotechnology and Biocatalysis, Ghent University, Coupure Links 653, B-9000, Belgium

[†]Denotes equal authorship

Corresponding Authors:
Ajikumar Parayil
Manus Biosynthesis
1030 Massachusetts Avenue, Suite 300,
Cambridge, MA, 02138, USA
Tel: +1-617-253-4583
Fax: +1-617-253-3122
Email: pkaji@manusbio.com

Marjan De Mey
Centre for Industrial Biotechnology and Biocatalysis,
Ghent University, Coupure Links 653, B-9000, Belgium
Tel: +32 9 264 60 28
Email: Marjan.DeMey@UGent.be

Table of Contents

Additional methods and materials

Strains, plasmids, oligonucleotide, genes, and chemical and physical materials. *Escherichia coli* K12 MG1655 △*recA*△*endA* DE3 was used as the parent for all strain engineering experiments and was carried from previous work.[1] Pathway construction and DNA manipulations were performed following standard molecular cloning protocols. A full list of the strains and plasmids used in this study can be found in Supplemental Table 1. All oligonucleotides used in this study are listed in Supplemental Table 2. All chemicals were purchased from Sigma (St. Louis, MO, USA) unless otherwise stated. Yeast extract was purchased from Bio Springer (Milwaukee, WI, USA). Magnesium sulfate, ammonium phosphate (dibasic), and citric acid were purchased from Fisher Scientific (Pittsburg, PA, USA). HEPES, calcium chloride, and thiamine hydrochloride were purchased from Alfa Aesar (Ward Hill, MA, USA). GC-MS grade n-Hexanes and methanol were purchased from EMD Millipore (Billerica, MA, USA). Isopropyl β-D-1-thiogalactopyranoside (IPTG) was purchased from Zymo Research (Irvine, CA, USA). Glycerol and Tris (4G, 1M, pH 8.0) were purchased from BDH Solvents (Dawsonville, GA, USA). Qiagen (Valencia, CA, USA) bacterial kits were used for all DNA and RNA preparations. All primers were purchased from Integrated DNA Technologies (Coralville, IA, USA). All sequencing services were conducted by Genewiz (Cambridge, MA, USA). Agarose, gel materials, and Cybersafe were purchased from Life Technologies (Carlsbad, CA, USA). Luria Broth (LB), Terrific Broth (TB), and LB agar were purchased from BD Biosciences (San Jose, CA, USA). Stock concentrations for antibiotics were spectinomycin (50 mg/L), chloramphenicol (34 mg/L), and kanamycin (50 mg/L). Antibiotic stocks were diluted 1000x for cell culture experiments. All incubation was carried out in Eppendorf New Brunswick Anova Incubators (Enfield, CT, USA).

The sequences of geranylgeranyl pyrophosphate synthase, Taxadiene synthase, CYP725A4 cytochrome P450 Taxadiene 5α-hydroxylase and Taxus NADPH:cytochrome P450 reductase (CPR) were obtained from *Taxus canadensis*, *Taxus brevifolia*, and *Taxus cuspidata* (Genbank accession codes: AF081514, U48796, AY289209 and AY571340). For functional expression in *E. coli* all sequences were codon optimized and the N-terminal plastid transit peptides were removed as previously described.[1]

Construction of the TG pathway plasmids with FRT-Km-FRT cassette for chromosomal integration. To construct the FRT-Km-FRT cassette plasmids used for amplifying the sequences for chromosomal pathway localization, p10T7-TG[1] was digested with XhoI/SalI. The FRT-Km-FRT cassette was amplified from the pKD4 plasmid using the primers Fw-KmFRT-XhoI and Rv-KmFRT-SalI. The amplified DNA was digested with XhoI/SalI and cloned into the XhoI/SalI digested p10T7-TG plasmid yielding p10T7-TG-KmFRT.

Construction of the P450 module plasmids with FRT-Cm-FRT cassette for chromosomal integration. Vectors pMB1-CmFRT-Trc-GFP and pMB1-CmFRT-T7-GFP were created for easy integration in loci *malT*, *lacY* and *araA* of *E. coli* MG1655. These vectors were constructed by amplification of the Trc-promoter and T7-promoter from p10Trc and p10T7 using primers Fw-Trc-SacI & Rv-Trc-NcoI and Fw-T7-SacI & Rv-T7-NcoI, respectively. Following, they were digestion with SacI/NcoI and subsequently cloned in SacI/NcoI site of pMB1-CmFRT-GFP (an in-house developed plasmid).

Next, the Trc-(8RP)P450-*l*-(MA)CPR module and T7-(8RP)P450-*l*-(MA)CPR module were amplified from p5Trc-(8RP)P450-*l*-(MA)CPR and p5T7-(8RP)P450-*l*-(MA)CPR using Fw-Trc-SphI & Rv-CPR-PvuI and Fw-T7-SphI & Rv-CPR-PvuI, respectively. Then, the PCR fragments were cut with SphI/PvuI and cloned in the corresponding sites of pMB1-CmFRT-TrcGFP and pMB1-CmFRT-T7GFP yielding pMB1-CmFRT-Trc-(8RP)P450-*l*-(MA)CPR and pMB1CmFRT-T7-(8RP)P450-*l*-(MA)CPR, respectively.

The (8RP)CPR was amplified using Fw-(8RP)CPR-NdeI & Rv-CPR-PvuI from p5Trc-(8RP)P450-*l*-(MA)CPR and subsequently cut with NdeI/PvuI. The NdeI/PvuI digested PCR fragment was ligated in the NdeI/PvuI site of pMB1CmFRT-T7-(8RP)P450-*l*-(MA)CPR yielding pMB1CmFRT-T7-(8RP)CPR. All primers used for these steps can be found in Supplemental Table 2.

Assay of MMME plasmid expression strength. Fluorescence signal intensity was used to characterize the expression strength among the selected promoters and plasmid copy number combinations. Host cell *E. coli* MG1655 △*recA* △*endA* DE3 carrying Trc-GFP or T7-GFP module on *malT* locus or transformed with MMME expression plasmids (p5Trc/T7-GFP, p10Trc/T7-GFP) was grown overnight in LB with appropriate antibiotic at 37°C, 250 rpm.

The next morning, 2 mL fresh LB with appropriate antibiotic + 0.1 mM IPTG was inoculated with 20 µl overnight culture in 15 mL Corning (Corning, NY, USA) tubes (optical density at 600 nm (OD_{600}) of 0.1 in 96 well plate), and subsequently 150 µL cell culture was transferred to Corning 96-well fluorescence plate (Corning 96 Flat bottom black, clear bottom). Cells in fluorescence plate were grown at 30°C with orbital shaking (540 seconds, 2 mm amplitude, every 5 seconds) in a Tecan Infinite 200 Pro plate reader (Morrisville, NC, USA). Cell optical density and expression of green fluorescence protein were simultaneously detected every 10 minutes. Optical density was read at 600 nm and the excitation and emission wavelengths for GFP were set at 485 ± 9 nm and 535 ± 20 nm, respectively. Normalized expression strength was determined as GFP/OD_{600} relative to GFP/OD_{600} of *E. coli* MG1655 △*recA* △*endA* DE3 △*malT*::Trc-GFP. All experiments were performed in triplicates.

Omics strain preparation. Omics samples were prepared identical to the outline for the small scale Hungate production fermentations. However, the proteomic and transcriptomics samples tubes contained no dodecane overlay to prevent interference with downstream processing. Samples were removed for analysis at 8, 24, 48, and 96 hours. Transcriptomics and proteomics experiments were performed in triplicate.

RNA Purification. Total RNA from *E. coli* was isolated from 1 OD_{600}, 1 mL equivalent of cells in duplicate at each sampled time point. After harvesting, a 2x volume of RNAprotect Cell Reagent (Qiagen) was added to the samples, which were vortexed and then immediately frozen in liquid nitrogen and stored at -80° Celsius for 2 weeks. The RNA was purified using the RNeasy Mini Kit (Qiagen) according to Protocol 4 for gram negative bacteria in complex media with cell lysis and proteinase K. RNA concentration was quantified using a Thermo Scientific NanoDrop 2000c spectrophotometer (Waltham, MA, USA) at the 260/280 and 260/230 ratios to evaluate protein and solvent contamination. The RNA was then diluted to a concentration of 16.7 ng/μL.

Creation of Primers and Probes. Taqman Primers and probes were created for each gene of interest based on the design parameter criteria described by Beacon Designer™ and AlleleID® (Premier Biosoft). Amplicon lengths of between 20-30 base pairs were chosen. Primers and probes were ordered from Integrated DNA Technologies (IDT) with fluorophores added to the 5' gene end and a ZEN-Iowa Black FQ quencher was added to the 3' gene end. Five fluorophores compatible with the BioRad RT-PCR Thermocycler were chosen including FAM, TEX615, TYE705, HEX, and CY5. Two fluorophores were dedicated solely to two reference genes and the remaining three fluorophores were used for the pathway genes of interest. Primers were hydrated in nuclease free H_2O and fluorophores were rehydrated in TE buffer (Qiagen recipe) and both were stored in dark conditions at -20°C.

Primer and Probe Verification. In order to ensure accuracy, all Taqman primers and probes were verified using the standard curve method prior to experimentation. Probes were verified in 5-plex with RNA isolated from strain **7**. Six 1:1 serial dilutions of RNA were performed so that the total RNA per well ranged from 200 ng to 6.25 ng. All preparations were complete out of direct light to prevent degradation of the TaqMan probes. All components were kept on ice throughout the experiment. The RT-PCR was performed in accordance to the Quantitect Multiplex RT-PCR (Qiagen) protocol for 4-Plex with the following modifications: a fifth compatible probe was added at the same volume with a reduction in final water volume, the 20x primer-probe mixes were prepared fresh to a final concentration of 4 μM forward primer, 4 μM reverse primer, and 4 μM probe for each targeting gene, and the final total volume of the reaction was scaled down to 25 μL. The resulting master-mix was aliquotted into a Bio-Rad Multiplate™ Low-Profile 96-well Unskirted PCR plate (Hercules, CA, USA) kept on ice. The template RNA was then added to the wells, covered in Bio-Rad Microseal® 'B' Adhesive film and centrifuged for 30 seconds at 2500 rpm at 4ºC. The plate was run in the BioRad RT-PCR Thermocycler according to the following protocol: one time 15 minute 95ºC step for the PCR initial activation, then a 94ºC 60 seconds step for denaturation, and a 60ºC 90 seconds step for annealing and extension. The denaturation and elongation steps were repeated for an additional 45 cycles. Standard curves were analyzed using the Bio-Rad CFX Manager 3.1 software to determine efficiency. Probes and primers were accepted within an efficiency range of 90-110%.

Cell Lysis for Obtaining Soluble and Insoluble Protein Fractions. Cell pellets were removed from -80˚C and immediately re-suspended in 100 μL BugBuster Master Mix (EMD Millipore). The cell suspension was incubated at room temperature for 30 minutes with periodic shaking to ensure adequate mixing and lysis. After 30 minutes, the suspension was centrifuged for 2 minutes at 1500 rpm using an Eppendorf 5424 Bench-top Centrifuge. 70 μL of supernatant was transferred to a new 1.5 mL Eppendorf for soluble fraction analysis. The remaining cell pellet was washed with 1 mL MilliQ water and centrifuged for one minute at 1,500 rpm. The supernatant was discarded via aspiration and the resulting pellet was considered the insoluble fraction. Both the soluble and insoluble fractions were then subjected to protein extraction.

Protein Extraction via Chloroform/Methanol Precipitation. 400 μL methanol and 100 μL chloroform were added to each Eppendorf tube containing either the soluble or insoluble protein fraction. Eppendorf tubes were then vortexed vigorously for 10 seconds. 300 μL MilliQ water was then added to the mixture and vortexed for an additional 10 seconds. The cell suspension was centrifuged for 1 minute at 1,500 rpm to generate three different phase layers: an aqueous top layer, a circular flake of protein in the interphase, and a smaller chloroform layer at the bottom. The top layer was next carefully removed via aspiration. 400 μL methanol was then added to the solution and vortexed for 10 seconds. Next, the tube was centrifuged 3 minutes at 1,500 rpm to pellet the protein. The supernatant was removed carefully via aspiration and the resulting protein pellet was placed in an Eppendorf Vacufuge at 30˚C for 30 minutes to dry. After 30 minutes of drying, the protein pellet was reconstituted in 90 μL of 50 mM ammonium biocarbonate in 10% acetonitrile with 0.1 μg/μL BSA and 1 mM $CaCl_2$ and placed in a sonication bath for 5 minutes.

Trypsin Digestion. After sonication, Eppendorf tubes containing reconstituted cell pellets were removed and placed on the benchtop. 10 μL of 0.1 μg/μL trypsin was added to each Eppendorf tube, and incubated at 37˚C for 12 hours. After 12 hours of

incubation, the reaction was terminated by the addition of 23 µL of 1 M formic acid and centrifuged for 1 minute at 1,500 rpm to pellet any remaining particulates. 60 µL of the supernatant was then transferred to an HPLC vial containing a glass insert, labeled and stored at -20˚C until subjected to LC-MS targeted proteomics analysis.

Ultra-Pressure Liquid Chromatography Separation. An Agilent Zorbax Eclipse Plus (Santa Clara, CA, USA) C18 2.1 x 150 mm, 95A column was used for UPLC separation. The Agilent 1200 series HPLC system consisted of a high-performance binary pump, a 108 well plate autosampler set at 4˚C, and a column compartment set at 40˚C. The solvent system included two mobile phases. Mobile phase A consisted of 0.1% Formic acid in water/acetonitrile (95:5). Mobile phase B consisted of 0.1% formic acid in acetonitrile/water (95:5). A total of 3 µL of the sample was drawn up and ejected onto the column at a rate of 300 µL/min. The flow rate of solvent through the column was 300 µL/min. The analytical column was primed with 100% solvent A for 3 minutes. At 3 minutes, solvent B was programmed to increase linearly from 0 to 50% over a period of 23.50 minutes. At 26.50 minutes, solvent A was programmed to increase linearly from 50% to 100% over a period of 3 minutes, after which the flow remained at 100% solvent A for one additional minute.

Offline fermentation analytics. Throughout the fermentation glycerol and acetate measurements were taken using an Agilent 1100 Series HPLC with a Hi-Plex H (300 x 7.7 mm) column. Flow rate was set to 0.7 mL/min using 100% Buffer A (14 mM H_2SO_4 in MilliQ H_2O) at a run temperature of 60°C. Total run time per sample was 20 minutes. Two MilliQ water washes preceded any set of samples being run. RID detection was carried out using an Agilent 1260 Infinity Series RID, with standard curves used to quantify amount of glycerol and acetate. Using this method the retention times for glycerol and acetate were 11.45 min and 13.1 min, respectively. OD_{600} measurements were taken using a Beckman Coulter DU 720 General Purpose UV/Vis Spectrophotometer. Samples for OD_{600} measurement were diluted to fall in the range of 0.1 to 0.7 for reading.

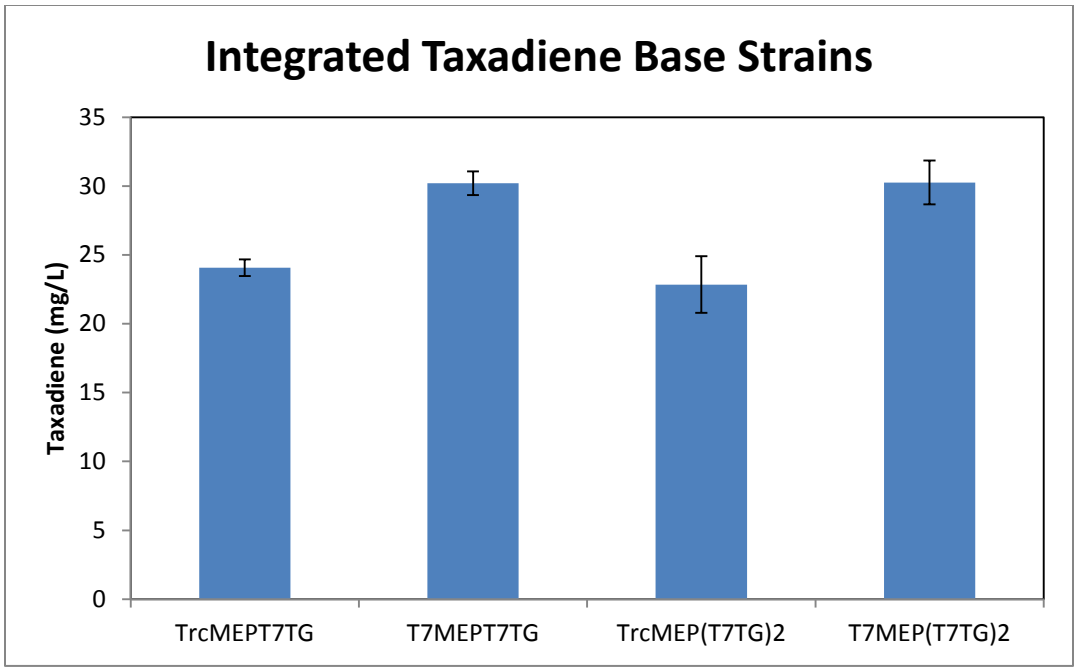


Figure 1 – Taxadiene production from chromosomally integrated, taxadiene producing strains. Variation is based upon promoter strength and copy number. The strains were fermented in 2 mL micro-aerobic Hungate tube fermentations with 10% dodecane overlay. Samples were harvested at 96 hours and quantified with GC-MS. Error bars represent standard deviation of triplicates.

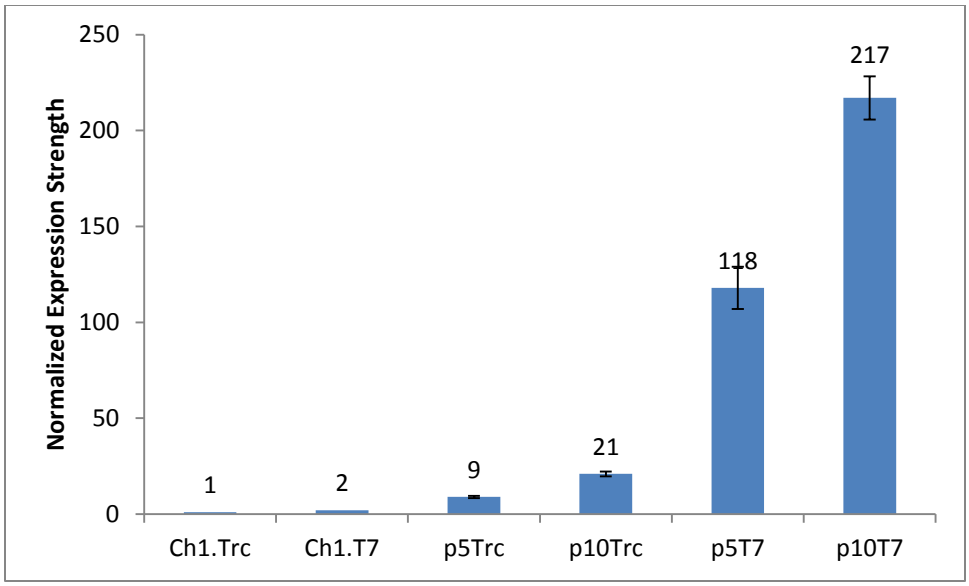


Figure 2 – MMME Expression System. We constructed and systematically evaluated the qualitative, relative protein expression of promoter and copy number variations used in this study, substituting GFP for our gene of interest. Each value is normalized to a single, chromosomally integrated copy under Trc promoter. Error bars represent the standard deviation of triplicates.

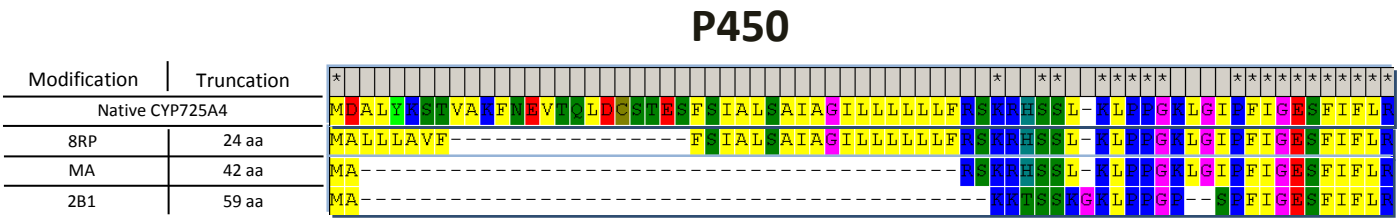


Figure 3 - P450 N-terminal modifications, truncations and sequence alignment. Image depicts specific N-terminal amino acid sequence for each modification. Asterisks indicate residues conserved overall modifications.

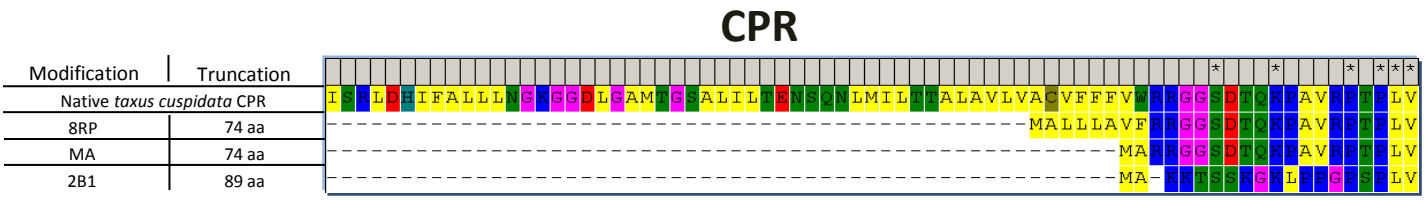
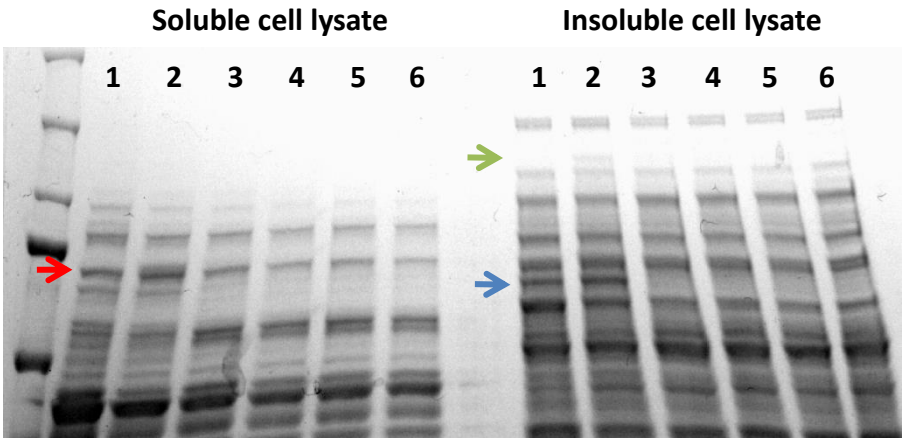


Figure 4 - *Taxus cuspidata* CPR N-terminal modifications, truncations, and sequence alignment. Image depicts specific N-terminal amino acid sequence for each modification. Asterisks indicate residues conserved overall modifications.



1 MEP TG
2 MEP TG - p5Trc(8RP)5aCYP FP
3 MEP TG - p5Trc(8RP)5aCYP(mut) FP
4 MEP TG - p5Trc(MA)5aCYP(mut) FP
5 MEP TG - p5Trc(2b1)5aCYP(mut) FP
6 MEP TG - p5Trc(MA)5aCYP(mut)-(MA)CPR

Figure 5 - Mutation of the internal RBS eliminated expression of truncated CYP-CPR fusion protein. The cells were grown in 15mL glass tubes according to the standard protocol, and the cells were sampled at 72h after inoculation The green arrow indicates the 8RP-5αCYP-CPR protein (120kDa). The red arrow indicates the truncated 8RP-5αCYP-CPR protein (75kDa, not well separated from an endogenous protein this time). The blue arrow indicates an unknown protein (~65kDa, likely to be dxs) whose expression level was significantly reduced when the RBS was mutated.

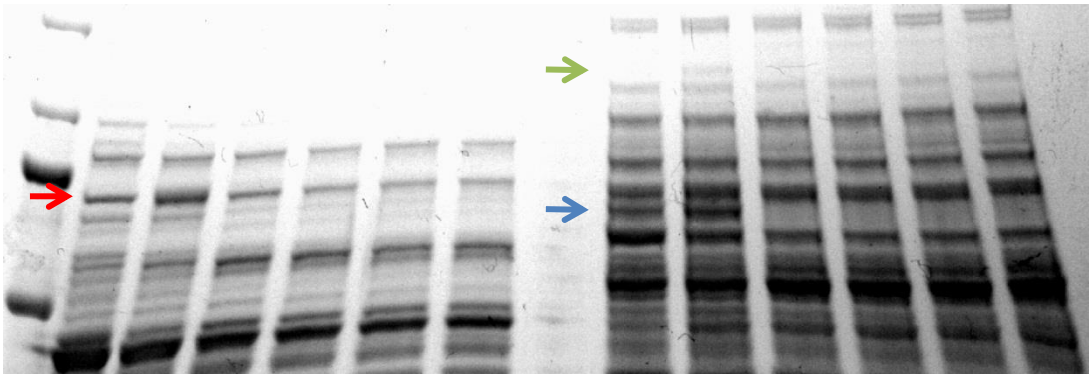


Figure 6 - Gel replicate from above.

Additional Omics Figures

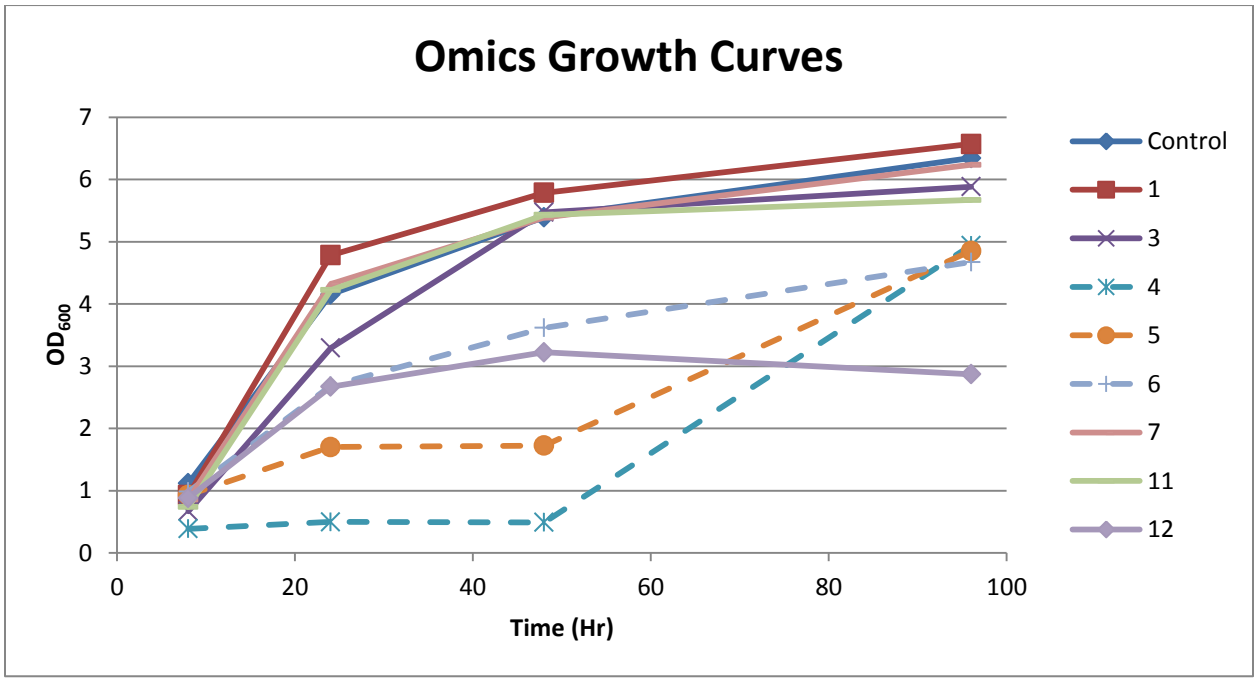


Figure 7 - Growth curve of strains used in the targeted transcriptomics and proteomics experiments. Sample time points at 8, 24, 48, and 96 hours. Dotted lines indicate strains that did not produce terpenes products. Control indicates integrated taxadiene producing strain with no P450 module.

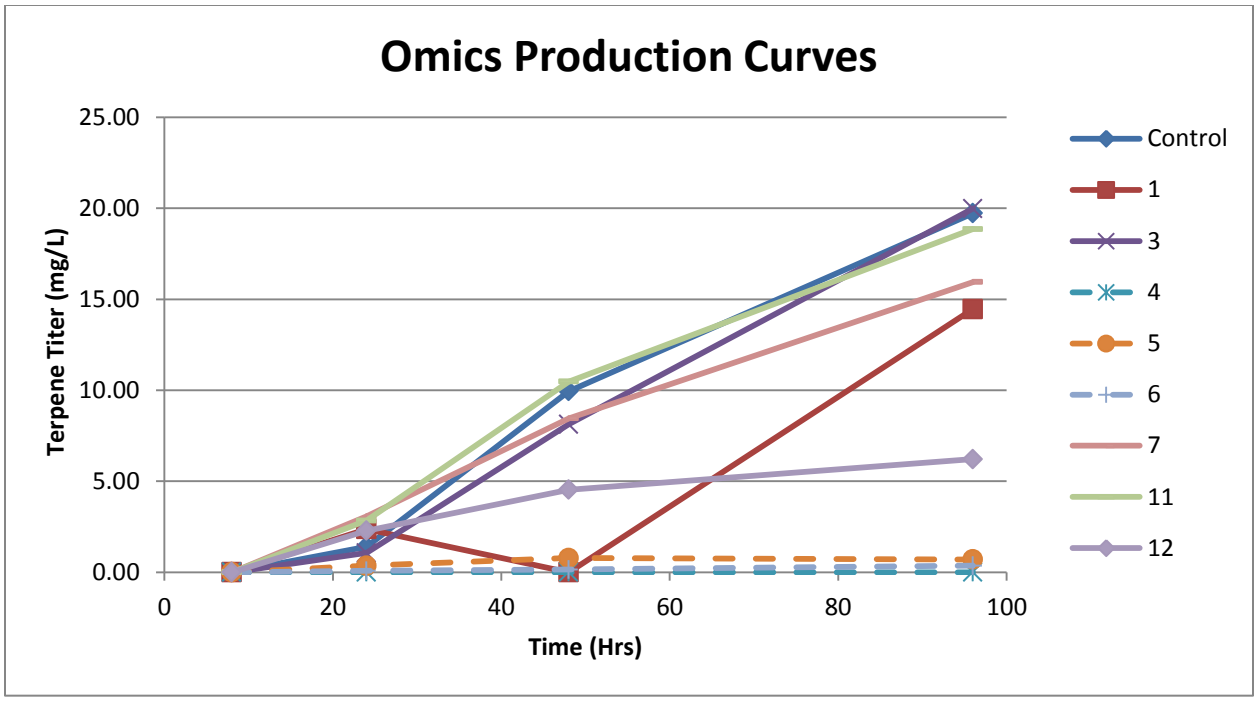


Figure 8 - Omics production curves. Control indicates chromosomally integrated taxadiene "upstream" strain without a P450 module integrated. Each other strain accounts for total terpenes (taxadiene and oxygenated taxanes). In excess of the control, only strains 2 and 8 had unconverted taxadiene at the end of the fermentation, and only minimal (<3 mg) amounts. Dotted lines indicate poor performing strains.

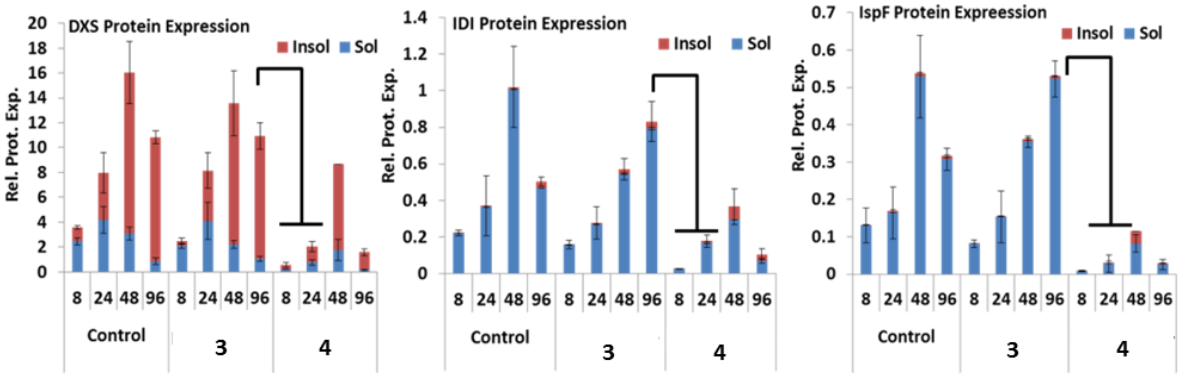


Figure 9 - Time course of proteome-level disturbance for upstream MEP enzymes, dxs, idi, ispF. The MEP pathway is expressed as an operon in the order of dxs, idi, ispDF under a chromosomally integrated T7 promoter for each strain. The CYP/CPR expression increases from 5 to 10 copy plasmids under a Trc promoter for strains 3 to 4, respectively.

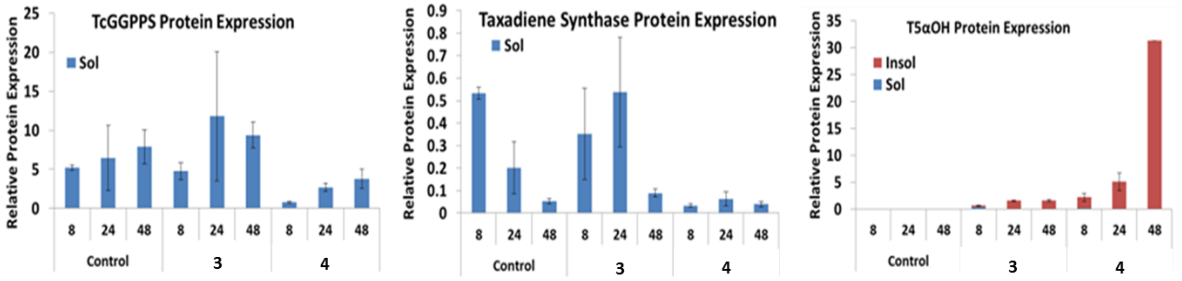


Figure 10 – Time course of Taxus GGPPS (TcGGPPS), taxadiene synthase, and CYP expression profiles. Synthases (GGPPS and TxS) are expressed from the chromosome under a T7 promoter in all strains. The CYP/CPR expression increases from 5 to 10 copy plasmids under a Trc promoter for strains 3 to 4, respectively.

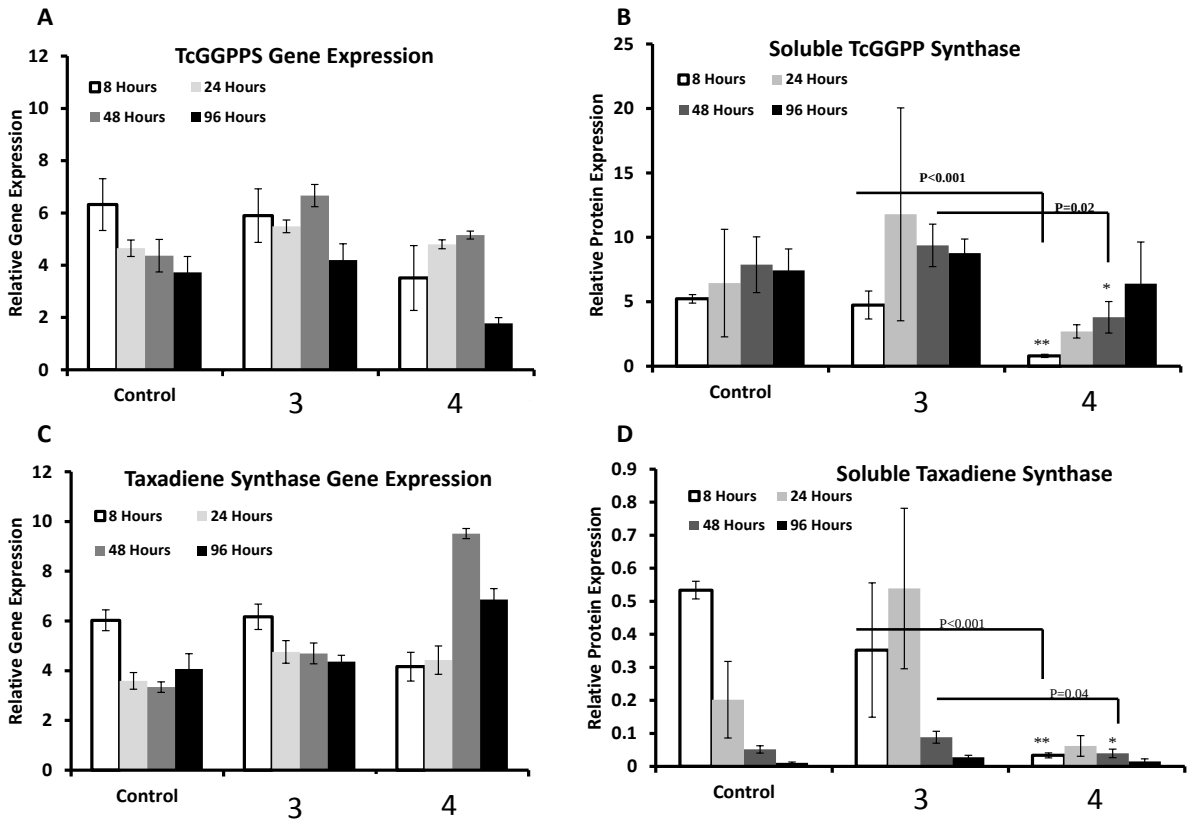


Figure 11 - Heterologous proteins are regulated at the proteome. (A) Transcriptomics results show that increased expression of the P450 module does not measurably affect the heterologous TcGGPP Synthase and (C) Taxadiene Synthase gene expression. However, (B) the same increase in P450 expression a decrease in soluble TcGGPP Synthase and (D) soluble Taxadiene Synthase expression, as determined by targeted proteomics. Y-axis represents relative gene expression (A, C) and relative protein expression (B, D). X-axis represents control strain, a high-producing strain (3, p5Trc CYP-L-CPR) and a low-producing strain (4, p10Trc CYP-L-CPR). Error bars represent standard deviation of biological and technical triplicates. Notations for calculation of significance - *p<0.05, **p<0.001.

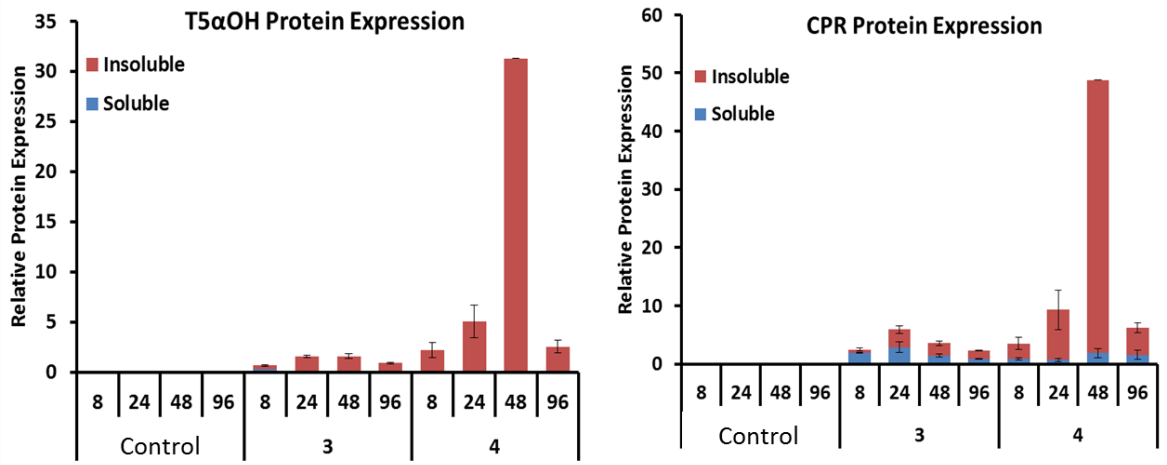


Figure 12 - Time course comparison for T5αOH and CPR relative protein expression across control, high producing (3), and low producing strains (4). As can be seen, significant increases in insoluble protein can be seen for both the P450 and CPR in the higher expression, lower producing strain (4). This is in addition to a modest decrease in soluble CPR. The decrease in overall production at 96 hours is most probably a function of the cells reaching the stationary phase. The impact of this increase in expression on other heterologous proteins can be seen in Supplemental Figure 7. Error bars represent standard deviation of biological and technical triplicates.

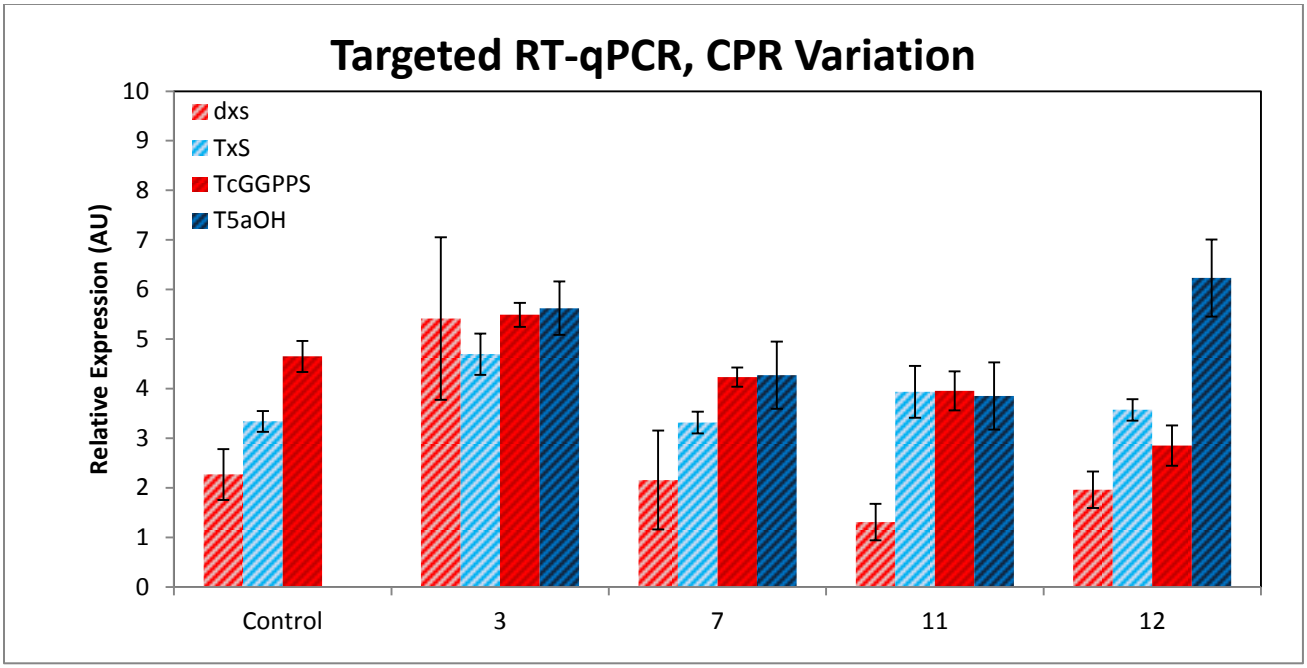


Figure 13 - Additional Targeted RT-qPCR for strains with varying P450 to CPR expression strength. Strain 3 is the linker P450-L-CPR. Strain 7 is the P450-O-CPR operon. Strain 11 has the P450 on five-copy plasmid, with the CPR integrated into the chromosome as a single under T7 promoter. Strain 12 is the five-copy plasmid with the CPR on the plasmid with its own T7 promoter. Transcript data did not show any unexpected trends other than the slightly higher expression of the P450 in strain 12 compared to strain 3 and 7. This might be attributed to leakiness of T7 terminator system in strain 12.[2]

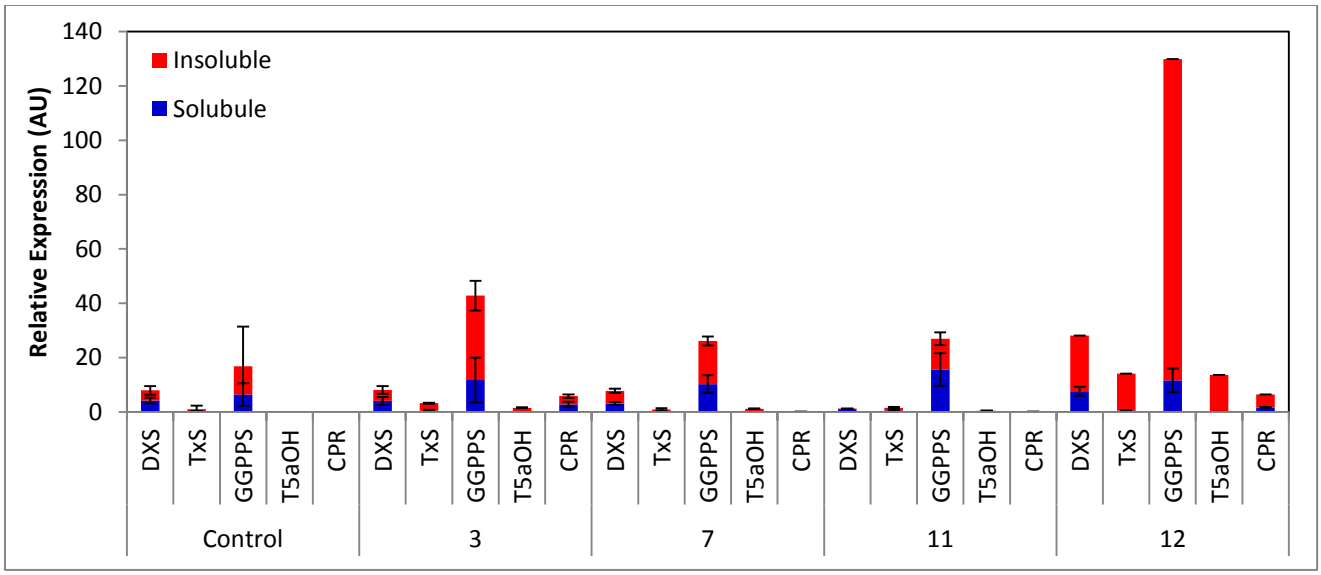


Figure 14 - Additional proteomics for P450 to CPR variation strains at 24 hour time point. Insoluble fraction is indicated in the top red fraction and soluble is indicated by the blue fraction on the bottom. For each strain shown, the P450 is on a five-copy plasmid under a Trc promoter. Strain 3 is the chimera linked P450 to CPR. Strain 7 is the P450 and CPR in operon format. Strain 11 has the CPR integrated as a single copy in the chromosome under a T7 promoter. Strain 12 has the CPR on the five-copy plasmid bearing the P450, but with its own separate T7 promoter.

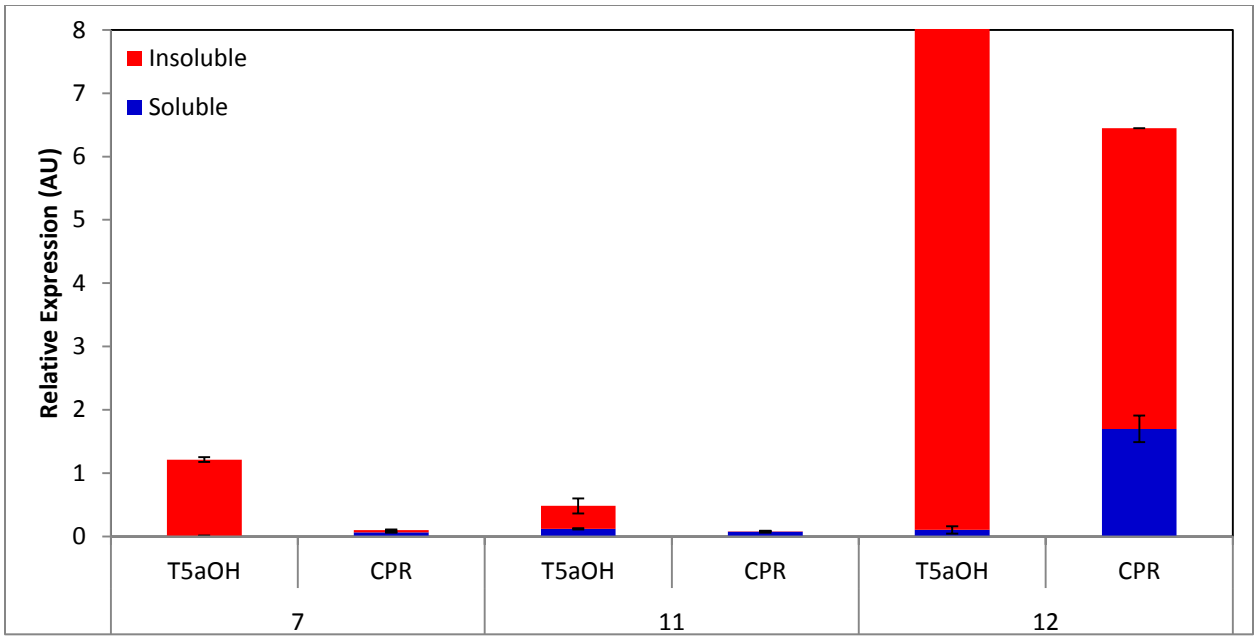


Figure 15 - Additional P450 vs. CPR proteomics. Zoomed in for only P450 and CPR proteomics and focusing on only strains 7, 11, and 12 from Figure 8. Strain 12 has an additional T7 terminator which may increase both enzymes' expression. Insoluble fraction is indicated in the top red fraction and soluble is indicated by the blue fraction on the bottom.

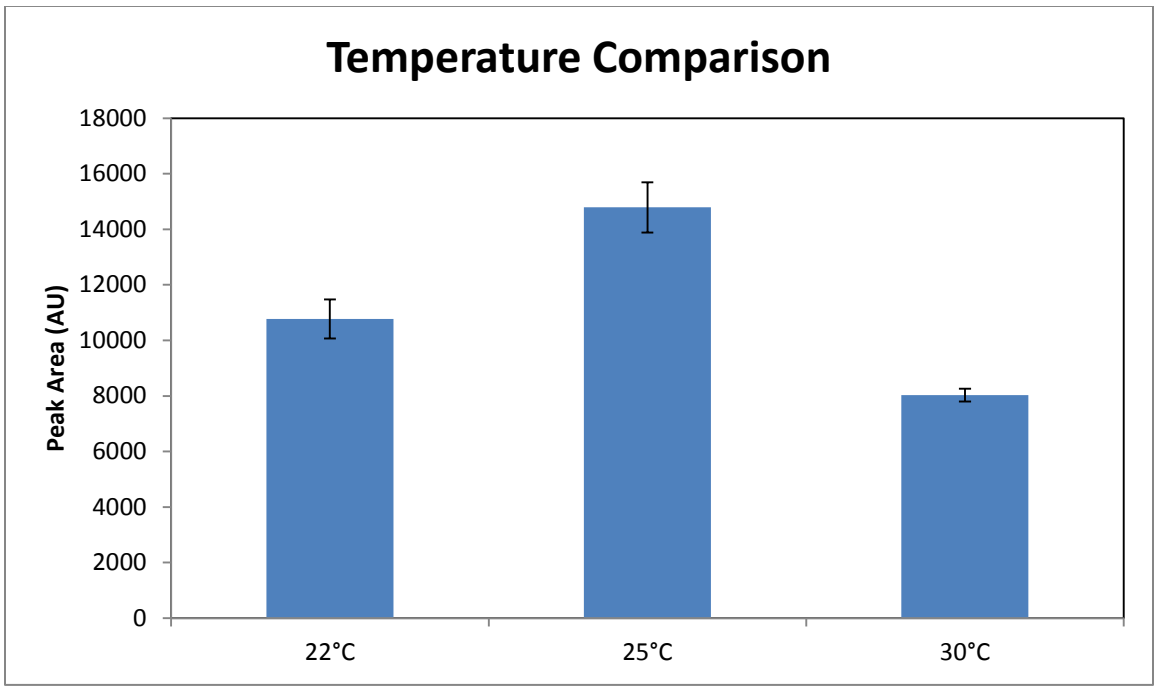


Figure 16 – Strain 7 Small-scale (2 mL) temperature screen. Hungate tube cell culture screening was done in triplicate for the top producing identified strain, strain 7, to evaluate its performance at different temperatures. Four day fermentations were carried out with 10% dodecane overlay at 22°C, 25°C, and 30°C. Error bars represent the standard deviation of triplicates.

Duplicate Fermentation.

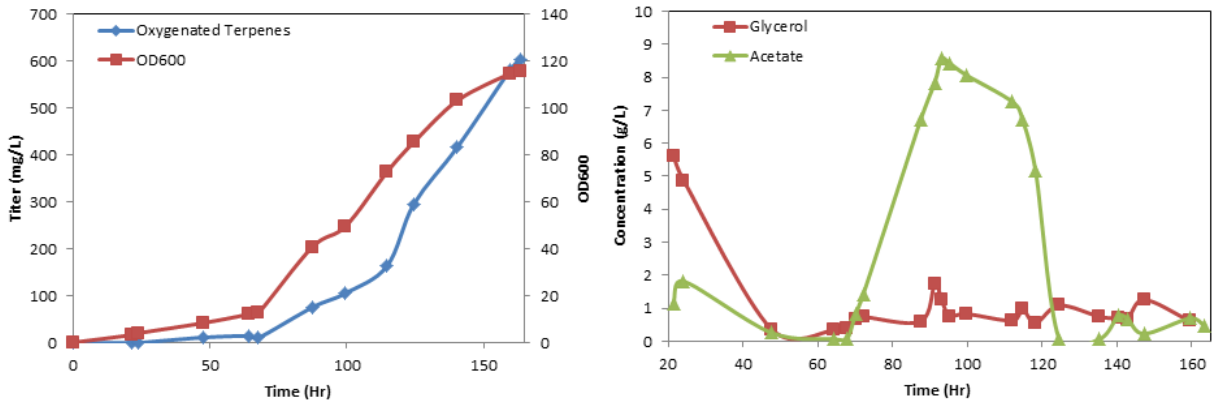


Figure 17 – Replicate fermentation. Replicate fermentation run in 2L Sartorius-Stedim benchtop bioreactors. The left figure indicates oxygenated taxane production and optical density at 600nm (OD_{600}). The right figure shows glycerol and acetate concentrations throughout the duration of the fermentation.

GC-MS.

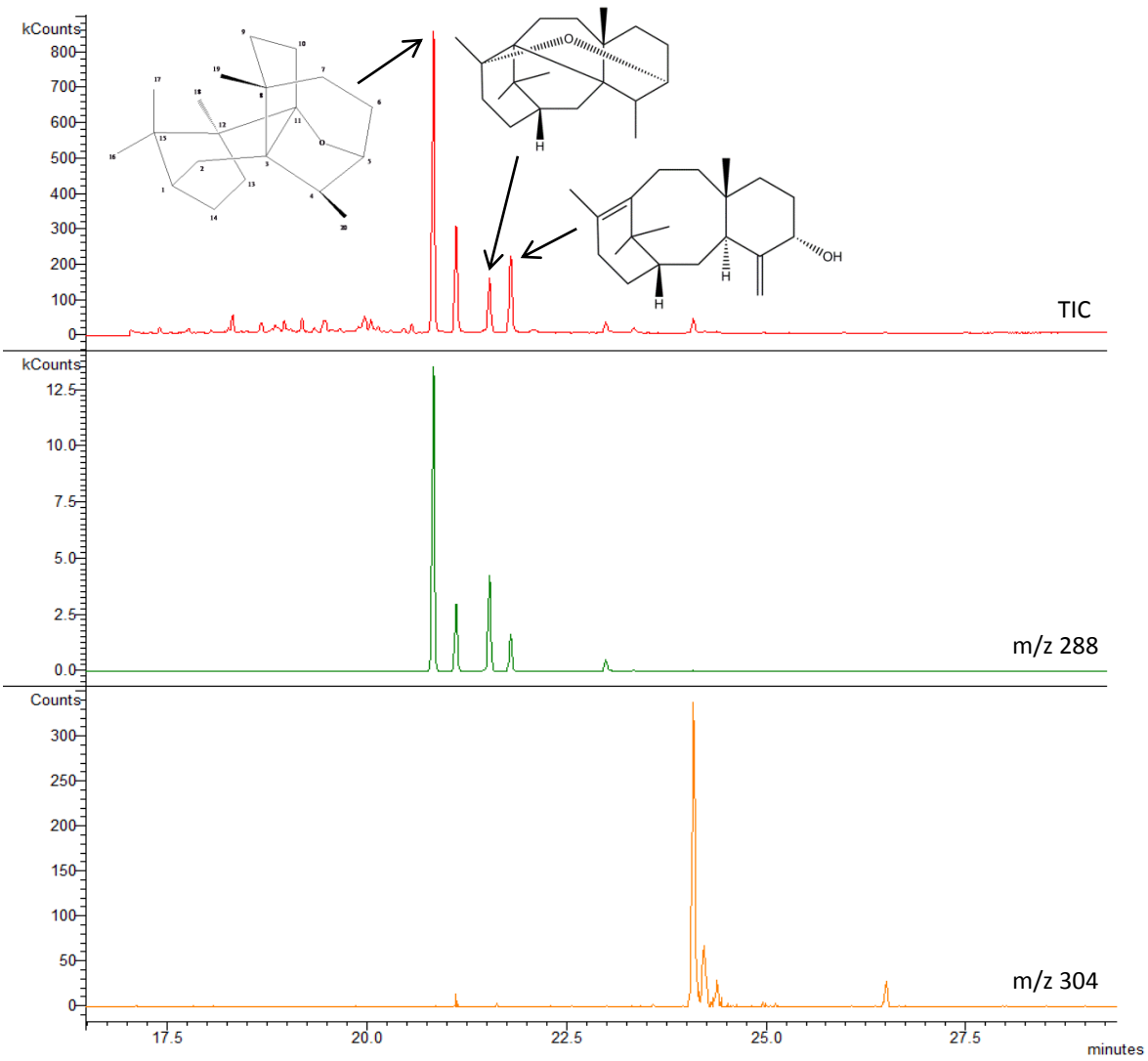


Figure 18 - Overall Fermentation GC. Top panel total is for the ion count and has the structure for the NMR solved compounds represented next to their respective peak. The middle panel shows the GC filtered for m/z 288. The bottom panel shows the GC for m/z 304, indication a potentially double hydroxylated product. The details of each peak and corresponding compounds can be found in Supplemental Figures 14-19.

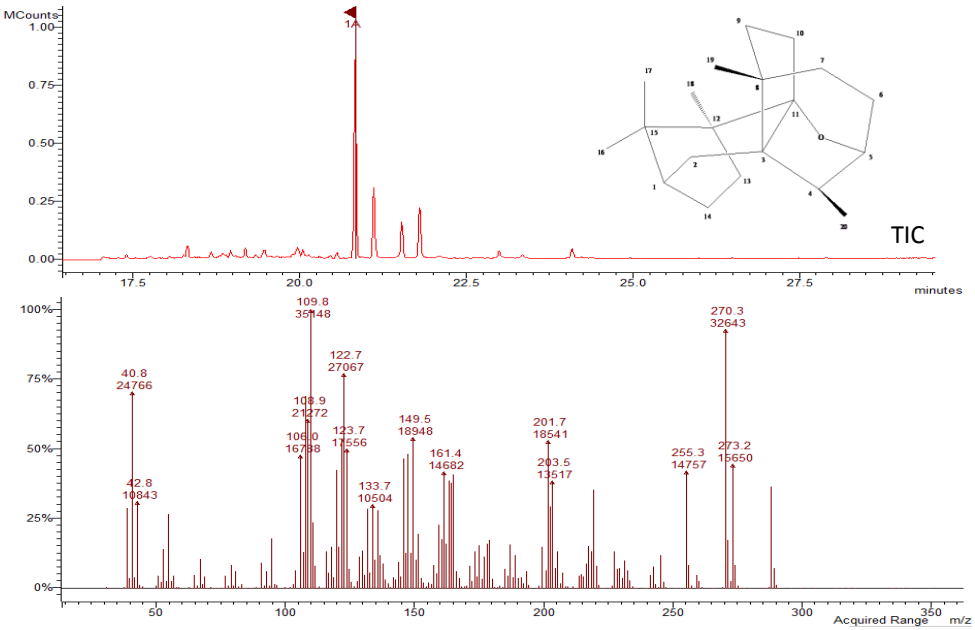


Figure 19 - GC-MS for retention time 20.838 min. This corresponds to iso-OCT. The structure was been previously solved by 2D-NMR and is currently submitted for publication.

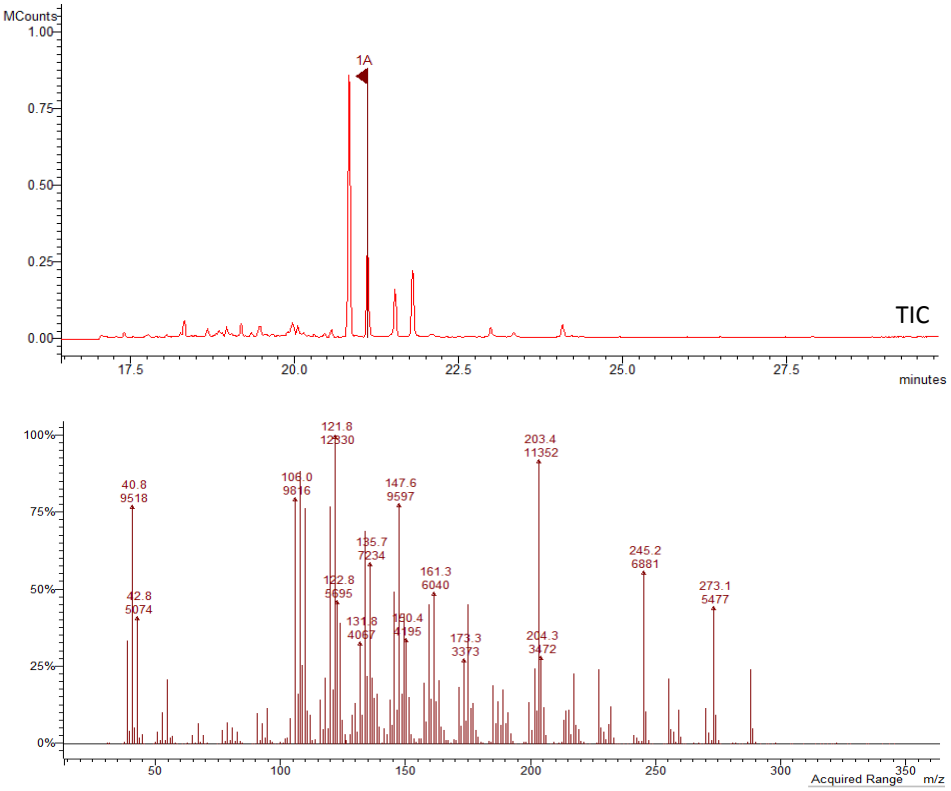


Figure 20 - GC-MS for retention time 21.114 min. This represents an unknown byproduct with m/z 288. It is likely another oxygenated taxane, although it is less stable than other oxygenated taxanes. Accordingly, its structure has yet to be solved.

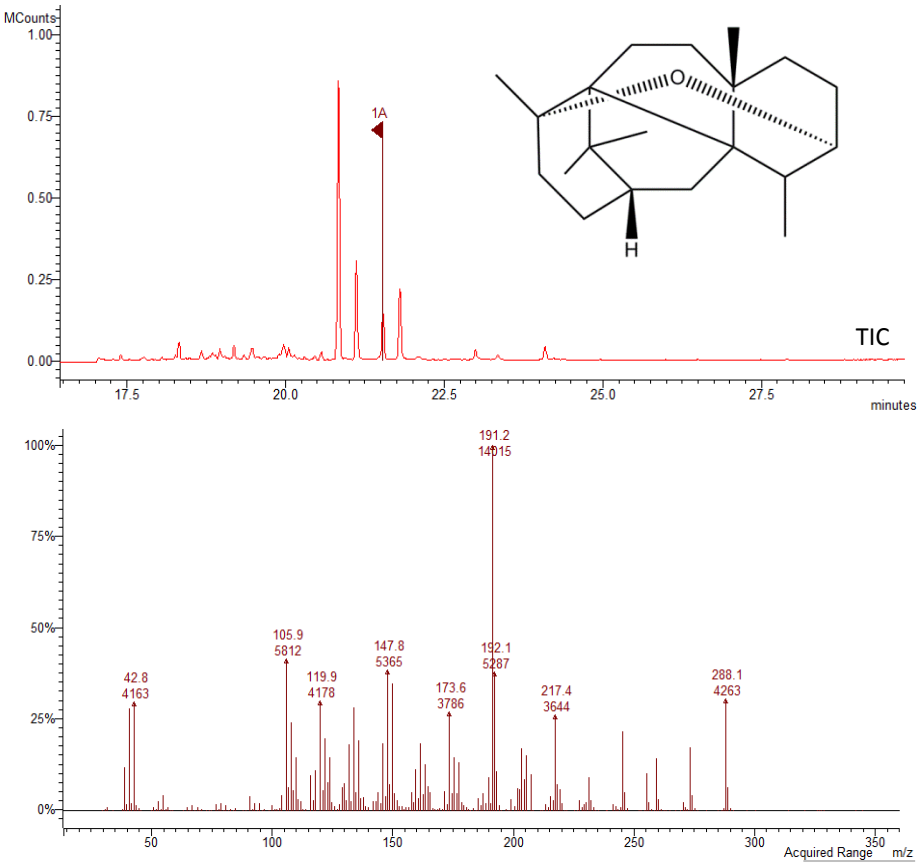


Figure 21 - GC-MS for retention time 21.536 min. This corresponds to OCT.

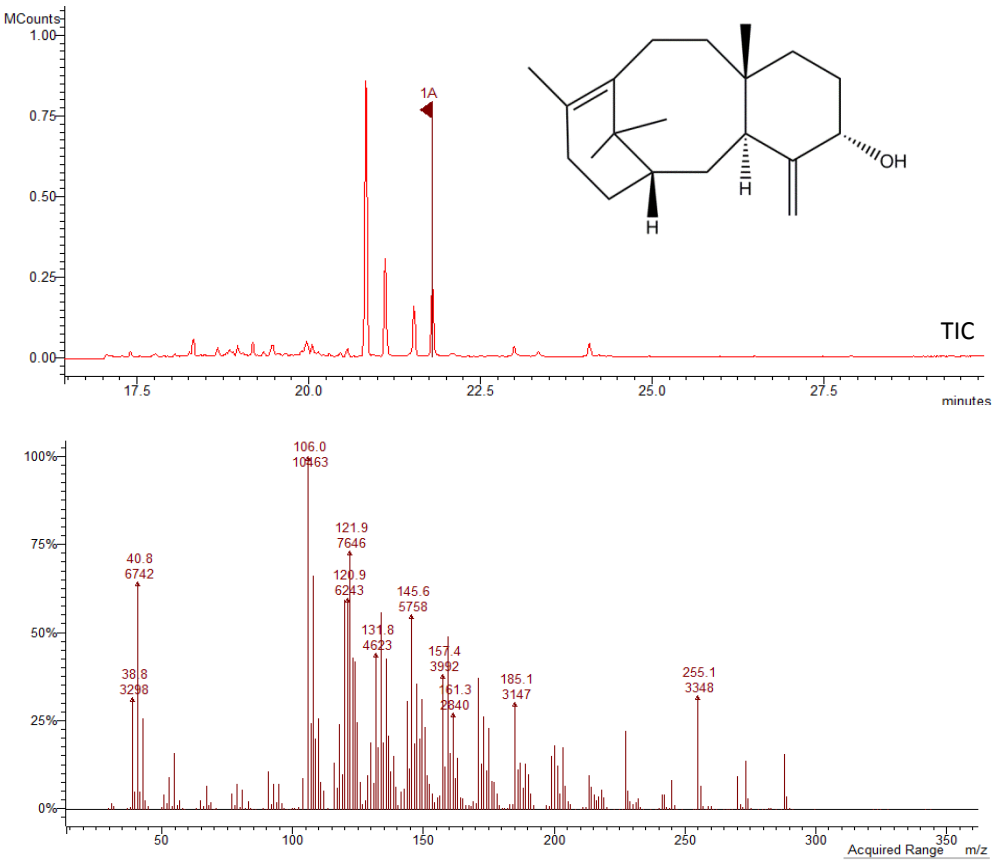


Figure 22 - GC-MS for retention time 21.800 min. This corresponds to taxadiene-5alpha-ol.

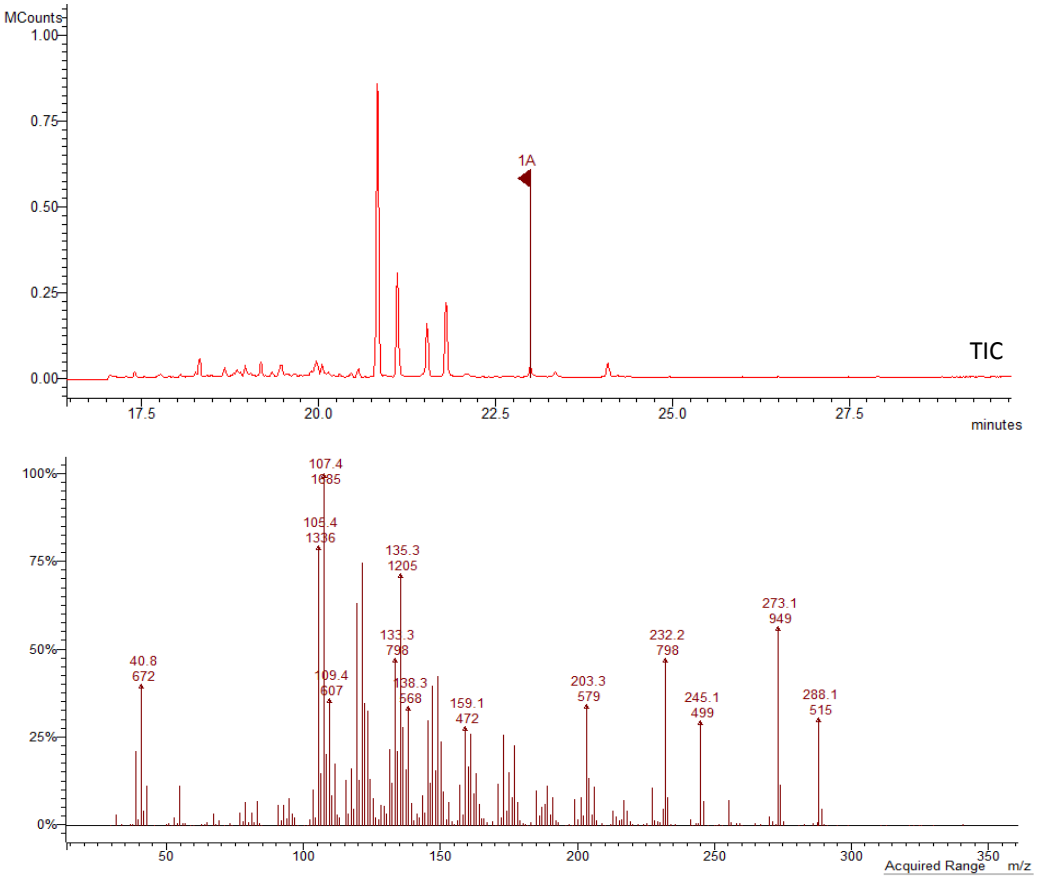


Figure 23 - GC-MS for retention time 22.995 min. This compound likely represents another unknown oxygenated taxane structure, indicated as per its m/z 288 and 273 ion fragmentation. Its minimal accumulation precludes isolation and structural identification.

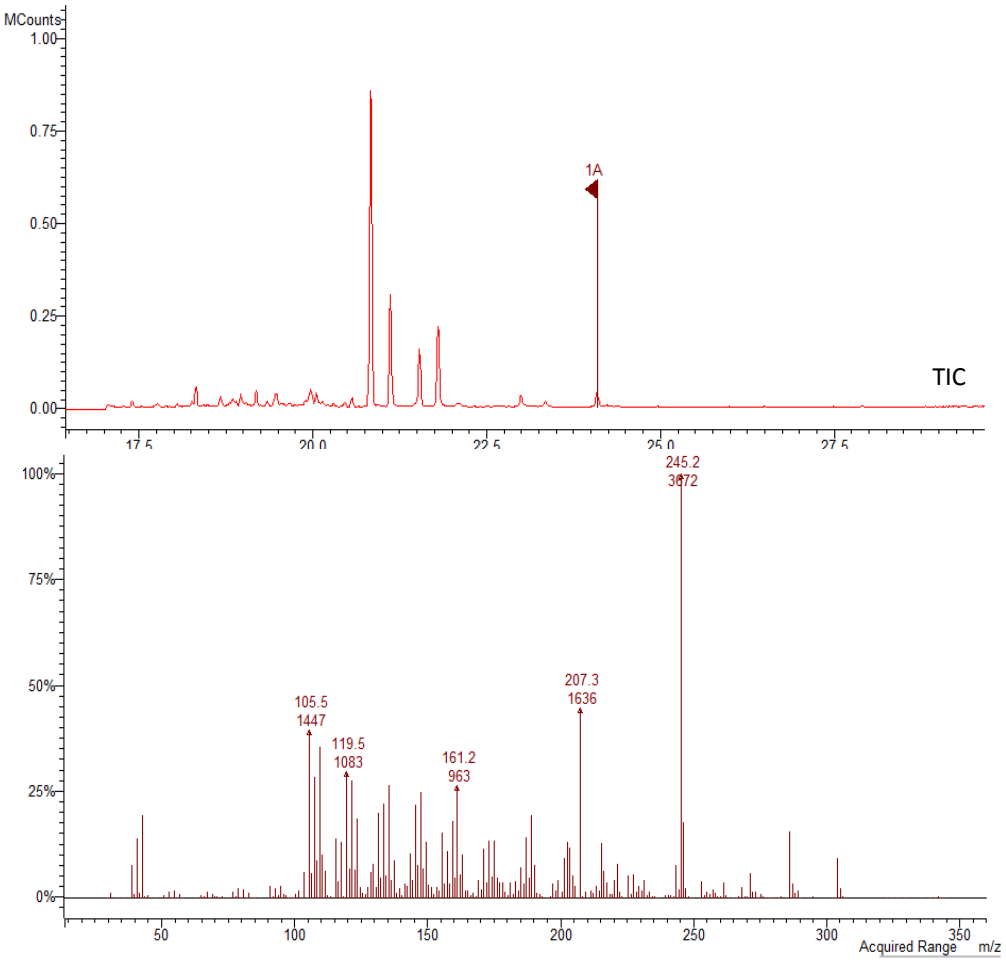


Figure 24 - GC-MS for retention time 24.090 min. The unknown, potentially doubly oxygenated taxane product indicated by the m/z 304 peak (272 + 16 + 16).

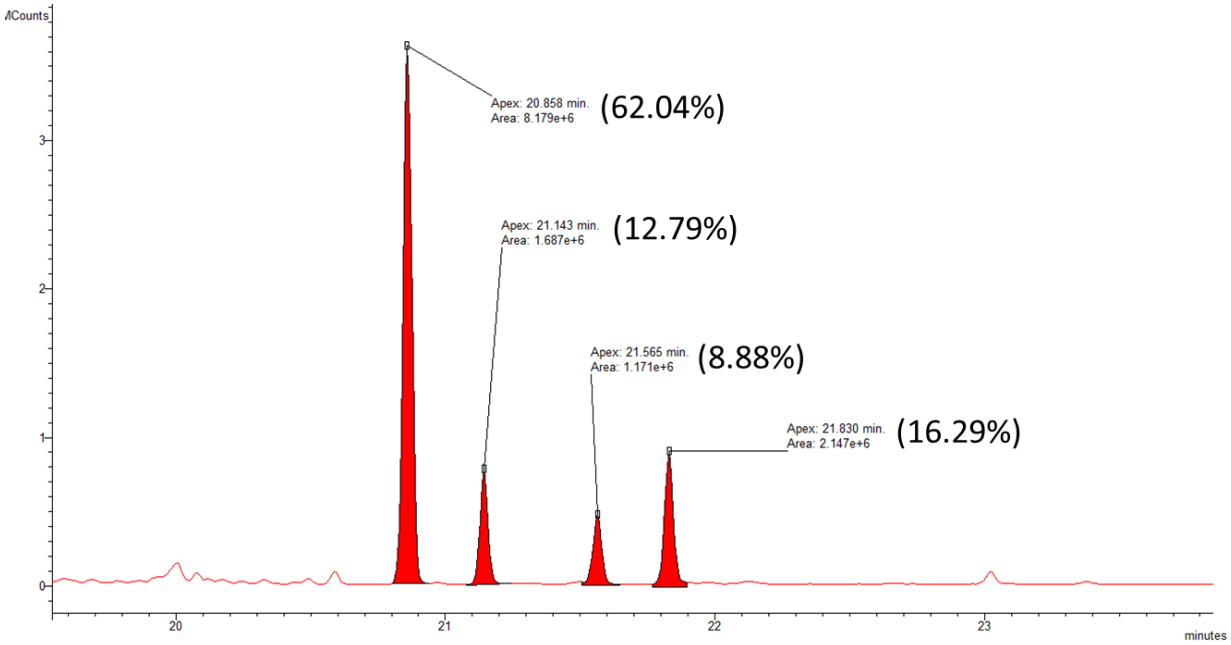


Figure 25 – Fermentation GC-MS for oxygenated species percentages. Total peak area is summed, and the respective percentage of each compound in terms of the total oxygenated taxane titer is presented. From left to right, the compounds are iso-OCT, an unidentified oxygenated taxane, OCT, and taxadiene-5α-ol.

Table 1 - List of plasmids and strains

Plasmid or strain	Relevant properties or genotype	Source or reference
Genetic engineering plasmids		
pKD46, pCP20	Helper plasmids for homologous recombination	[3]
pMB1CmFRT-GFP	Chromosomal integration plasmid	in house made
pMB1CmFRT-TrcGFP	Chromosomal integration plasmid	this paper
pMB1CmFRT-T7GFP	Chromosomal integration plasmid	this paper
p10T7-TG-KanFRT	Chromosomal integration plasmid	this paper
pMB1CmFRT-Trc-(8RP)P450-*l*-(MA)CPR	Chromosomal integration plasmid	this paper
pMB1CmFRT-T7-(8RP)P450-*l*-(MA)CPR	Chromosomal integration plasmid	this paper
pMB1CmFRT-T7-(8RP)CPR	Chromosomal integration plasmid	this paper
MMME expression plasmids		
p5Trc, p10Trc, p5T7, p10T7, p20Trc	MMME expression plasmids	[1]
p5Trc-GFP	p5TrcGFP; pSC101 *ori*, *Trc* promoter, SpR MMME expression level determination	in house made
p5T7-GFP	p5T7GFP; pSC101 *ori*, *T7* promoter, SpR MMME expression level determination	in house made
p10Trc-GFP	p10TrcGFP; p15A *ori*, *Trc* promoter, CmR MMME expression level determination	in house made
p10T7-GFP	p10T7GFP; p15A *ori*, *T7* promoter, CmR MMME expression level determination	in house made
P450 module expression plasmids		
p5Trc(8RP)P450-*l*-CPR	pSC101 *ori*, *Trc* promoter, chimera, 8RP N-terminus of P450	this paper
p10Trc(8RP)P450-*l*-CPR	P15A *ori*, *Trc* promoter, chimera, 8RP N-terminus of P450	this paper
p5T7(8RP)P450-*l*-CPR	pSC101 *ori*, *T7* promoter, chimera, 8RP N-terminus of P450	this paper
p10T7(8RP)P450-*l*-CPR	P15A *ori*, *T7* promoter, chimera, 8RP N-terminus of P450	this paper
p5Trc(8RP)P450-*o*-(8RP)CPR	pSC101 *ori*, *Trc* promoter, operon, 8RP N-terminus of P450 and CPR	this paper
p10Trc(8RP)P450-*o*-(8RP)CPR	P15A *ori*, *Trc* promoter, operon, 8RP N-terminus of P450 and CPR	this paper
p5T7(8RP)P450-*o*-(8RP)CPR	pSC101 *ori*, *T7*promoter, operon, 8RP N-terminus of P450 and CPR	this paper
p10T7(8RP)P450-*o*-(8RP)CPR	P15A *ori*, *T7* promoter, operon, 8RP N-terminus of P450 and CPR	this paper
p5Trc(8RP)P450	pSC101 *ori*, *Trc* promoter, 8RP N-terminus of P450	this paper
p5Trc(8RP)P450-T7(8RP)CPR	pSC101 *ori*, *Trc* and *T7* promoter, 8RP N-terminus of P450 and CPR	this paper
p5Trc(8RP)P450-*o*-(MA)CPR	pSC101 *ori*, *Trc* promoter, operon, 8RP N-terminus of P450 and MA N-terminus of CPR	this paper
p5Trc(8RP)P450-*o*-(2b1)CPR	pSC101 *ori*, *Trc* promoter, operon, 8RP N-terminus of P450 and 2b1 N-terminus of CPR	this paper

p5Trc(MA)P450-*o*-(MA)CPR	pSC101 *ori*, *Trc* promoter, operon, MA N-terminus of P450 and MA N-terminus of CPR	this paper
p5Trc(MA)P450-*o*-(2b1)CPR	pSC101 *ori*, *Trc* promoter, operon, MA N-terminus of P450 and 2b1 N-terminus of CPR	this paper
p5Trc(2b1)P450-*o*-(MA)CPR	pSC101 *ori*, *Trc* promoter, operon, 2b1 N-terminus of P450 and MA N-terminus of CPR	this paper
p5Trc(2b1)P450-*o*-(2b1)CPR	pSC101 *ori*, *Trc* promoter, operon, 2b1 N-terminus of P450 and 2b1 N-terminus of CPR	this paper
Strains		
E. coli MG1655 △*recA* △*endA* DE3	Wild-type	[1]
E. coli MG1655 △*recA* △*endA* DE3 △*malT*::Trc-GFP	Ch1.TrcGFP MMME expression level determination	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*malT*::T7-GFP	Ch1.T7GFP MMME expression level determination	this paper
E. coli MG1655 △*recA* △*endA* DE3 + p5Trc-GFP	p5TrcGFP MMME expression level determination	this paper
E. coli MG1655 △*recA* △*endA* DE3 + p5T7-GFP	p5T7GFP MMME expression level determination	this paper
E. coli MG1655 △*recA* △*endA* DE3 + p10Trc-GFP	p10TrcGFP MMME expression level determination	this paper
E. coli MG1655 △*recA* △*endA* DE3 + p10T7-GFP	p10T7GFP MMME expression level determination	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::Trc-MEP △*lacY*::T7-TG	Ch1.TrcMEP Ch1.T7TG Taxadiene expression strain	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG	Ch1.T7MEP Ch1.T7TG Taxadiene expression strain =control strain	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::Trc-MEP △*lacY*::T7-TG △*malT*::T7-TG	Ch1.TrcMEP Ch2.T7TG Taxadiene expression strain	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG △*malT*::T7-TG	Ch1.T7MEP Ch2.T7TG Taxadiene expression strain	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG △*malT*::Trc(8RP)P450-*l*-CPR	T7MEPT7TG Ch1.Trc(8RP)P450-*l*-CPR, Oxygenated taxadiene expression strain = strain 1	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG △*malT*::T7(8RP)P450-*l*-CPR	T7MEPT7TG Ch1.T7(8RP)P450-*l*-CPR, Oxygenated taxadiene expression strain = strain 2	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p5Trc(8RP)P450-*l*-CPR	T7MEPT7TG + p5Trc(8RP)P450-*l*-CPR, Oxygenated taxadiene expression strain = strain 3	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p5T7(8RP)P450-*l*-CPR	T7MEPT7TG + p5T7(8RP)P450-*l*-CPR, Oxygenated taxadiene expression strain = strain 4	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p10Trc(8RP)P450-*l*-CPR	T7MEPT7TG + p10Trc(8RP)P450-*l*-CPR, Oxygenated taxadiene expression strain = strain 5	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p10T7(8RP)P450-*l*-CPR	T7MEPT7TG + p10T7(8RP)P450-*l*-CPR, Oxygenated taxadiene expression strain = strain 6	this paper

E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p5Trc(8RP)P450-*o*-(8RP)CPR	T7MEPT7TG + p5Trc(8RP)P450-*o*-(8RP)CPR, Oxygenated taxadiene expression strain = strain 7	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p5T7(8RP)P450-*o*-(8RP)CPR	T7MEPT7TG + p5T7(8RP)P450-*o*-(8RP)CPR, Oxygenated taxadiene expression strain = strain 8	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p10Trc(8RP)P450-*o*-(8RP)CPR	T7MEPT7TG + p10Trc(8RP)P450-*o*-(8RP)CPR, Oxygenated taxadiene expression strain = strain 9	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p10T7(8RP)P450-*o*-(8RP)CPR	T7MEPT7TG + p10T7(8RP)P450-*o*-(8RP)CPR, Oxygenated taxadiene expression strain = strain 10	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG △*malT*::T7(8RP)CPR + p5Trc(8RP)P450	T7MEPT7TG Ch1.T7(8RP)CPR + p5Trc(8RP)P450, Oxygenated taxadiene expression strain = strain 11	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p5Trc(8RP)P450-T7(8RP)CPR	T7MEPT7TG + p5Trc(8RP)P450-T7CPR, Oxygenated taxadiene expression strain = strain 12	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p5Trc(8RP)P450-*o*-(MA)CPR	T7MEPT7TG + p5Trc(8RP)P450-*o*-(MA)CPR, Oxygenated taxadiene expression strain = strain 13	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p5Trc(8RP)P450-*o*-(2b1)CPR	T7MEPT7TG + p5Trc(8RP)P450-*o*-(2b1)CPR, Oxygenated taxadiene expression strain = strain 14	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p5Trc(MA)P450-*o*-(MA)CPR	T7MEPT7TG + p5Trc(MA)P450-*o*-(MA)CPR, Oxygenated taxadiene expression strain = strain15	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p5Trc(MA)P450-*o*-(2b1)CPR	T7MEPT7TG + p5Trc(MA)P450-*o*-(2b1)CPR, Oxygenated taxadiene expression strain = strain 16	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p5Trc(2b1)P450-*o*-(MA)CPR	T7MEPT7TG + p5Trc(2b1)P450-*o*-(MA)CPR, Oxygenated taxadiene expression strain = strain 17	this paper
E. coli MG1655 △*recA* △*endA* DE3 △*araA*::T7-MEP △*lacY*::T7-TG + p5Trc(2b1)P450-*o*-(2b1)CPR	T7MEPT7TG + p5Trc(2b1)P450-*o*-(2b1)CPR, Oxygenated taxadiene expression strain = strain 18	this paper

Table 2 - List of oligonucleotides

Nr	Oligonucleotide	Sequence (5' →3')*
1	Fw-(8RP)P450-NdeI	CGCGGCGAGGCGGTGGCGGCATATGGCTCTGTTATTAGCAGTTTTTTTAG CATCGCTTTGAGTGCAATTG
2	Fw-(MA)P450-NdeI	TGGCGCATATGGCTCGCTCGAAACGTCATAGTAG
3	Fw-(2b1)P450-NdeI	TACGCATATGGCTAAGAAAACGAGCTCTAAAGGGAAGCTCCCACCAGGACC TAGCCCGTTTATCGGTGAGTCC
4	Rv-P450-l-BamHI-SalI	TGCAGTCGACTATAGGATCCGGTGCTGCCCGGACGAGGGAACAGTTTG
5	Rv-P450-o-BamHI-SalI	AAGTGTCGACTAGCGGATCCTTACGGACGAGGGAACAG
6	Fw-l-(MA)CPR-BamHI	TATAGGATCCCGCCGTGGTGGAAGTGATAC
7	Fw-o-(8RP)CPR-BamHI	GCGTGCGCGGATCCAAGGAGATATACCATGGCTCTGTTATTAGCAGTTTTTC GCCGTGGTGGAAGTGATAC
8	Fw-o-(MA)CPR-BamHI	TGACGGATCCAAGGAGATATACCATGGCTCGCCGTGGTGGAAGTGATAC
9	Fw-o-(2b1)CPR-BamHI	TGCCGGATCCAAGGAGATATACCATGGCTAAGAAAACGAGCTCTAAAGGG AAGCTCCCACCAGGACCTAGCCCTCTTGTTAAAGAAGAGGACG
10	Rv-CPR-SalI	TGCGAGCGTCGACTTACCAAATATCCCGTAAGTAGC
11	Fw-KmFRT-XhoI	TGCGACCTCGAGGAGCAATAACTAGCATAACC
12	Rv-KmFRT-SalI	TTGATCCGTCGACCTGCAGTTCGAAG
13	Fw-Trc-SacI	TGCGCGAGCTCCTGAAATGAGCTGTTGACAA
14	Rv-Trc-NcoI	GCCGCCATGGTTATTCCTCCTTATTTAATC
15	Fw-T7-SacI	TTATAGAGCTCGACTCTCTTCCGGGCGCTATC
16	Rv-T7-NcoI	TGCCGCAGGCGCCCATGGTATATCTCCTTATTAAAG
17	Fw-Trc-SphI	GCTGGCAGCATGCGCTGTTGACAATTAATCATCC
18	Fw-T7-SphI	GTGCCGCATGCTAATACGACTCACTATAGG
19	Fw-(8RP)CPR-NdeI	TGGCGCATATGGCTCTGTTATTAGCAGTTTTTCGCCGTGGTGGAAGTGATAC
20	Rv-CPR-PvuI	AGCTCGGCGATCGTTACCAAATATCCCGTAAGTAGC
21	Fw-KI-T7TG-KmFRT-H1LacY	ATGATATGTTGGTCGGATAAGGCGCTCGCGCCGCATCCGACATTGATTGCC GGTGCCTAATGAGTGAGCTAAC
22	Rv-KI-T7TG-KmFRT-H2LacY	AATAACCGGGCAGGCCATGTCTGCCCGTATTTCGCGTAAGGAAATCCATTCT AGTTATTGCTCAGCAAGC
23	Fw-KI-T7TG-KmFRT-H1malT	AGCGCAAAAAAAATAATATTTCCTCATTTTCCACAGTGAAGTGATTAACTCC GGTGCCTAATGAGTGAGCTAAC
24	Rv-KI-T7TG-KmFRT-H2malT	GCGCGTTATCCGGCTAAACTTACACGCCGTACCCCATCATCTTCAGCAATTCT AGTTATTGCTCAGCAAGC
25	Fw-KI-pMB1CmFRT-H1LacY	ATGATATGTTGGTCGGATAAGGCGCTCGCGCCGCATCCGACATTGATTGCT GTAGGCTGGAGCTGCTTGG
26	Rv-KI-pMB1CmFRT-H2LacY	AATAACCGGGCAGGCCATGTCTGCCCGTATTTCGCGTAAGGAAATCCATTAT ATGCCGCCCTCGAGGTAC
27	Fw-KI-pMB1CmFRT-H1malT	TTCTGGCCGACCTTATAACC
28	Rv-KI-pMB1CmFRT-H2malT	CGTAACGCGGGTGAGCATTG
29	Fw-LacY-out	GTAAGCCTTCGCACATATCG
30	Rv-LacY-out	CCGGTCGCTACCATTACCAG
31	Fw-malT-out	CGTCATCGCTTGCATTAG
32	Rv-malT-out	CCCATCCCTCATGCCATCTG

* Restriction enzyme recognition site is underlined; N-terminus modification is dotted underlined; homology region is dashed underlined

Table 3 - List of fermented production strains

Strain #	Upstream	P450 Module	Construction	P450, N-terminus	CPR, N-terminus
Control	T7 MEP, T7 TG	N/A	N/A	N/A	
1	T7 MEP, T7 TG	Ch1 Trc	Chimera	8RP	
2	T7 MEP, T7 TG	Ch1 T7	Chimera	8RP	
3	T7 MEP, T7 TG	p5 Trc	Chimera	8RP	
4	T7 MEP, T7 TG	p10 Trc	Chimera	8RP	
5	T7 MEP, T7 TG	p5 T7	Chimera	8RP	
6	T7 MEP, T7 TG	p10 T7	Chimera	8RP	
7	T7 MEP, T7 TG	p5 Trc	Operon	8RP	8RP
8	T7 MEP, T7 TG	p10 Trc	Operon	8RP	8RP
9	T7 MEP, T7 TG	p5 T7	Operon	8RP	8RP
10	T7 MEP, T7 TG	p10 T7	Operon	8RP	8RP
11	T7 MEP, T7 TG	p5 Trc, Ch1 T7	Independent	8RP	8RP
12	T7 MEP, T7 TG	p5 Trc, T7	Independent	8RP	8RP
13	T7 MEP, T7 TG	p5 Trc	Operon	8RP	MA
14	T7 MEP, T7 TG	p5 Trc	Operon	8RP	2B1
15	T7 MEP, T7 TG	p5 Trc	Operon	MA	MA
16	T7 MEP, T7 TG	p5 Trc	Operon	MA	2B1
17	T7 MEP, T7 TG	p5 Trc	Operon	2B1	MA
18	T7 MEP, T7 TG	p5 Trc	Operon	2B1	2B1

Media Tables

Table 4 - Basal Hungate R-media Composition

Component	Amount (g/L)
KH_2PO_4 (monobasic)	13.3
$(NH_4)_2HPO_4$ (dibasic)	4.00
Citric Acid Monohydrate	1.70
Yeast Extract	5.00
HEPES	23.83

Table 5. R-media complete for "Hungate" small scale, 2 mL fermentations.

Component	Amount (mL/L)
Basal R-media	954.57
1 M MgSO4	4.88
0.1 M Fe(III) Citrate	2.45
Trace Element Solution	1.00
1 M Thiamine HCl	0.0134
32% (v/v) Glycerol	37.14
Antibiotic	1.00
IPTG	0.200

Table 6. Basal fermentation R-media composition.

Component	Amount (g/L)
KH_2PO_4 (monobasic)	13.3
$(NH_4)_2HPO_4$ (dibasic)	4.00
Citric Acid Monohydrate	1.70
Yeast Extract	5.00

Table 7. R-media complete for bioreactor.

Component	Amount (mL/L)
Basal R-media	928
1 M MgSO4	4.88
0.1 M Fe(III) Citrate	2.45
Trace Element Solution	1
1 M Thiamine HCl	0.0134
Antibiotic	1
R-media Pre-culture	50
60% v/v glycerol	12.66

Table 8- Trace element composition for both Hungate and bioreactor R-media. (Diluted 100x)

Component	Stock Concentration (g/L)
EDTA	8.4
H_3BO_3	3
$Zn(CH_3COO)_2$	8
$CoCl_2 \cdot 6H_2O$	4.6
$CuCl_2 \cdot 2H_2O$	1.9
$MnCl_2 \cdot 4H_2O$	24
$Na_2MoO_4 \cdot 2H_2O$	2.9

Table 9. Supplemented glycerol feeding solution composition.

Component	Amount (mL/L)
60% (v/v) Glycerol	784.99
Sterile DI H2O	130.5
1 M MgSO4	81.3
Trace Elements	1.6
0.1 M Fe(III) Citrate	0.96

References

1. Ajikumar, P. K. *et al.* Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in Escherichia coli. *Science* **330,** 70–74 (2010).
2. Mairhofer, J., Wittwer, A., Cserjan-Puschmann, M. & Striedner, G. Preventing T7 RNA Polymerase Read-through Transcription - A Synthetic Termination Signal Capable of Improving Bioprocess. *ACS Synth. Biol.* **4,** 265–273 (2015).
3. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97,** 6640–5 (2000).