# Definitive screening for accelerated Taxol biosynthetic pathway optimization and scale up in *Saccharomyces cerevisiae* cell factories

Laura E. Walls1,2,3, José L. Martinez3, E. Antonio del Rio Chanona4, Leonardo Rios-Solis1,2\*

1. *Institute for Bioengineering, School of Engineering, University of Edinburgh,Edinburgh, UK*
2. *Centre for Synthetic and Systems Biology (SynthSys), University of Edinburgh, Edinburgh, UK*
3. *Department of Biotechnology and Biomedicine, Section for Synthetic Biology, Technical University of Denmark, Kongens Lyngby, Denmark*
4. *Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, UK*

*\* Corresponding author: Leo.Rios@ed.ac.uk*

# Abstract

Recent technological advancements in synthetic and systems biology have enabled the construction of microbial cell factories expressing diverse heterologous pathways in unprecedentedly short time scales. However, the translation of such laboratory scale breakthroughs to industrial bioprocesses remains a major bottleneck. In this study, an accelerated bioprocess development approach was employed to optimize the biosynthetic pathway of the blockbuster chemotherapy drug, Taxol. Statistical design of experiments approaches were coupled with an industrially relevant high-throughput microbioreactor system to optimize production of key Taxol intermediates, Taxadien-5α-ol and Taxadien-5α-yl-acetate, in engineered yeast cell factories. The optimal factor combination was determined via data driven statistical modelling and validated in 1L bioreactors leading to a 2.1-fold improvement in taxane production compared to a typical defined media. Elucidation and mitigation of a nutrient limitation enhanced product titers a further two-fold and titers of the critical Taxol precursors, Taxadien-5α-ol and Taxadien-5α-yl-acetate were improved to 34 and 11 mg/L, representing a three-fold improvement compared to the highest literature titers in *S. cerevisiae*. Comparable titers were obtained when the process was scaled up a further five-fold using 5 L bioreactors. The results of this study highlight the benefits of a holistic design of experiments guided approach to expedite early stage bioprocess development.

**Keywords:** Definitive screening design, Bioprocess optimization, Saccharomyces cerevisiae, Taxol, High-throughput microbioreactor

# Introduction

The construction of microbial cell factories for the sustainable bioconversion of renewable feedstocks into complex natural products represents an auspicious alternative approach to traditional organic synthesis routes. Recent advances in synthetic and systems biology have accelerated progress towards this through enabling the construction of strains expressing wide ranging heterologous pathways in unprecedentedly short timescales.[1] However, the translation of such laboratory scale breakthroughs to industrial scale bioprocess remains a major challenge.[2] Despite lacking infrastructure for online monitoring and control of parameters such as pH and dissolved oxygen (DO), which are deemed critical to product quality at industrial scale, simple batch microtiter plate or shake flask cultivations are ubiquitously employed in the early stages of bioprocess development. In addition, the largely discredited intuition led one-factor-at-a-time (OFAT) approach is typically employed to guide screening experiments, further limiting process insight during this crucial developmental stage. [3–5]

Efficient, standardized methods for the detailed characterization of candidate strains under process relevant conditions are critical to alleviating key bottlenecks and accelerating bioprocess development.[6–8] Quality by design is a strategic approach, which aims to maximize process insight and understanding from the outset to minimize the risks associated with bioprocess scale-up. Statistical design of experiments (DoE), a key tool in the implementation of quality by design, provide a structured approach to exploring relationships between factors of interest and the measured response.[4] As factors are varied systematically and simultaneously, interaction effects, which are omnipresent in biological systems and neglected by OFAT approaches, can be estimated. The resulting data may be used to drive the derivation of statistical models to maximize process understanding and expedite the optimization process, whilst minimizing the experimental burden. Such DoE approaches can be readily coupled with recently developed advanced microbioreactor tools to facilitate high-throughput screening with dramatically increased process insight and closer mimicry of larger scale conditions compared to traditional microplates.[9,10] The microscale BioLector platform, for example, enables the online monitoring of critical process parameters such as pH, DO, fluorescence and cell growth for 48 simultaneous fermentations. Such data is extremely beneficial for bioprocess characterization, rendering the BioLector an excellent instrument for the execution of design of experiments studies. The platform has been successfully employed to screen the effect of media and processing conditions on recombinant β-carotene production in *Yarrowia lipolytica.*[11,12]Comparable trends were observed in the microscale platform and larger bench top bioreactor (0.5-1L) cultivations, highlighting scalability.

The densely functionalized and intricate structure of the highly effective chemotherapy drug, Taxol, renders it an excellent model compound for metabolic engineering studies. As a result, its microbial synthesis has been a major research focus for decades.[13–15] However, the pathway (Figure 1B) is yet to be fully elucidated and low and variable yields of early Taxol precursors have been achieved in microbial hosts to date. The highly promiscuous and multi-specific enzyme, taxadiene-5α-hydroxylase (CYP725A4), which is responsible for catalyzing the first oxidation step (Figure 1B), is a particular bottleneck. Both the activity and selectivity of the enzyme are highly sensitive to external conditions. Substantial differences in product spectra and overall titer have been observed in response to deviations in medium composition, pH, extraction solvent and production host.[16,17] The characterization of the superfluous side products generated during this initial oxidation step is desirable for future metabolic engineering efforts to improve enzyme selectivity. Nevertheless, current production titers are insufficient for purification, hindering such characterization. The development of an optimized and scalable bioprocess for the robust production of Taxol precursors and the corresponding side products is critical to accelerating pathway development.

Previous studies involving Taxol precursor production in *S. cerevisiae* have predominantly been performed in complex media.[17,18] Chemically defined media are inherently more desirable as they allow precise metabolic analysis of the production host and greater reproducibility compared to complex media.[19,20] Uncontrollable variables such as batch-batch variation in media can pose a major reproducibility challenge in complex media.[21] It was therefore hypothesized that the development of an optimized, chemically defined cultivation medium could be beneficial for Taxol precursor overproduction.

In this work, a DoE guided accelerated bioprocess development approach (Figure 1A) was employed with the aim of alleviating a key bottleneck in the biosynthetic pathway of the complex diterpenoid chemotherapy drug, Taxol. Statistical design of experiments were coupled with a high-throughput microbioreactor screening platform to optimize production of taxadien-5α-ol and the subsequent Taxol intermediate, taxadien-5α-yl-acetate, in yeast cell factories. Critical process parameters including pH and DO were monitored online for 48 simultaneous cultivations to enhance process insight. Statistical modelling was employed to identify and optimize factors significantly affecting productivity and the optimal combination was validated at 1L and 5L bioreactor scales under industrially relevant, controlled conditions.

# Materials and Methods

## Yeast strains

The *Saccharomyces cerevisiae* strains used in this study were *LRS5 (MATa, leu2‐3, 112::HIS3MX6‐GAL1p‐ERG19/GAL10p‐ERG8;ura3‐52::URA3‐GAL1p‐MvaSA110G/GAL10p‐MvaE [codon optimized]; his3Δ1::hphMX4‐GAL1p‐ERG12/GAL10p‐IDI1; trp1‐289::TRP1\_GAL1p‐CrtE(X.den)/GAL10p‐ERG20;YPRCdelta15::NatMX‐GAL1p‐CrtE(opt)/GAL10p‐CrtE; ARS1014::GAL1p‐TASY‐GFP; ARS1622b::GAL1p‐MBP‐TASY‐ERG20; ARS1114a::TDH3p‐MBP‐TASY‐ERG20)* and *LRS6* (*LRS5* ARS511b::GAL1p-CYP725A4-PGK1t/GAL3p-CPR-ENO2t; RKC3::GAL1p-TAT-CYC1t) as described previously,[17,18] originating from CEN.PK2-1C (EUROSCARF collection). All reagents were obtained from Fisher Scientific UK at the highest available purity unless otherwise stated.

## Experimental design and scoping

### Definitive screening design

Following the preliminary microscale cultivations (Section 2.3.1.), the CSM-URA was replaced with Yeast Synthetic Drop-out Medium Supplements without uracil (YSM, Sigma-Aldrich, UK). Uracil was added at a concentration of 40 mg/g YSM. The effect of six factors: initial OD600 along with the initial concentrations of galactose, Yeast Nitrogen Base without Amino Acids and Ammonium Sulphate (YNB), ammonium sulphate, additional MgSO4 and YSM) on *LRS6* productivity was investigated. A three-level definitive screening design (DSD) was selected to investigate the effect of the six factors on *LRS6* productivity as shown in Table 1.

### Shake flask scoping experiment

The *LRS6* strain was cultivated in 250 mL Erlenmeyer flasks for 3 days. Inocula preparation was achieved by transferring single colonies to 5 mL of rich YPD medium (1% yeast extract; 2% peptone; and 2% glucose) and incubating at 30 °C and 250 rpm overnight. The scoping trial involved the following three media compositions:

1. Least forcing: all factors set to their lowest level – Galactose 10 g/L, YNB 1.7 g/L, Nitrogen 1 g/L, MgSO4 0 g/L, YSM 2 g/L, Initial OD600 = 0.1
2. Centre point: all factors set halfway between the low and high range – Galactose 30 g/L, YNB 2.55 g/L, Nitrogen 4.5 g/L, MgSO4 0.3 g/L, YSM 3 g/L, Initial OD600 = 0.55
3. Most forcing: all factors set to their highest level – Galactose 50 g/L, YNB 3.4 g/L, Nitrogen 10 g/L, MgSO4 0.6 g/L, YSM 4 g/L, Initial OD600 = 1

Each condition was tested in duplicate 250 mL flasks, each with a culture volume of 20 mL. Uracil was added at a concentration of 40 mg/g YSM. A 5 mL dodecane overlay was also added giving a final working volume of 25 mL. Taxane production was analyzed via gas chromatography‐mass spectrometry (GC‐MS) at the end of the cultivation.

## Microbioreactor cultivation

Microscale cultivations were performed using a BioLector Pro (mp2‐labs) microbioreactor‐screening platform. Inocula were prepared as described for the shake flasks. The temperature was maintained at 30 °C under agitation of 1000 rpm with a shaking diameter of 3 mm in 48‐well FlowerPlates (mp2‐labs). Temperature, biomass, dissolved oxygen (DO), and pH were monitored online using the inbuilt optical sensors. Taxane production was analyzed via GC‐MS at the end of the cultivation.

### Preliminary scoping trials

Preliminary cultivation experiments were performed in the BioLector FlowerPlates using a typical synthetic defined galactose (SDG) medium (Yeast Nitrogen Base without Amino Acids, 6.7 g/L; Complete Supplement Mixture minus Uracil (CSM-URA), 0.77 g/L; Uracil, 20 mg/L; galactose, 20 g/L). Control cultivations were performed with YPG (yeast extract, 10g/L; peptone, 20 g/L; galactose 20 g/L). Aliquots of the preculture were diluted to give an 800 µL culture with an initial OD600 = 1. A 200‐µL dodecane overlay was also added to each well giving a total working volume of 1 mL.

### High-throughput screening

For the high-throughput screening experiments, aliquots of the preculture were diluted with each of the medium combinations indicated in Table 1 to give 800 μL cultures with the appropriate initial OD600. A 200 µL dodecane overlay was also added to each well giving a total working volume of 1 mL. Each factor combination was tested in duplicate or triplicate in a 48-well FlowerPlate.

## Bioreactor cultivation

## Larger scale cultivations were conducted in 1  L and 5 L BIOSTAT Q plus bioreactors (Sartorius‐Stedim Biotech S.A.) with working volumes of 500 mL and 2.5 L, respectively. Pre-inoculum cultures were prepared by transferring from a single colony to 5 mL of YPD and incubating at 30 ℃ and 250 rpm for 24 hours. The resulting culture was subsequently used to inoculate a secondary 50 mL culture to an OD600 = 1 and incubated overnight. An aliquot of the resulting culture was diluted with the optimized medium to give a culture with an initial OD600 = 1.

Antifoam 204 (Sigma Aldrich, Denmark) was added to prevent excess foam production. A 20 % dodecane overlay was also added. Temperature, DO, and pH were measured online. MFCS software (version 3.0, Sartorius‐Stedim Biotech S.A.) was employed to control the cultivation, pH was maintained at six through the automatic addition of 2 M NaOH or 2 M H2SO4, and temperature was maintained at 30 ℃. A constant airflow of 1 vvm was maintained and stirrer speed was adjusted manually to maintain DO above 30 %. Off‐gas analysis was performed online via mass spectrometry (Prima Pro, Thermo Fisher Scientific). Samples were taken twice daily for taxane and metabolite quantification via GC‐MS and high‐performance liquid chromatography (HPLC).

## Taxane and metabolite identification and quantification

Taxane identification and quantification was achieved via GC‐MS. The organic dodecane layer was separated from the culture medium through centrifugation and a 1‐μl sample was injected into a TRACE™ 1300 Gas Chromatograph (Thermo Fisher Scientific) coupled to an ISQ LT single quadrupole mass spectrometer (Thermo Fisher Scientific). Chromatographic separation was achieved using a Trace Gold TG‐SQC gas chromatography column using a previously described method.[22] To identify and quantify the production of compounds by *LRS5* and *LRS6*, pure standards of taxadiene, kindly supplied by Baran Lab (The Scripps Research Institute) and GGOH, obtained from Sigma Aldrich (UK), were used. Additional product concentrations were estimated relative to standard taxadiene concentrations. In the bioreactor cultivation, ethanol, acetate, and glycerol production were analyzed via ion‐exchange HPLC. Following filtration using a 0.45‐μm filter, 20 μl samples were injected into a Bio‐Rad Aminex HPX‐87H column for analysis. The eluent was 5 mM H2SO4, flowrate 0.6 mL/min, and the temperature was 60 °C. A RID‐detector was used for quantification.

## Statistical analysis

Design of experiments and statistical modelling were performed using JMP Pro 15 statistical software. Forward stepwise regression with a p-value to enter of 0.1 was performed using JMP to fit the regression model for the definitive screening design. A full quadratic analysis was performed, thereby considering all main effects and any second order interactions. The resulting model contained those factors which produced a significant effect on the response. The prediction profiler was then used to optimize the settings of these factors.

# Results and discussion

## Preliminary microscale growth experiments

Previous studies at micro (1 mL) and bioreactor (0.5-1 L) scale involving the cultivation of *S. cerevisiae* strain, *LRS6,* which has been engineered for heterologous expression of the first three genes of the Taxol biosynthetic pathway (Figure 1B), were performed exclusively in complex media.[17] In order to characterize performance in defined media preliminary microscale growth experiments were performed using a typical synthetic defined media (SDG, Section 2.3.1). Galactose was selected as the carbon source to maximize expression of the heterologous pathway, which was under the control of galactose inducible promoters (Figure 1B) Control experiments using a typical complex media (YPG, Section 2.3.1.) were included for comparison. The results of this investigation are summarized in Figure 2.

The initial pH of the SDG media was significantly lower than that of YPG at around 5.2 compared to 6.5. Previous studies revealed that the CYP725A4 is sensitive to deviations in external pH,[17] indicating pH adjustment may be beneficial in future runs. CYP725A4 has an optimal pH of 7.2 [23] and little activity was observed in cultivations where the pH dropped below 5.[17] However, as *LRS6* growth has been found to be dramatically reduced at neutral pH,[17] a pH of 6 was deemed adequate for this process to maximize productivity. After 48 hours of cultivation the OD600 values were 31.2 ± 4.0 and 19.9 ± 2.0 for the YPG and SDG cultures respectively. Although SDG was supplemented with a source of amino acids and vitamins, the complex YPD media is a much richer source. The reduced nutrient availability likely contributed the reduced biomass accumulation observed. This is consistent with previous studies which demonstrated that the growth rate of *S. cerevisiae*  is reduced in standard defined media compared to complex YPD.[24] Total taxane yields were 81 ± 17 and 46 ± 8 for the YPG and SDG cultures, respectively. The lower taxane titer observed in the SDG cultures was expected as a strong positive correlation has been elucidated between biomass and taxane accumulation.[17,18]

## Definitive screening design development and validation

In order to improve taxane productivity, a definitive screening design (DSD) was developed to screen a range of defined media compositions. Using a synthetic complete media as a basis, a range of media compositions were screened to determine the optimal for taxane accumulation. Preliminary studies revealed a significantly reduced growth rate in synthetic compared to complex media (Figure 2A). As the reduced amino acid and nutrient availability likely contributed, an alternative supplement mixture (YSM, Sigma-Aldrich, UK) with a richer supply of amino acids was used and its concentration along with that of yeast nitrogen base were included as factors in the DSD. As the critical taxadiene synthase enzyme relies on the co-factor Mg2+ for activation, it was hypothesized that supplementing the media with an additional magnesium supply may improve activity. The highest reported oxygenated taxane yields reported in *E. coli* were achieved using an optimized media supplemented with around 0.6 g/L MgSO4.[25] This concentration was therefore selected for this optimization study. As taxane accumulation has been found to be strongly correlated to *LRS6* growth,[17]it was hypothesized that the initial OD600 of the cultivation may influence taxane accumulation and initial OD600 was therefore included as a factor. The carbon to nitrogen (C/N) ratio of the growth media can have a major influence on cellular metabolism. It is widely accepted that increasing the C/N ratio has a positive effect on lipid accumulation by oleaginous yeast;[26] however, the effect on isoprenoid production is more ambiguous.[27] Like lipids, isoprenoids rely on acetyl-CoA as a universal precursor. In addition, as hydrophobic isoprenoids such as the triterpene, squalene[28] and carotene, β-carotene[29] have been found to accumulate in cytoplasmic lipid particles, increasing the C/N ratio may be beneficial. In a study by Olson et al., 2016, β-carotene production was found to increase up to fourfold when the C/N ratio was increased 5.7-fold in an engineered *S. cerevisiae* strain. Similarly, increasing lipid accumulation has been found to improve lycopene accumulation in *S. cerevisiae.*[30] The concentrations of the carbon (galactose) and nitrogen (ammonium sulphate) were therefore deemed important factors in this study. A six factor, three-level DSD was created as summarized in Table 1 (Section 2.2.1). In order to validate the design a scoping experiment was performed in shake flasks. This involved three conditions in which, the levels of each factor were set to their low (least forcing), medium (center point) and high (most forcing) levels, respectively, as described in section 2.2.2. The results of the scoping trial are summarized in Figure 3.

The least and most forcing conditions gave rise to the lowest and highest taxane titers, respectively. The center point titer was between that of the least and most forcing condition (Figure 3A). In addition, there was little variation between the duplicate shake flasks, the experimental factors were therefore deemed appropriate. The use of defined media resulted in substantial changes in product spectra in comparison to complex media (Figure 3C). The titer of diterpenoid 1, an isomer of the first oxygenated Taxol intermediate, taxadien-5α-ol (T5αol), which was the major oxygenated product of the strain in complex media investigations,[17] was negligible. The main oxygenated products were OCT and Iso-OCT. Relative production of the di-oxygenated taxadiendiol compound was enhanced, and additional products were observed at 8.67 and 9.70 minutes as shown in Figure 3C and Supplementary Figure 1. Although the mass spectra corresponding to these peaks was similar to the other oxygenated compounds it is possible that they could be endogenous products. In order to determine whether the compounds were products of the heterologous CYP725A4 or TAT enzymes rather than endogenous gene products a further experiment was performed. The *LRS5* strain which expresses TASY alone was also cultivated in the most forcing media as a control. The resulting gas chromatogram is shown in Figure 3B. The product spectra for *LRS5* grown in the most forcing synthetic media was very similar to that observed in complex media.[17,18] Taxadiene was the main product of the strain and the additional products observed in the *LRS6* cultures were not observed. This indicates that the additional peaks likely resulted from CYP725A4 or TAT enzymatic activity or products.

## Defined medium optimization using definitive screening design of experiments

Following validation, the DSD outlined in Table 1 was implemented at microscale using the BioLector microbioreactor screening platform. The key response variable, total taxane accumulation was evaluated at the end of the cultivation as summarized in Figure 4B.

Of the 17 conditions tested, factor combinations six, nine and eleven gave rise to the highest total taxane titers of 155 ± 3, 194 ± 46 and 145 ± 12 mg/L, respectively. The product profile obtained for these factor combinations were highly similar to those observed in the most forcing shake flask cultivations (Figure 3C). The effect of each of the six factors on taxane production was evaluated via forward stepwise regression using JMP with a p-value to enter of 0.1. A full quadratic model was derived, thereby considering all main effects and any second order interactions. The resulting statistical model revealed initial OD600 (p = 6 x 10-6) along with initial YNB (p = 0.021) and galactose (p = 0.011) concentrations were significant main effects. According to the statistical model the optimal galactose, YNB and initial OD600 settings were 50 g/L, 3.4 g/L and 1, respectively (Supplementary Figure 2). As the concentrations of ammonium sulphate (p = 0.285), YSM (p = 0.990) and MgSO4 (p = 0.690) did not significantly affect productivity, the lowest values of 1, 2 and 0 g/L were used in subsequent cultivations to conserve resources.

The final titer of the key products, iso-taxadiene, taxadiene, iso-OCT, diterpenoid 1, OCT and T5αol were plotted for each of the 17 factor combinations as shown in Figure 4D. The observed product spectra in the high yielding defined media (factor combinations 6, 9, 11, 11 and 12; Figure 4D) were highly different to those observed in the YPG control (Figure 2) and previous cultivations with complex media in BioLector FlowerPlates.[17] CYP725A4 activity and selectivity has been previously found to be highly sensitive to media composition. Edgar et al., 2016, reported a relative increase in iso-OCT production in minimal media compared to rich media when the pathway was expressed in *S. cerevisiae* and *Yarrowia lipolytica*. Differences in media composition may have therefore contributed to the deviations in product spectra observed in this study. As in the shake flask cultivations (Figure 3C), the major oxygenated products were OCT and Iso-OCT along with smaller quantities of diterpenoid 1 and T5αol. The novel diterpenoids with retention times of 8.67 and 9.70 minutes were also detected along with two additional peaks at 10.23 and 10.57 minutes (Supplementary Figure 3). The diagnostic T5αAc peaks at m/z 287 (P+ - CH3CO), 270 (P+ - CH3COOH) and 255 (P+-CH3COOHCH3) [14] were observed in the mass spectra of the peak at 9.43 (Supplementary Figure 4). However, a large peak at 245 was also detected along with peaks at m/z 304 and 286 (P+-H2O), which are characteristic of taxadiendiol. Therefore, although T5αAc was likely produced in the BioLector cultivations, accurate quantification was not possible due to co-elution with the di-oxygenated compound.[17]

Although the initial pH of the cultivation medium was adjusted to pH 6, acidification of the cultivation medium was observed (Figure 4C). The rate of acidification was greater in the higher yielding treatments (6,9,11,12), with a rapid decrease in pH to between 5.1 and 5.3 observed in the first 30 hours of cultivation. This was considerably lower than the minimum pH of 5.98 ± 0.04 detected in the standard YPG cultivation (Figure 2). Metabolism of galactose by *LRS6* is respiratory-fermentative, during fermentative growth the yeast acidifies the cultivation medium through a combination of proton pumping and organic acid secretion.[31] As the initial concentrations of galactose and biomass were at the high level of 50 g/L and OD600 = 1 for factor combinations 6, 9, 11 and 12, the rate of fermentation was likely greater in these cultivations resulting in an increased rate of acidification. The rate of acidification was much lower during the lower yielding treatments (4, 5, 13, 14, 15, 17), likely due to the reduced growth and fermentation rates, with the minimum pH observed at the end of the 72 hour cultivation. Interestingly, diterpenoid 1 was the sole oxygenated product detected in these cultivations (4, 5, 13, 14, 15, 17; Figure 4D and Supplementary Figure 5A). In addition, the buffering capacity of yeast extract peptone medium has been found to be greater than that of defined medium supplemented with amino acids [32], which may have contributed to the increased acidification. The *CYP725A4* enzyme has been previously found to be highly sensitive to deviations in pH. A dramatic reduction in enzyme activity was observed when the pH of the cultivation medium dropped to 4.9 during an uncontrolled bioreactor cultivation of the strain.[17] The fluctuations in the external pH of the cultivation medium may have therefore contributed to the deviations in enzyme selectivity observed in defined medium at microscale. Between 30 and 72 hours a gradual increase in pH toward the initial pH of 6 was observed for treatments 6, 9, 11 and 12. This was likely due to the assimilation of less favorable carbon and nitrogen sources in the later stages of the cultivation.

## Scale up of optimized defined media in 1 L bioreactors

The optimal conditions elucidated at microscale (Supplementary Figure 2) were subsequently scaled up using 1 L bioreactors. A control was also ran with a standard defined media similar to the SDG media (Figure 2), except the CSM supplement was replaced with the more nutrient rich YSM. The results of this investigation are summarized in Figure 5.

At bioreactor scale the total taxane titer was 164 mg/L, this represented a 2.1-fold improvement in titer compared to the control (Figure 5F) and was just 13% lower than that predicted by the statistical model (Supplementary Figure 2) derived from the microscale cultivation data. However, the maximum OD600 value was just 22.5, significantly lower than those obtained in the higher yielding BioLector cultivations (Figure 4). Biomass accumulation (Figure 5A and C) and CO2 production (Figure 5E) were similar for the control and optimized cultivations during the first 24 hours. However, after 24 hours a drop in CO2 production was observed in the optimized cultivation (Figure 5E) despite excess galactose availability (Figure 5A), this was indicative of nutrient limitation. As ammonium sulphate concentration did not have a significant effect on taxane accumulation at microscale (p = 0.285), the concentration was set to the low level of 1 g/L at this scale. As all other nutrients were in excess it was hypothesized that nitrogen limitation may have been responsible for the reduced growth rate observed. A second fermentation was therefore performed with an increased initial ammonium sulphate as summarized in Figure 6.

Through increasing the initial ammonium concentration, the nutrient limitation was effectively eliminated. Biomass accumulation was improved fourfold with a maximum OD600 of 84 at the end of the 120 hour cultivation (Figure 6A). Total taxane production was also improved 1.8-fold, with a maximum titer of 302 mg/L obtained at 54 hours (Figure 6B). At 54 hours the galactose had been depleted and the yeast metabolized the produced ethanol in the subsequent 55 hours of cultivation (Figure 6A). As the heterologous pathway was expressed under the control of galactose inducible promoters, no additional taxanes were produced during growth on ethanol and the total titer declined to around 218 mg/L at the end of the cultivation. As this drop in titer was proportional for each of the key products (Figure 6C), it may have been due to loss via air stripping. Although the addition of a dodecane overlay has been found to reduce dramatically reduce air-stripping of volatile taxane compounds, with the relatively high air flow rate (1 vvm) and long cultivation time in the bioreactor, its application may not have been sufficient to eliminate the phenomenon entirely.

Interestingly, the nature of the taxane products generated in the 1L bioreactors (Supplementary Figure 6) was highly similar to those obtained for *LRS6* grown in complex media in 1L BIOSTAT reactors.[17] The major product of CYP725A4 was diterpenoid 1 under each of the three conditions tested in the bioreactor (Figure 5B and D, Figure 6C). Under the optimized condition with increased ammonium sulphate a maximum diterpenoid 1 titer of 73 mg/L was achieved (Figure 6C). The confirmed Taxol intermediates taxadiene, iso-taxadiene, T5αol and T5αAc were also detected with maximum titers of 63, 6, 35 and 12 mg/L, respectively. This represented 1.5 and 3.2-fold improvements in titers of the critical Taxol precursors, T5αol and T5αAc, compared to the highest reported literature titers in *S. cerevisiae.*[17] OCT and iso-OCT were also produced, however, the additional compounds observed in the BioLector and shake flask cultivations were not detected. A further diterpenoid compound, diterpenoid 4 (Supplementary Figure 6), which was previously produced by the strain in bioreactor cultivations with complex media[17] was also detected. The differences in product spectra observed in the smaller scale cultivations were therefore unlikely to be solely the result of the differences in medium composition. The acidification of the cultivation medium in the absence of pH control at smaller scale likely contributed to the deviations in selectivity.

## Further scale up of the optimized defined media cultivation in 5 L bioreactor

To further investigate the scalability of the optimized 1L process, a further fivefold scale up was performed using a 5L bioreactor. The results of this experiment are summarized in Figure 7.

At 5 L scale the lag phase of growth was increased to around 72 hours (Figure 7A). Despite this, the final OD600 was 87.5, highly comparable to the final OD600 of 84 obtained in the 1 L bioreactor cultivation (Figure 6A). A maximum total taxane titer of 270 mg/L was observed at 126 hours (Figure 7B), just 10 % lower than the maximum titer of 302 mg/L obtained at 1 L scale (Figure 6B). A slight decrease in titer to 260 mg/L was observed in the final 17 hours of the cultivation. The investigated process was demonstrated to be robust despite a 2500-fold scale up in cultivation volume from micro- to 5 L bioreactor scale. To our knowledge this represents the greatest scale-up in production of oxygenated and acetylated taxanes by yeast microbial cell factories.

# Concluding remarks

In this study, a design of experiments guided approach was applied to develop an optimized bioprocess to produce Taxol precursors using an engineered *S. cerevisiae* strain. A definitive screening design elucidated key factors affecting taxane accumulation in the strain at microscale, facilitating statistical model derivation. The optimal factor combination predicted by the statistical model was validated in highly instrumented 1 L bioreactors with comparable total taxane titers of 164 mg/L and 189 mg/L achieved. This titer was 2.1-fold higher than that in the parallel control cultivation in standard defined media. However, the final OD600 was just 19.0, 2.6-fold lower than the highest yielding microscale cultivation and closer inspection of the offgas data revealed nitrogen limitation. Mitigation of this nutrient limitation led to maximum total and oxygenated taxane titers of 302 and 197 mg/L, respectively. This represented 1.8 and 2.5-fold improvements compared to the highest reported titers for yeast, achieved in rich media. A further fivefold scale up of this optimized process was performed using 5 L bioreactors. At this scale a final OD600 of 87.5 and maximum total taxane titer of 270 mg/L was achieved, highly comparable to the 84 and 302 mg/L obtained at 1 L scale. The ability of strategic quality by design approaches to dramatically improve the efficiency of the early phases of bioprocess development was demonstrated in this study. Production of critical oxygenated and acetylated Taxol intermediates was successfully scaled up 2500-fold in yeast cell factories. Future work should focus on the incorporation of such methods during microbial cell factory construction and optimization to ensure robust growth under industrially relevant conditions and further reduce the risks associated with bioprocess scale up.

# Acknowledgements

The authors would like to thank Ms. Caroline Delahoyde and Mr. Stuart Martin at The School of Engineering, University of Edinburgh, UK and Dr Aaron John Christian Andersen and Mette Amfelt at the DTU Metabolomics Core, DTU, Denmark for their kind assistance and technical support with GC-MS analysis. Thanks to Tina Johansen at the DTU Fermentation Core, DTU, Denmark for her assistance with the BIOSTAT cultivations and HPLC analyses. Thanks to Professor Phil Baran's Lab at The Scripps Research Institute, San Diego, California for providing the taxadiene standard. Thanks to the Edinburgh Genome Foundry at the University of Edinburgh for providing access to the BioLector platform. This work was supported by the Engineering and Physical Sciences Research Council (Grant number EP/R513209/1), the Royal Society (Grant Number RSG\R1\180345), The British Council (Grant Number: 527429894) and the Novo Nordisk Foundation within the frame of the Fermentation Based Biomanufacturing initiative (Grant Number: NNF17SA0031362).

# Conflicts of interest

The authors declare no conflict of interest.

# Author contributions

Laura E. Walls designed and performed the experiments and all of the data analysis. Laura E. Walls wrote the manuscript with input from all authors. E. Antonio del Rio Chanona provided support with design of experiments, statistical modelling and reviewing the manuscript. Leonardo Rios-Solis and José L. Martinez supervised and coordinated the study.

# References

[1] Malcı, K., Walls, L. E., & Rios-Solis, L. (2020). Multiplex Genome Engineering Methods for Yeast Cell Factory Development . In *Frontiers in Bioengineering and Biotechnology*  (Vol. 8, p. 1264). https://www.frontiersin.org/article/10.3389/fbioe.2020.589468

[2] Chubukov, V., Mukhopadhyay, A., Petzold, C. J., Keasling, J. D., & Martín, H. G. (2016). Synthetic and systems biology for microbial production of commodity chemicals. *Npj Systems Biology and Applications*, *2*(1), 16009. https://doi.org/10.1038/npjsba.2016.9

[3] Wilk, P., Halim, M., & Rios-Solis, L. (2019). Recent Advances and Impacts of Microtiter Plate-Based Fermentations in Synthetic Biology and Bioprocess Development. In E. M. T. El-Mansi, J. Nielsen, D. Mousdale, & R. P. Carlson (Eds.), *Fermentation Microbiology and Biotechnology, Fourth Edition* (Fourth, p. 14). CRC Press. https://doi.org/https://doi.org/10.1201/9780429506987

[4] Gilman, J., Walls, L., Bandiera, L., & Menolascina, F. (2021). Statistical Design of Experiments for Synthetic Biology. *ACS Synthetic Biology*, *10*(1), 1–18. https://doi.org/10.1021/acssynbio.0c00385

[5] Fellermann, H., Shirt-Ediss, B., Kozyra, J., Linsley, M., Lendrem, D., Isaacs, J., & Howard, T. (2019). Design of experiments and the virtual PCR simulator: An online game for pharmaceutical scientists and biotechnologists. *Pharmaceutical Statistics*, *18*(4), 402–406. https://doi.org/10.1002/pst.1932

[6] Campbell, K., Xia, J., & Nielsen, J. (2017). The Impact of Systems Biology on Bioprocessing. *Trends in Biotechnology*, *35*(12), 1156–1168. https://doi.org/10.1016/j.tibtech.2017.08.011

[7] Hemmerich, J., Noack, S., Wiechert, W., & Oldiges, M. (2018). Microbioreactor Systems for Accelerated Bioprocess Development. *Biotechnology Journal*, *13*(4), 1700141. https://doi.org/https://doi.org/10.1002/biot.201700141

[8] Fink, M., Cserjan-Puschmann, M., Reinisch, D., & Striedner, G. (2021). High-throughput microbioreactor provides a capable tool for early stage bioprocess development. *Scientific Reports*, *11*(1), 2056. https://doi.org/10.1038/s41598-021-81633-6

[9] Hsu, W.-T., Aulakh, R. P. S., Traul, D. L., & Yuk, I. H. (2012). Advanced microscale bioreactor system: a representative scale-down model for bench-top bioreactors. *Cytotechnology*, *64*(6), 667–678. https://doi.org/10.1007/s10616-012-9446-1

[10] Funke, M., Diederichs, S., Kensy, F., Müller, C., & Büchs, J. (2009). The baffled microtiter plate: Increased oxygen transfer and improved online monitoring in small scale fermentations. *Biotechnology and Bioengineering*, *103*(6), 1118–1128. https://doi.org/10.1002/bit.22341

[11] Jacobsen, I. H., Ledesma-Amaro, R., & Martinez, J. L. (2020). Recombinant β-Carotene Production by Yarrowia lipolytica – Assessing the Potential of Micro-Scale Fermentation Analysis in Cell Factory Design and Bioreaction Optimization . In *Frontiers in Bioengineering and Biotechnology*  (Vol. 8, p. 29). https://www.frontiersin.org/article/10.3389/fbioe.2020.00029

[12] Back, A., Rossignol, T., Krier, F., Nicaud, J.-M., & Dhulster, P. (2016). High-throughput fermentation screening for the yeast Yarrowia lipolytica with real-time monitoring of biomass and lipid production. *Microbial Cell Factories*, *15*(1), 147. https://doi.org/10.1186/s12934-016-0546-z

[13] Huang, K., Huang, Q., Wildung, M. R., Croteau, R., & Scott, A. I. (1998). Overproduction, inEscherichia coli,of Soluble Taxadiene Synthase, a Key Enzyme in the Taxol Biosynthetic Pathway. *Protein Expression and Purification*, *13*(1), 90–96. https://doi.org/https://doi.org/10.1006/prep.1998.0870

[14] Walker, K., Ketchum, R. E. B., Hezari, M., Gatfield, D., Goleniowski, M., Barthol, A., & Croteau, R. (1999). Partial Purification and Characterization of Acetyl Coenzyme A: Taxa-4(20),11(12)-dien-5α-olO-Acetyl Transferase That Catalyzes the First Acylation Step of Taxol Biosynthesis. *Archives of Biochemistry and Biophysics*, *364*(2), 273–279. https://doi.org/https://doi.org/10.1006/abbi.1999.1125

[15] McElroy, C., & Jennewein, S. (2018). Taxol® Biosynthesis and Production: From Forests to Fermenters. In W. Schwab, B. M. Lange, & M. Wüst (Eds.), *Biotechnology of Natural Products* (pp. 145–185). Springer International Publishing. https://doi.org/10.1007/978-3-319-67903-7\_7

[16] Edgar, S., Zhou, K., Qiao, K., King, J. R., Simpson, J. H., & Stephanopoulos, G. (2016). Mechanistic Insights into Taxadiene Epoxidation by Taxadiene-5α-Hydroxylase. *ACS Chemical Biology*, *11*(2), 460–469. https://doi.org/10.1021/acschembio.5b00767

[17] Walls, L. E., Malcı, K., Nowrouzi, B., Li, R. A., d’Espaux, L., Wong, J., Dennis, J. A., Semião, A. J. C., Wallace, S., Martinez, J. L., Keasling, J. D., & Rios-Solis, L. (2020). Optimizing the biosynthesis of oxygenated and acetylated Taxol precursors in Saccharomyces cerevisiae using advanced bioprocessing strategies. *Biotechnology and Bioengineering*, *n/a*(n/a). https://doi.org/10.1002/bit.27569

[18] Nowrouzi, B., Li, R. A., Walls, L. E., d’Espaux, L., Malcı, K., Lungang, L., Borrego, N. J., Escalera, A., Morones-Ramirez, J. R., Keasling, J., & Rios-Solis, L. (2020). Enhanced production of taxadiene in Saccharomyces cerevisiae. *Microbial Cell Factories*, *19*(200). https://doi.org/10.1186/s12934-020-01458-2

[19] Lee, S. Y., & Kim, H. U. (2015). Systems strategies for developing industrial microbial strains. *Nature Biotechnology*, *33*(10), 1061–1072. https://doi.org/10.1038/nbt.3365

[20] Rangel, A. E. T., Gómez Ramírez, J. M., & González Barrios, A. F. (2020). From industrial by-products to value-added compounds: the design of efficient microbial cell factories by coupling systems metabolic engineering and bioprocesses. *Biofuels, Bioproducts and Biorefining*, *14*(6), 1228–1238. https://doi.org/https://doi.org/10.1002/bbb.2127

[21] Zhang, J., Reddy, J., Buckland, B., & Greasham, R. (2003). Toward consistent and productive complex media for industrial fermentations: Studies on yeast extract for a recombinant yeast fermentation process. *Biotechnology and Bioengineering*, *82*(6), 640–652. https://doi.org/10.1002/bit.10608

[22] Reider Apel, A., Sachs, D., Tong, G. J., d’Espaux, L., Wehrs, M., Garber, M., Nnadi, O., Mukhopadhyay, A., Keasling, J. D., Hillson, N. J., Li, R. A., & Zhuang, W. (2016). A Cas9-based toolkit to program gene expression in Saccharomyces cerevisiae. *Nucleic Acids Research*, *45*(1), 496–508. https://doi.org/10.1093/nar/gkw1023

[23] Hefner, J., Rubenstein, S. M., Ketchum, R. E. B., Gibson, D. M., Williams, R. M., & Croteau, R. (1996). Cytochrome P450-catalyzed hydroxylation of taxa-4(5),11(12)-diene to taxa-4(20),11(12)-dien-5a-o1: the first oxygenation step in taxol biosynthesis. *Chemistry & Biology*, *3*(6), 479–489. https://doi.org/https://doi.org/10.1016/S1074-5521(96)90096-4

[24] Roberts, T. M., Kaltenbach, H.-M., & Rudolf, F. (2020). Development and optimisation of a defined high cell density yeast medium. *Yeast*, *37*(5–6), 336–347. https://doi.org/https://doi.org/10.1002/yea.3464

[25] Biggs, B. W., Rouck, J. E., Kambalyal, A., Arnold, W., Lim, C. G., Mey, M. De, Neil-johnson, M. O., Starks, C. M., Das, A., & Ajikumar, P. K. (2016). Orthogonal Assays Clarify the Oxidative Biochemistry of Taxol P450 CYP725A4. *ACS Chemical Biology*, *11*, 1445–1451. https://doi.org/10.1021/acschembio.5b00968

[26] Beopoulos, A., Nicaud, J.-M., & Gaillardin, C. (2011). An overview of lipid metabolism in yeasts and its impact on biotechnological processes. *Applied Microbiology and Biotechnology*, *90*(4), 1193–1206. https://doi.org/10.1007/s00253-011-3212-8

[27] Braunwald, T., Schwemmlein, L., Graeff-Hönninger, S., French, W. T., Hernandez, R., Holmes, W. E., & Claupein, W. (2013). Effect of different C/N ratios on carotenoid and lipid production by Rhodotorula glutinis. *Applied Microbiology and Biotechnology*, *97*(14), 6581–6588. https://doi.org/10.1007/s00253-013-5005-8

[28] Spanova, M., Czabany, T., Zellnig, G., Leitner, E., Hapala, I., & Daum, G. (2010). Effect of Lipid Particle Biogenesis on the Subcellular Distribution of Squalene in the Yeast Saccharomyces cerevisiae. *Journal of Biological Chemistry* , *285*(9), 6127–6133. https://doi.org/10.1074/jbc.M109.074229

[29] Olson, M. L., Johnson, J., Carswell, W. F., Reyes, L. H., Senger, R. S., & Kao, K. C. (2016). Characterization of an evolved carotenoids hyper-producer of Saccharomyces cerevisiae through bioreactor parameter optimization and Raman spectroscopy. *Journal of Industrial Microbiology & Biotechnology*, *43*(10), 1355–1363. https://doi.org/10.1007/s10295-016-1808-9

[30] Ma, T., Shi, B., Ye, Z., Li, X., Liu, M., Chen, Y., Xia, J., Nielsen, J., Deng, Z., & Liu, T. (2019). Lipid engineering combined with systematic metabolic engineering of Saccharomyces cerevisiae for high-yield production of lycopene. *Metabolic Engineering*, *52*, 134–142. https://doi.org/https://doi.org/10.1016/j.ymben.2018.11.009

[31] Walker, M. G., & Stewart, G. G. (2016). Saccharomyces cerevisiae in the Production of Fermented Beverages. In *Beverages*  (Vol. 2, Issue 4). https://doi.org/10.3390/beverages2040030

[32] Thomas, K. C., Hynes, S. H., & Ingledew, W. M. (2002). Influence of medium buffering capacity on inhibition of Saccharomyces cerevisiae growth by acetic and lactic acids. *Applied and Environmental Microbiology*, *68*(4), 1616–1623. https://doi.org/10.1128/aem.68.4.1616-1623.2002

Tables

**Table 1: DSD experimental design for defined media optimization**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Medium Number | Galactose (g/L) | YNB (g/L) | Ammonium Sulphate (g/L) | MgSO4 (g/L) | Initial OD600 | YSM (g/L) |
| 1 | 10 | 3.40 | 10.0 | 0.6 | 0.55 | 2 |
| 2 | 30 | 2.55 | 5.5 | 0.3 | 0.55 | 3 |
| 3 | 10 | 3.40 | 10.0 | 0.0 | 1.00 | 4 |
| 4 | 50 | 3.40 | 10.0 | 0.0 | 0.10 | 2 |
| 5 | 30 | 1.70 | 10.0 | 0.0 | 0.10 | 4 |
| 6 | 50 | 2.55 | 10.0 | 0.6 | 1.00 | 4 |
| 7 | 10 | 1.70 | 5.5 | 0.0 | 1.00 | 2 |
| 8 | 10 | 1.70 | 1.0 | 0.6 | 1.00 | 4 |
| 9 | 50 | 3.40 | 1.0 | 0.0 | 1.00 | 3 |
| 10 | 50 | 3.40 | 5.5 | 0.6 | 0.10 | 4 |
| 11 | 50 | 3.40 | 1.0 | 0.6 | 1.00 | 2 |
| 12 | 50 | 1.70 | 10.0 | 0.3 | 1.00 | 2 |
| 13 | 10 | 3.40 | 1.0 | 0.3 | 0.10 | 4 |
| 14 | 50 | 1.70 | 1.0 | 0.6 | 0.10 | 2 |
| 15 | 10 | 1.70 | 10.0 | 0.6 | 0.10 | 3 |
| 16 | 50 | 1.70 | 1.0 | 0.0 | 0.55 | 4 |
| 17 | 10 | 2.55 | 1.0 | 0.0 | 0.10 | 2 |

Figure legends

**Figure 1: Summary of the accelerated bioprocess development approach and Taxol biosynthetic pathway.** A) The accelerated process optimization and scale up approach involved DoE guided screening of a range of process parameters using advanced high-throughput microbioreactors (1 mL). The optimal factor combination was then elucidated via statistical modelling and validated at 1L and 5L scale. B) Heterologous Taxol pathway in *S. cerevisiae* cell factories. Galactose is converted into the universal diterpenoid precursor, GGPP, via the mevalonate pathway. Taxadiene synthase (*TASY*), expressed by both strains *LRS5* and LRS6 catalyzes the cyclisation of GGPP to taxadiene, along with small amounts of iso-taxadiene and additional diterpene side products. CYP725A4 and TAT, expressed by *LRS6* only, catalyze the subsequent hydroxylation and acetylation reactions, respectively. A further 16 enzymatic steps are believed to be required for the conversion of taxadiene-5-yl acetate to Taxol.

**Figure 2: Preliminary microscale cultivation of *LRS6***. LRS6 was cultivated in SDG (Yeast Nitrogen Base without Amino Acids, 6.7 g/L; Complete Supplement Mixture minus Uracil (CSM-URA), 0.77 g/L; Uracil, 20 mg/L; galactose, 20 g/L) and YPG (Yeast extract, 10 g/L; Peptone 20 g/L; Galactose, 20 g/L). (A) Biomass measured as optical density at 600nm, B) Taxane production C) pH D) DO. Measurements are mean ± standard deviation for triplicate cultivations.

**Figure 3: Scoping trial: validation of the definitive screening design.** A) Total taxane production was evaluated for shake flask cultivation of *LRS6* under least forcing (all factors set to low level), centre point (all factors set to mid-point) and most forcing (all factors set to high level) conditions. Gas chromatographs resulting from cultivation under most forcing condition. B) Gas chromatograph resulting from control cultivation of *LRS5 (expressing TASY only)* in the most forcing medium, C) Gas chromatograph resulting from cultivation of *LRS6 (expressing TASY, CYP725A4 + CPR and TAT)* in the most forcing medium.

**Figure 4: Definitive screening design results**. A) Online biomass data, B) Final total taxane titer and optical density values, C) Online pH data D) Final titer of each of the key pathway intermediates. Data is mean ± standard deviation for duplicate or triplicate cultivations under each of the 17 factor combinations of the DSD as described in Table 1 (Section 2.2.1.) at microscale.

**Figure 5: Defined media validation in 1L bioreactors.** Biomass, galactose and secondary metabolite

concentrations were monitored for *LRS6* cultivations using (A) optimized defined (3.4 g/L, ammonium sulphate 1 g/L, YSM 2 g/L, galactose 50g/L) and (B) standard defined media (YNB 1.7 g/L, ammonium sulphate 5 g/L, YSM 2 g/L, galactose 20g/L). Total taxane production was also monitored for the two cultivations (F). Accumulation of the key Taxol precursors and side products was also plotted for the optimized (B) and control (D) cultivations. Offgas CO2 data is also shown for both the control and optimized cultivations (E).

Figure 6: **Optimized defined media with increased ammonium sulphate**. *LRS6* was cultivated in the optimal defined media with a higher initial ammonium sulphate concentration (YNB 3.4 g/L, ammonium sulphate 10 g/L, YSM 2 g/L, galactose 50g/L). A) Biomass, galactose and secondary metabolite concentration. B) Total taxane accumulation. C) Accumulation of key products in the Taxol pathway.

Figure 7: **Optimized defined media with increased ammonium sulphate at 5L scale**. *LRS6* was cultivated in the optimal defined media with a higher initial ammonium sulphate concentration (YNB 3.4 g/L, ammonium sulphate 10 g/L, YSM 2 g/L, galactose 50g/L). A) Biomass and galactose concentration, B) Total taxane concentration, C) Accumulation of key products in the Taxol pathway.