scientific reports

Check for updates

OPEN

Pichia pastoris growth—coupled heme biosynthesis analysis using metabolic modelling

Agris Pentjuss[1,2]✉, Emils Bolmanis[3], Anastasija Suleiko[2,4], Elina Didrihsone[2], Arturs Suleiko[2,4], Konstantins Dubencovs[2,4], Janis Liepins[1], Andris Kazaks[3] & Juris Vanags[2,4]

Soy leghemoglobin is one of the most important and key ingredients in plant-based meat substitutes that can imitate the colour and flavour of the meat. To improve the high-yield production of leghemoglobin protein and its main component—heme in the yeast *Pichia pastoris*, glycerol and methanol cultivation conditions were studied. Additionally, *in-silico* metabolic modelling analysis of growth-coupled enzyme quantity, suggests metabolic gene up/down-regulation strategies for heme production. First, cultivations and metabolic modelling analysis of *P. pastoris* were performed on glycerol and methanol in different growth media. Glycerol cultivation uptake and production rates can be increased by 50% according to metabolic modelling results, but methanol cultivation—is near the theoretical maximum. Growth-coupled metabolic optimisation results revealed the best feasible upregulation (33 reactions) (1.47% of total reactions) and 66 downregulation/deletion (2.98% of total) reaction suggestions. Finally, we describe reaction regulation suggestions with the highest potential to increase heme production yields.

Meat is one of the significant sources of dietary protein. It is frequently recognised as a high-quality protein source due to its nutritional qualities and favourable sensory properties such as texture and flavour. However, the rising global population has led to a rise in the production and consumption of meat around the world[1], which in turn has raised environmental concerns regarding the usage of land and water, as well as its impact on pollution and climate change, greenhouse gas emissions, and the loss of biodiversity[2]. Recommendations to limit meat consumption have been made because meat is a critical source of protein and in numerous high-income countries, protein utilisation outperforms dietary necessities[3]. Plant-based meat can be an efficient alternative to livestock, as for the production of the primary plant proteins, fats, carbohydrates and vitamins are used. Plant-based meat is defined as a meat-like substance made from vegetarian-friendly ingredients like proteins (from soy and potatoes), fat (from coconut and sunflower oils), carbohydrates (potato starch, corn starch), nutrition additives (yeast extract, vitamins) and flavours (beet juice extract, apple extract and plant hemoglobin)[4]. The market for plant-based meat is anticipated to reach a significant milestone of $30.9 billion by 2026[5] due to the rise in the number of vegetarians, vegans, and flexitarians in Europe during the past several years[6]. The switch from an animal to a non-animal diet is also influenced by religious considerations and high production costs[7].

Heme-containing proteins are considered one of the crucial flavour and colourant contributors in animal red meat[8]. Also, hemoprotein addition to artificial meat or its analogues has been proven to display great potential to mimic the taste and smell similar to that of traditional meat. There are a number of hemoglobin analogues across many organisms. Leghemoglobins, are present in the root nodules of leguminous plants. Leghemoglobin (legH) is approved by the United States Food and Drug Administration (FDA) as a food colouring additive and can be used equivalently to hemoglobin and myoglobin in both vegetarian and non-vegetarian diets[9].

In plant-based meats legH unfolds upon cooking, releasing its heme cofactor to catalyse reactions that can transform the same biomolecules, isolated from plant-based sources, into the array of compounds that comprise the unique flavour and aroma of traditional meat[10]. The use of legH protein as a food additive also circumvents the contradictions of animal material used in the diet, especially for vegetarians.

Many attempts in biotechnology have already been made to increase hemoprotein yields, thus ensuring the feasibility of such technology for industrial-scale production. For example, to improve heme-containing protein production the addition of 5-aminolevulinic acid (ALA) to the growth medium and overexpression of

[1]Microbiology and Biotechnology Institute, University of Latvia, Jelgavas Street 1, Riga 1004, Latvia. [2]Latvian State Institute of Wood Chemistry, Dzerbenes Street 27, Riga 1006, Latvia. [3]Latvian Biomedical Research and Study Centre, Ratsupites Street 1 K-1, Riga 1067, Latvia. [4]Bioreactors.Net AS, Dzerbenes Street 27, Riga 1006, Latvia. ✉email: agris.pentjuss@gmail.com

heme transporter genes[11,12] was studied, furthermore, other groups studied extracellular hemin addition, which provided only 5–10% increase of heme-containing protein mass in respect to the total protein content[13]. In heme-containing protein production, extracellular precursor addition seems to not be an economically feasible approach and further studies should be focused to achieve commercially viable intracellular hemoglobin, myoglobin and legH production through genetic engineering.

Currently, microorganism platforms are the most perspective way of manufacturing legH in commercial amounts, for further integration into plant-based meat. Several legH-producing recombinant microorganisms are *Candida spp.*, *Hansenula spp., Pichia spp*., or *Toruplosis spp*. representatives, which are methylotrophic yeast and can utilize alcohols, also methanol, have been previously used for accomplishing the mentioned tasks[14]. Furthermore, alternative microorganisms, like *Kluyveromyces lactis,* which are capable of utilising cheap dairy industry by-products, like cheese whey, have been studied[15]. Another study reported that it was possible to increase legH secretion by 83-fold in a genetically engineered *P. pastoris* with respect to the wild-type strain, by overexpressing heme pathways and globin synthesis genes simultaneously[16]. A recent study demonstrated that by combining systems biology and genetic engineering approaches, heme production can be increased by 70-fold in *Saccharomyces cerevisiae*. This was done using metabolic modelling coupled with growth-coupled heme production analysis on the genome level, where the most promising analysis results were genetically implemented and experimentally tested[17].

In the last decade, genome-scale metabolic modelling is increasingly being used as an auxiliary tool for facilitating various genetic engineering tasks and for predicting potential effects. Metabolic models effectively encode gene-protein-reaction (GPR) associations[18] and explain genotype–phenotype interaction in unicellular and multi-cellular organisms[19,20].

Metabolic modelling allows us to predict the organisms phenotypes biotechnological properties and responses[21], and predicts environmental perturbation impacts on the metabolism and potential carbon distribution[22,23]. Genome-scale metabolic model (GEM) application in biotechnology already has proven its importance and helped to develop new nutrient uptake innovations[24], novel software developments[25,26], created new biotechnology applications[27,28].

GEM investigation has guided the development of strains with optimised yields of industrial metabolites (e.g., sesquiterpenes, vanillin, bioethanol, 2,3-butanediol, succinate, amorphadiene, β-farnesene, and dihydroxyacetone phosphate, 3-hydroxy propionate, fumaric acid)[29,30].

One of the most common approaches to metabolic engineering and computational biology during the last decade has spread to link the production of the desired metabolite with coupled growth[31]. This method is very effective in new strain design because metabolic production becomes a necessary by-product for cell growth. Furthermore, to increase desired metabolite production yields and production capabilities it is possible adapting laboratory evolution by selecting for maximum growth[32,33].

Over the past 20–30 years, the *Pichia pastoris* expression system has been extensively used to produce various recombinant proteins for research and industrial applications. This methylotrophic yeast is highly suitable for heterologous protein expression due to several key features: ease of genetic manipulation, high-frequency DNA transformation, cloning by functional complementation, high levels of intracellular and extracellular protein expression, and the ability to perform higher eukaryotic protein modifications (such as glycosylation, disulphide bond formation, and proteolytic processing)[34,35]. In addition, relatively low levels of native secreted proteins facilitate the purification of the secreted recombinant proteins. Taking into account the economic factors (such as high cell growth in minimal media and high product stability in prolonged processes) and the advanced genetic techniques available, *P. pastoris* is the system of choice for heterologous protein expression[36,37]. The versatility of *P. pastoris* is reinforced by the high yields of various recombinant proteins achieved in this microorganism[16,38–40].

Thus, to increase the production of legH protein to economically feasible quantities, global intracellular reactions/gene manipulations must be carried out.

This study aimed to perform biomass growth-coupled analyses to determine a list of the reactions necessary to up or down-regulate to increase heme biosynthesis in *Pichia pastoris*. We exploited fermentation data and FBA (Flux Balance Analysis) to estimate the potential of the *P. pastoris* biomass yield. To identify particular reactions, we developed a growth coupled analyses algorithm to systematically explore the *P. pastoris* genome-scale metabolism model to identify reactions with the highest impact on heme production.

Results and discussion

Process in *P. pastoris*. In the first step, *P. pastoris* is grown on glycerol to produce biomass. In the second step, *P. pastoris* is cultivated on methanol as a sole carbon source and legH synthesis is induced. We introduced the soya (*Glycine max*) leghemoglobin coding gene into *P. pastoris* X-33 strain genome under control of methanol inducible promoter (see details in Materials and methods). We performed *P. pastoris* biomass production followed by legH induction and growth on methanol. We analysed legH production using polyacrylamide gel.

First, we performed FBA to test if *P. pastoris* growth on glycerol or methanol are optimal. Then we built an algorithm to find out reaction sets linked to *P. pastoris* growth and heme production.

P. pastoris growth in methanol is close to the theoretical maximum

We tested the GEM ability to simulate physiologically relevant conditions and analyse heme biosynthesis growth-coupled production. We performed experimental fermentations with *Pichia pastoris* strain X-33 containing genome-integrated legH gene (see details Materials and methods). In GEM analyses we assumed that each legH protein contains one heme molecule[41]. Therefore, legH production is linked to biomass growth and is limited by intracellular heme production.

To develop *P. pastoris* reactions regulation design sets for genetic engineering purposes and increase intracellular heme production rate, we used *P. pastoris* iMT1026 GEM, which consists of 1708 metabolites, 2243 reactions, 2 different biomass compositions for glycerol and methanol fermentation optimisations, has 9 different compartments and transport intercompartment[42]. iMT1026 model was updated and missing mandatory information filled to be compatible with CobraToolbox 3.0[43].

GEM includes metabolite name, charged formula and compartment name in the cell. As iMT1026 has no metabolite or reaction database references, thus it was difficult to test new functionality in comparison with scientific literature. GEM is considered to work at pH 7.2, 30 °C environment temperature.

We performed *P. pastoris* cultivation in synthetic media on glycerol and methanol and measured the specific growth rate ($\boldsymbol{\mu}$). We found, that the growth rate of glycerol was 6.3–4.75-times higher ($\boldsymbol{\mu} \sim 0.19\ h^{-1}$) than consuming methanol ($\boldsymbol{\mu} \sim 0.04$–$0.03\ h^{-1}$), which is in the scope of previously published data (1.7–8.5 times faster)[44]. The O_2 consumption rate on glycerol growth was 2.17–2.29 $mmol*g^{-1} * h^{-1}$, but CO_2 production was a little less (1.61–1.7 $mmol*g^{-1} * h^{-1}$) than O_2 consumption (Supplementary Materials 1). Data from S6.13 contains detailed GEM constraints from experiments.

After obtaining growth, O_2 and CO_2 uptake rates in the glycerol/methanol-driven cultivations, the GEM model was accordingly constrained with experimentally measured exometabolomics data. After *P. pastoris* GEM optimisations, results suggested that the GEM model can achieve a steady state and produce heme. Experimentally *P. pastoris* biomass growth was measured by a turbidity sensor (Optek, Germany) or permittivity probe (Hamilton, Incyte, Switzerland). Both sensors delivered similar tendencies during the first stage of the fermentation (glycerol). However, culture permittivity results started to decrease gradually during the methanol fermentation stage, thus indicating loss of viability when methanol feed was increasing (see materials and methods). The same conclusions have been made by earlier studies[11–13], where it was concluded, that extracellular addition of substrates is a less effective strategy and has its limitations in comparison to genetic engineering potential. For example, 70 times increase in heme production was achieved in *S. cerevisiae* when using GEM-driven metabolic engineering[17].

Growth coupled analyses of heme production

Thus, the next step in this study was to use *P. pastoris* GEM to make a systemic reaction impact analysis on potential heme production. Previously described experimental and GEM optimisation results comparisons suggested a potential to increase heme production at the expense of biomass production. This could be one of the GEM modelling bottlenecks in the biomass reaction itself. In GEM, a biomass objective function[45] exists, which defines necessary metabolic resources for cell process machinery and proliferation needs. But the biomass objective function data, experimentally measured in strict cultivation conditions, is expensive to measure and sometimes even has a large error distribution. Also, it is known that biomass objective function composition and macromolecule weighting factors can change during different environmental conditions[46].

In this study, we improved the previously published growth-coupled GEM algorithm and adjusted it to new GEM requirements. To exclude biomass objective function limitations, we decided that growth-coupled GEM algorithm functionality (see materials and methods) will systematically analyse all metabolic reactions' impact on heme production.

The weighted factor shows how much percentage each flux value changes versus the mean value and is summed together. As not all reaction flux changes are in linear (Supplementary Materials 6 Table S6.2), then we summed all reaction changes and calculated the sum of percentages. The bigger the number, the higher the impact on the target product (increased heme production) (Formula (1)). All genetic regulation suggestions are done using methanol as a carbon source.

As a result, we found the most promising upregulation (33 reactions) (1.47% of total reactions) (Supplementary Materials 6 Table S6.7) and 66 downregulation/deletion (2.98% of total) reaction candidates (Supplementary Materials 6 Table S6.8) (Formula (1)). All reactions in the materials and methods will be described as:

Reaction name (ID in the model) (ID in the Metacyc).

Upregulated reactions

To find out reactions which should be upregulated when coupling *P. pastoris* cell growth with legH production—we listed all reactions whose rate increased when the biomass rate decreased. To increase heme intracellular production, the algorithm has suggested upregulating reactions from the dihydroxyacetone cycle combined with glycolysis reactions (Supplementary Materials 6, Table S6.1) and heme biosynthesis pathways: tetrapyrrole biosynthesis II (Supplementary Materials 6, Table S6.5) and aerobic heme b biosynthesis I (Supplementary Materials 6, Table S6.6).

Interestingly the highest step-weighted factor for upregulation is mitochondrial pyrroline-5-carboxylate reductase. The conversion of glutamate to proline and vice versa is thought to play a role in ATP increase[47,48]. This could be related to 5-aminolevulinate synthase (ALASm) (5-AMINOLEVULINIC-ACID-SYNTHASE-RXN), where Succinyl CoA (succoa) (SUC-COA) is consumed in large amounts as a 5-Aminolevulinate (5aop) (5-AMINOLEVULINIC-ACID-SYNTHASE-RXN) precursor. In the mitochondrial TCA cycle- (PWY-5690) Succinate–CoA ligase (ADP-forming) reaction (SUCCCOASYN-RXN), which is one of the suggested upregulation reactions is used to produce Succinyl CoA (SucCoA) (SUC-COA) and consume large amounts of ATP (atp) (ATP), which could be a limiting step for improved intracellular heme production if less ATP (atp) (ATP) is available in mitochondria (Supplementary Materials 3 Formaldehyde assimilation III).

To increase intracellular heme bioproduction the most perspective upregulation reactions are 5-aminolevulinate synthase (ALASm), porphobilinogen synthase (PPBNGS), hydroxymethylbilane synthase (HMBS), uroporphyrinogen-III synthase (UPP3S) from heme b biosynthesis I pathway (heme-BIOSYNTHESIS-II)

(Supplementary Materials 6 Table S6.6) and uroporphyrinogen decarboxylase (UPPDC1), coproporphyrinogenase (CPPPGO), protoporphyrinogen oxidase (PPPGOm), protoporphyrin ferrochelatase (FCLTm) from tetrapyrrole biosynthesis II (from glycine) pathway (PWY-5189) (Supplementary Materials 6 Table S6.5), which together forms the heme biosynthesis metabolic pathway (Porphyrin and Chlorophyll Metabolism in GEM). These reactions have one of the highest step-weighted factors and have the highest impact on heme production increase as a whole pathway. In *S. cerevisiae* the *HEM3* gene (in GEM HMBS reaction) was identified as rate limiting[49], but recent studies in *P. pastoris* showed that *HEM1* (in GEM ALASm)[16] is limiting. They analysed single gene upregulation of heme b biosynthesis I pathway (heme-BIOSYNTHESIS-II) and tetrapyrrole biosynthesis II (from glycine) (PWY-5189) pathways and concluded that ALASm reaction upregulation had the most impact on legH production compared to other heme biosynthesis pathway reactions. The same conclusions indirectly suggest GEM pyrroline-5-carboxylate reductase step-weighted factor algorithm results. Thus, different results are found in different yeast species, which suggests heme production in various species is differently regulated. More the study suggested that *HEM1* (ALASm), *HEM2* (PPBNGS), *HEM3* (HMBS) and *HEM12* (FCLTm) upregulation showed a 41% increase and all heme production genes overexpression led to 1.5-fold increased heme production, leaving the potential for much more improvements as it was recently found in *S. cerevisiae*[17].

The next reaction group with step-weighted factor (11–14%) are (a) formyltetrahydrofolate dehydrogenase (FTHFDH) and methylenetetrahydrofolate dehydrogenase (NADP) (MTHFD) (Supplementary Material 6 Table S6.2), which are one-carbon metabolism pathway reactions responsible for the interconversion of different tetrahydrofolate (THF) forms. Glycine hydroxymethyltransferase (GHMT2r) (GLYOHMETRANS-RXN) is involved in anaerobic purine degradation. Proposed candidate upregulation could re-route additional resources to increased heme production.

(b) Phosphoglycerate dehydrogenase (PGCD), phosphoserine transaminase (PSERT), and phosphoserine phosphatase (L-serine) (PSP_L) (Supplementary Materials 6 Table S6.3) are serine biosynthesis reactions. Glutamate dehydrogenase (NADP) (GLUDyi) and pyruvate carboxylase (PC), which in GEM are glycine (gly) precursors. The glycine is the precursor of heme b biosynthesis I pathway (heme-BIOSYNTHESIS-II) and tetrapyrrole biosynthesis II (from glycine) (PWY-5189), which produces heme.

As the last group of upregulation suggestions are the Pentose Phosphate pathway Dihydroxyacetone synthase (DAS) and Dihydroxyacetone kinase (DHAKx) from the methanol metabolism pathway, which together form formaldehyde assimilation III (dihydroxyacetone cycle) (P185-PWY) (Supplementary Materials 6 Table S6.1). These reactions have less step-weighted factor (5–9%) and have less impact on heme biosynthesis. These reactions are closely interconnected and allow methylotrophs to metabolise methanol to formaldehyde (fold), which later forms Glyceraldehyde 3-phosphate (G3P) and dihydroxyacetone (DHA)—the building block chemicals for all other metabolic processes, including of heme biosynthesis.

Downregulation/deletions calculations

Relying only on an upregulation strategy to increase heme intracellular production is not sufficient enough. Previous studies on *Escherichia coli*[50] proved that metabolic modelling proposed unnecessary metabolic process deletion for glycerol uptake allowed to increase succinate production yields. This demonstrated that some metabolic processes might be not essential and even lower product yield.

We found 66 individual reactions, in which flux rates decrease while heme production increases (change inversely proportional to heme production).

The largest group of downregulation/deletion candidates are different amino acids and their intermediate interconversion-related metabolism reactions with step-weighted factor 17.3 (Supplementary Materials 4). The exception is Aspartate 1-decarboxylase (ASP1DC) with a lower step-weighted factor 15. The group consists of 26 reactions from 66 (~ 40% of the total), involving 8 different amino acid pathways (Fig. 1) (Supplementary Materials 6 Table S6.9).

The largest amount of candidate reactions are:

- Valine, Leucine, and Isoleucine metabolic pathway: 7 reactions (25, 93%);
- Threonine and Lysine pathway: 6 reactions (22, 22%)
- Tyrosine, Tryptophan and Phenylalanine pathway: 5 reactions (18, 52%);
- Cysteine pathway: 3 reactions (11, 11%);
- Methionine pathway: 3 reactions (11, 11%)

We assume that in our experimental model, *P. pastoris* amino acid metabolism for *biomass growth* competes with heme biosynthesis for glycine. Additionally, *P. pastoris* upregulate the C1 metabolism pathway in mitochondria to increase glycine synthesis necessary for heme (heme-BIOSYNTHESIS-II and PWY-5189) biosynthesis (see chapter "upregulations").

The downregulation of amino acid metabolism could be related to the allocation of resources towards glycine and heme synthesis. Threonine is starting point for leucine and isoleucine synthesis, it is produced by the threonine aldolase from glycine. By lowering leucine, valine and isoleucine metabolism—the system retains more glycine to allocate for heme synthesis. Detailed biochemical pathways and reactions description is found in Supplementary Materials 4.

The next largest group is the nucleotide metabolism pathway 11 (~ 16.5% of total) reactions, in which the step-weighted factor is 17.3–18.1% (Fig. 2). The algorithm suggests that these reactions group downregulation/deletions are not essential and will not affect cell growth rate. Purine biosynthesis synthesis depends on glycine and folate supply—one glycine and two formyltetrahydrofolate molecules are necessary to synthesize one purine molecule[51]. Our results reveal, that DHFRim (dihydrofolate reductase, mitochondrial) should be downregulated

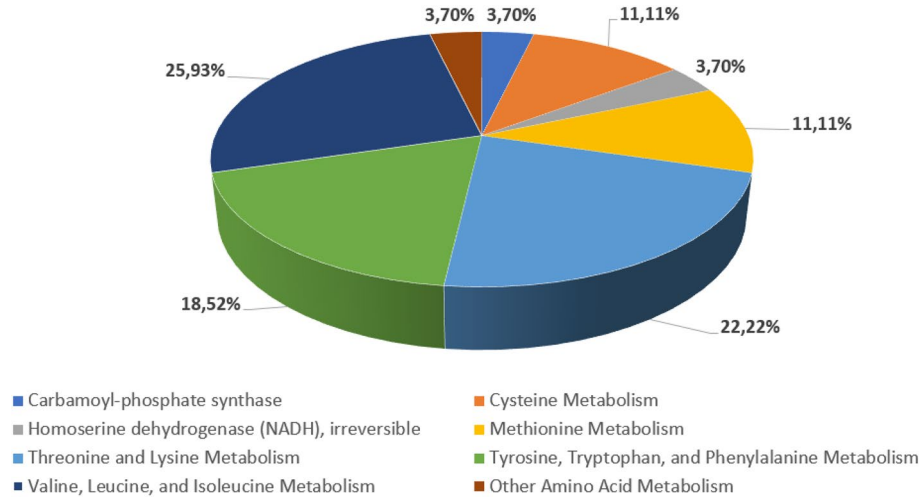

Figure 1. Fractional distribution of amino acids downregulation/deletion reaction candidates.

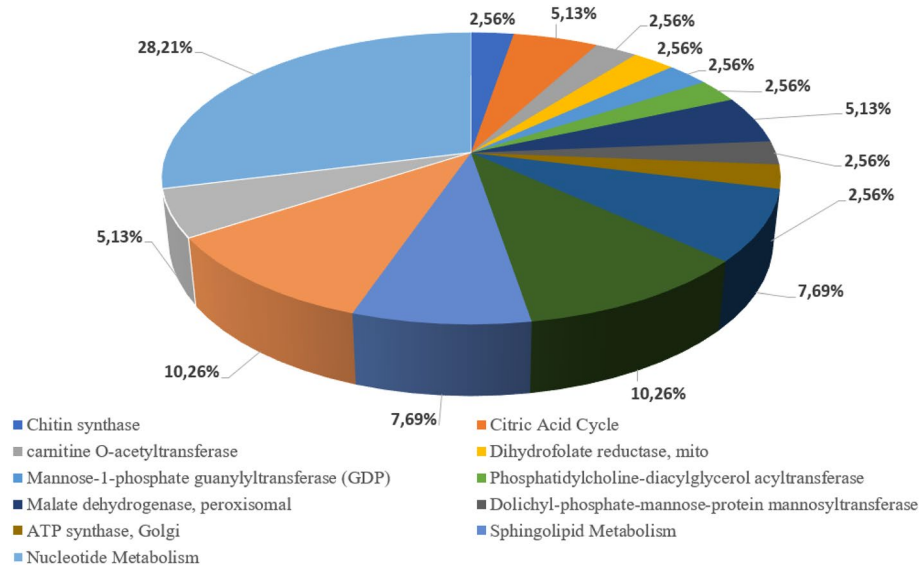

Figure 2. Downregulation/deletion suggestions for other metabolic pathways and compartments.

to achieve max heme synthesis. There are 2 reaction chains from a 5, 10-methylenetetrahydrofolate to tetrahydrofolate in yeasts and DHFRim is glycine dependent, while the parallel reaction chain is not.

Therefore, by keeping upregulating C1 metabolism and glycine synthesis and downregulating their consumption elsewhere—heme yield can be improved. Moreover, purine synthesis uses 4 ATP per 1 purine molecule; if down-regulating the purine synthesis pathway, some ATP could be reallocated to heme synthesis.

The third large group consists of 23 reactions from different metabolic pathways (Supplementary Materials 6 Table S6.9), which the GEM algorithm suggests as potential downregulation/deletion reactions (Fig. 2). Detailed biochemical pathways and reactions description is found in Supplementary Materials 5.

The algorithm suggests that these reactions and pathways have an indirect effect on heme production and must be detailed and analysed with additional omics data measurements during the cultivation process.

Downregulation of sterol metabolism might be related to the fact, that within the sterols desaturation reactions in yeasts—proteins contain one or two heme and Fe atoms in their active sites[52], therefore maximising surplus heme production can be achieved by downregulation their consumption.

Summarising everything described above, we would like to point out the most promising up/down-regulation and deletion suggestions.

For upregulation, the best candidates are heme-BIOSYNTHESIS-II and PWY-5189, which include all heme biosynthesis pathways. Previous studies showed that HEM3 (HMBS reaction) is a rate-limiting step in *S. cerevisiae*[17], but in *P. pastoris* it is not. More, experimental data showed that *HEM1* (ALASm), *HEM2* (PPBNGS),

HEM3 (HMBS) and *HEM12* (FCLTm) upregulation leads to a 41% increase in intracellular heme production and all HEBE biosynthesis reactions upregulation lead only to 1,fivefold increase.

In comparison with *S. cerevisiae,* the latest report shows 70-fold intracellular heme increase potential by upregulating heme biosynthesis pathways and deleting not essential carbon or nitrogen-consuming reactions like serine hydroxymethyltransferase (GHMT2r), heme oxygenase (biliverdin-producing) (heme-OXYGENASE-DECYCLIZING-RXN in MetaCyc[53]) and glycine cleavage system (GLYCLm). Heme-OXYGENASE-DECYCLIZING-RXN is not found in the iMT1026 GEM, thus hasn't been calculated step-weighted factor, but in other organisms, the reaction converts heme to bilirubin, which would rapidly decrease intracellular heme concentration.

GLYCm reaction converts glycine and THF to 5, 10-Methylenetetrahydrofolate and is used as a response to high concentrations of the glycine, thus diverting and decreasing it from heme production.

GHMT2r is responsible for converting serine to glycine using thf as a co-factor. Our model suggests it as an upregulation candidate, but published results showed that deletion will in total increase intracellular heme. This is explained because our algorithm does not include quantitative omics measurements which more precisely explain enzyme activity, available amount, and indirect effects on other metabolic processes. Nevertheless, comparing within literature published data algorithm shows some inconsistencies, still, it shows feasible upregulation (33 reactions) (1.47% of total reactions) and 66 downregulation/deletion (2.98% of total) reaction suggestions results.

Conclusions

Our work is the first in silico *P. pastoris* GEM analysis that demonstrates the potential directions for metabolic engineering to improve *P. pastoris* as an expression platform for legH production.

P. pastoris metabolism, although well adapted to heterologous protein production, it's biomass flux can be increased. When comparing our fermentation data with in silico *P. pastoris* GEM optimisation results, we found, that the biomass growth rate on glycerol could be increased by 50%. However, *P. pastoris* growth on methanol is close to the theoretical maximum. Our experiments showed that the specific growth rate (μ) is 0.19 h^{-1}, which is a little higher than Looser et al.—0.15 h^{-1}[54] and 0.06–0.16 h^{-1}[55].A higher biomass growth rate on the glycerol is an opportunity to shorten the biomass production phase before switching to methanol induction. Since legH production is linked to biomass growth, legH production is limited with the "free heme" bioavailability. Thus, if part of the growth resources could be rerouted to the heme production—an increase in heme and legH production would be expected.

We developed an algorithm to systematically explore *P. pastoris* GEM response to the decrease in biomass growth rate and identified 500 reactions which had an impact on heme production. As the most promising targets—the algorithm filtered out 33 reactions for upregulation and 66—as downregulation targets.

For upregulation, the algorithm suggests that heme-BIOSYNTHESIS-II and PWY-5189 are the best candidates and formaldehyde assimilation III (dihydroxyacetone cycle) (P185-PWY) is the less effective candidate. For downregulations/deletion suggestions, most potential targets are identified within amino acid and nucleotide metabolism. Heme biosynthesis competes with amino acid metabolism for common NH2 donor glutamine and methyl group donor glycine.

A similar approach has been done by Ref.[56], where Yeast8 genome-scale metabolic model and ecYeast8 enzyme-constrained metabolic model combinatorial analysis were done to find out the best up/down-regulation gene candidates and experimental validation. In enzyme-constrained metabolic models, reaction rates are constrained by the corresponding intracellular enzyme concentration multiplied by its turnover number (kcat) by keeping fluxes at physiologically relevant levels[57]. Compared with GEM, enzyme-constrained GEM has a large reduction in flux variability for most of the metabolic reactions. Since *P. pastoris* GEM is less detailed than Yeast8 metabolic model[58] and our best knowledge has less available enzymes concentration experimental data, thus we chose to use the algorithm not related to enzyme concertation-based GEM analysis. This would force to implement data from other yeast species and could lead to biologically irrelevant results. Our results showed the 33 best upregulation candidate reactions comparing with 40 upregulation gene candidates in *S. cerevisiae*, which are quite a similar number of candidates.

In *P. pastoris* growth rate is tightly regulated to translation rate[59]. The next engineering task involving modelling and genome-wide engineering would be to decouple translation machinery from growth and to reallocate more resources for heterologous protein production. We speculate, that our analyses would help to choose strategies for metabolic engineering for a number of heme-containing proteins like peroxidases, catalases, etc. when heterologous expressed in *P. pastoris.*

Materials and methods

Genome-scale metabolic model

Growth-coupled analysis

To analyse GEM for growth-coupled reactions and reaction manipulation impact on heme biosynthesis, we improved *P. pastoris* iMT1026 GEM, which consists of 1708 metabolites, 2243 reactions, 2 different biomass compositions for glycerol and methanol fermentation optimisations, has 9 different compartments and transport intercompartment for Cobra toolbox 3.0 compatibility[42].

We developed an algorithm based on a previously published modelling approach for Lycopene production improvement for gene amplification targets[60]. The main concept of the framework is to select reaction amplification targets for improved product formation. The framework concept is to search for the candidate reaction, which flux changes have the most impact on target metabolite production under maximising biomass reaction flux. During analysis 3 different enzymatic reaction types were determined: flux-increasing reaction,

flux-decreasing reaction and unaffected. In the end, algorithm develops an MS Excel file with exact results in specific tables (Supplementary Materials 6), the content of each table is listed in Supplementary Materials 2.

Before using the growth-coupled GEM modelling algorithm, there is a need to find the non-growth ATP maintenance cost. The model must be able to reach a steady state and produce ATP, which is necessary for maintenance processes. This is shown in Fig. 3 where the biomass formation minimal flux value must be calculated and minimal ATP consumption must be maintained, otherwise, the model will not reach a steady state due to not enough ATP metabolic pool. After this value has been calculated, the algorithm finds the theoretical maximal biomass formation flux rate. When the minimal and maximal biomass flux rates are determined, the growth-coupled GEM modelling algorithm calculates the heme production maximal flux value by setting it as an objective function and maximising using the Flux Balance Analysis approach (FBA)[61]. In this stage, the growth-coupled GEM modelling algorithm has done initiation steps and is ready to find out each GEM reaction's impact on product formation rate. In our case product formation is the heme production rate. The next step was to count heme production flux value increase step count (Fig. 3 black dots), which was divided into 5 equal steps to obtain the most promising optimisation results. In each step, the algorithm determines the objective reaction flux value change ratio against product flux increase value. The results are saved in a Supplementary Materials 6 Table S6.2, where Reactions ID from GEM and reactions names are saved in the "Reactions_ID" and "Reactions_name" columns and all 5 forced flux values of each reaction are saved in the "FBA_results_" columns. This table contains all found reactions with direct and indirect connections with heme production. More, we included subsystem and reactions stoichiometry information from GEM in the "Subsystem" and in the "Var10" columns to improve later data analysis steps.

After initiation and heme production flux value step count determination, the optimisation task was done. During optimisation, the last step for the heme transport reaction was set less than 10% less of the maximal heme calculated value. This is important because in GEM reactions flux distribution is calculated by constraint-based flux analysis and, without additional omics data constraints implementation, the maximal flux rate constraint mostly is biologically irrelevant and biomass formation flux value becomes 0. In this case, there is no growth in the model and it is not possible to calculate growth-coupled reactions and genes manipulation impact on heme biosynthesis.

As a result, the algorithm determines 3 different GEM reaction types: flux-increasing reaction, flux-decreasing reaction and unaffected (Fig. 4). Flux-increasing reactions are considered for up-regulation, Flux-decreasing reactions are considered for down-regulation or deletion, for non-essential metabolites (the biomass building block

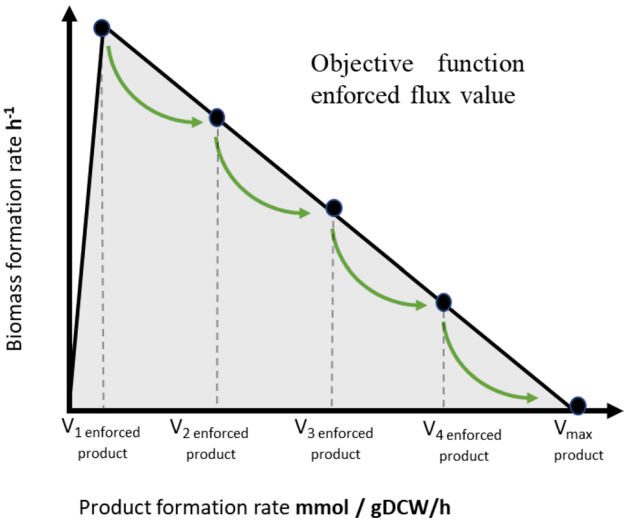

Figure 3. Growth-coupled GEM modelling algorithm concept.

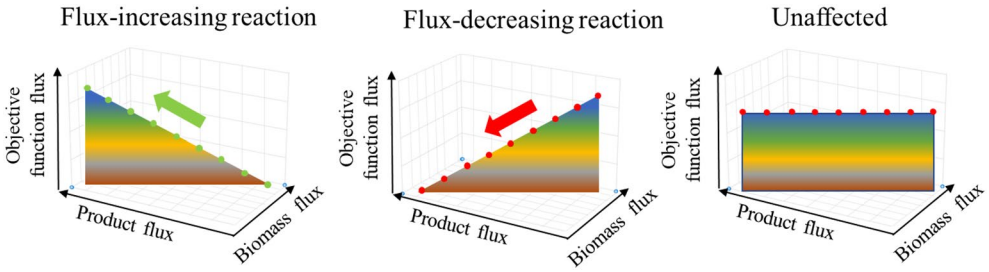

Figure 4. Reaction types identified by Growth-coupled GEM modelling algorithm.

chemicals or non-essential amino acids). There are also unaffected reactions, which do not impact heme production. The algorithm finds out all 3 types of reactions and saves them in Supplementary Materials 6 Table S6.2).

After growth-coupled GEM modelling algorithm optimisations, we made Excel additional reactions data filtering because not all irreversible reactions stoichiometry in GEM are in the same direction, which can lead to incorrect or incomplete optimisation results. Also, not all reactions had linearly changing flux values or even steady-state results, thus only reactions with all 5 results and steadily changing reaction flux values were chosen for the next analysis steps. As a result, the algorithm generates 4 additional MS Excel data tables:

1. "Positive_contra_proportional"- reactions with flux value from left to right and are inversely proportionally to forced product (heme) changing fluxes;
2. "Negative_contra_proportional"—reactions with flux value from right to left and are inversely proportionally to forced product (heme) changing fluxes;
3. "Positive_directly_proportional"—reactions with flux value from left to right and are directly proportionally to forced product (heme) changing fluxes;
4. "Negative_directly_proportional"—reactions with flux value from right to left and are directly proportionally to forced product (heme) changing fluxes.

When all upregulation and downregulation/deletion reactions are filtered out and saved in separate data tables (Supplementary Materials 6 Table S6.3–S6.6), then additional calculations are made. The algorithm finds and writes in the MS Excel file, the Boolean value if the reaction flux is left to right or right to left. This is necessary if the algorithm makes unpredicted inconsistencies, then manual filtering in the file is available.

Step-weighted factor

To determine the most direct proportional or inversely proportional impact on objective product improved production, we introduced the step-weighted factor (SWF), which was calculated as (Formula (1)):

$$\mathrm{SWF} = \sum_{1}^{5} \frac{Product\ flux}{mean\ flux} * 100\%. \tag{1}$$

The weighted factor shows how much percentage each flux value changes versus the mean value and is summed together. As not all reaction flux changes are in constant value (Supplementary Materials 6 Table S6.2), then we summed all reaction changes and calculated the sum of percentages. The bigger number the better impact on the target product (increased heme production).

Reactions with the largest step-weighted factor are the main candidates for reaction flux and gene regulation to increase heme production yields. Additionally, we calculated each reaction essentiality, which is reaction deletion essential for growth. For downregulation/deletion candidates only non-essential reactions were used (Supplementary Materials 6 Table S6.8).

Candidate reaction analysis

After all, the reactions are sorted out, and the algorithm will filter the most potential ones, which upregulation could lead to a heme production increase. The second step is to find out which reactions must be downregulated/deleted to increase heme production. Deletion of the specific reaction in living cells can sometimes lead to lethality. To find out essential reactions in the filtered most potential reactions list we performed a single-reaction deletion analysis using CobraToolbox 3.0. In the end, we got reactions deletion list and genes deletion list suggestions for experimental genetic manipulation strain designing.

Construction of the expression vector and selection of clones

An artificial gene encoding the soy legHemoglobin sequence (GenBank Acc. NP_001235248.2) was designed by GenScript and synthesized by BioCat GmbH (Heidelberg, Germany). The gene was cloned into the pPICZC vector (Invitrogen) using EcoRI and NotI restriction sites (Fig. 5). The plasmid was linearized with PmeI and transformed in *Pichia pastoris* strain X-33 by electroporation. Mut + transformants were produced on agarized YPD plates containing 800 µg/mL zeocin and analytically cultivated in flasks with rich BMMY medium for three days. The total protein was visualised on Coomassie-stained PAGE. The best legH-producing clone was selected for fermentation studies.

Fermentation conditions

Batch cultivation and feeding media solutions used in this study were prepared according to the "Pichia fermentation process guidelines" by Invitrogen Corporation[62]. 1.9 l of Basal Salts Medium: 26.7 ml L^{-1} H_3PO_4 85%, 0.93 g L^{-1} $CaSO_4$, 18.2 g L^{-1} K_2SO_4, 14.9 g L^{-1} $MgSO_4{\cdot}7H_2O$, 4.13 g L^{-1} KOH, 40.0 g L^{-1} glycerol and 4.35 ml L^{-1} PTM1 trace-element solution (0.02 g L^{-1} H_3BO_4, 5 ml L^{-1} H_2SO_4 98%, 6.0 g L^{-1} $CuSO_4{\cdot}5H_2O$, 0.08 g L^{-1} NaI, 3.0 g L^{-1} $MnSO_4{\cdot}H_2O$, 0.2 g L^{-1} $Na_2MoO_4{\cdot}2H_2O$, 0.5 g L^{-1} $Ca_2SO_4{\cdot}2H_2O$, 20.0 g L^{-1} $ZnCl_2$, 65.0 g L^{-1} $FeSO_4{\cdot}7H_2O$, 0.2 g L^{-1} biotin), were inoculated with 100 mL of inoculum grown in BMGY medium (10.0 g L^{-1} yeast extract, 20.0 g L^{-1} peptone, 100 mM potassium phosphate buffer, pH 6.0, 13.4 g L^{-1} yeast nitrogen base, 10.0 g L^{-1} glycerol, and 0.0004 g L^{-1} biotin) at 30 °C for 18–22 h in a shake flask at 250 RPM. Two feeding solutions were used—glycerol fed-batch solution (50% glycerol, 12 mL L^{-1} PTM1) and methanol fed-batch solution (100% methanol, 12 mL L^{-1} PTM1).

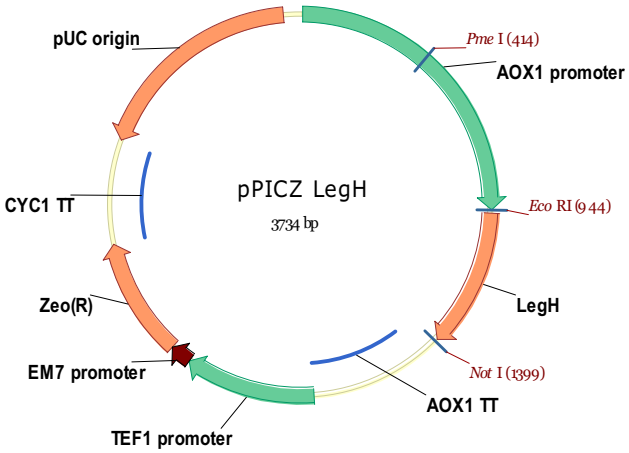

Figure 5. Schematic map of plasmid used for LegH production. Indicated are the *Pme*I site used for linearization as well as *Eco*RI and *Not*I sites used for cloning.

The bioreactor vessel was filled with distilled water and sterilized at 121 °C for 30 min, while the cultivation mediums and glycerol fed-batch solutions were autoclaved separately at the same conditions. The PTM1 and methanol-fed-batch solutions were sterilized by filtration through a 0.2 µm filter.

The fermentations were conducted in a 5 L bench-top fermenter (Bioreactors.net, EDF-5.4/BIO-4, Latvia) with a working volume of 2–4 L. The pH was monitored using a calibrated pH sensor probe (Hamilton, EasyFerm Bio, Switzerland) and adjusted to 5.0 ± 0.1 with a 28% NH_4OH solution before starting the cultivation process and maintained at the set value during fermentation. The temperature was controlled at 30.0 ± 0.1 °C, using a temperature sensor and by adjusting the temperature in the vessel jacket. The dissolved oxygen (DO) level was measured using a DO probe (Hamilton, Oxyferm Bio, Switzerland) and kept above 30 ± 5% by varying the stirrer speed (200–1000 RPM) (Cascade 1) or enriching the inlet air with pure O_2 (Cascade 2). A constant air or air/oxygen mixture at a flow rate of 3.0 slpm was maintained throughout the process. A condenser was used to condense moisture from outlet gasses, and antifoam 204 (Sigma) was added when necessary to control excessive foam formation. Substrate feed solutions were pumped using a high-accuracy peristaltic pump (Longer-Pump, BT100–2 J, China).

Process O_2 and CO_2 concentration were measured in the reactor exhaust gas using an O_2/CO_2 analyser (Bluesens, BlueInOneFerm, Germany). Culture turbidity was estimated using a turbidity probe (Optek, ASD19-EB-01, Germany) measuring the light absorption (transmission) within the 840–910 nm wavelength range. Finally, a permittivity probe (Hamilton, Incyte, Switzerland) was used to estimate the viable cell concentration during cultivation.

The cultivations began with a glycerol batch phase. After 18–24 h, the batch glycerol was depleted, and a glycerol fed-batch solution was fed into the reactor at a rate of 0.61 mL/min for 4 h, or until an optical density of 100–120 was achieved. After a brief feeding pause of 10–30 min, to allow the cells to consume all of the residual glycerol, the feed substrate was switched to methanol, and the solution was fed into the reactor at a rate of 0.12 mL/min for 5 h, followed by 0.24 mL/min for 2 h, and finally 0.36 mL/min for the remainder of the cultivation.

Analytical measurements

Cell growth was observed by offline measurements of the culture optical density (OD) at a wavelength of 590 nm (GRANAT, KFK-2, St. Petersburg, Russia). Wet cell weight (WCW) measurements were determined gravimetrically. Biomass samples were placed in pre-weighted Eppendorf® tubes and centrifuged at 15′500 g for 3 min. Afterwards, the supernatant was discarded, the cells resuspended in distilled water and centrifuged once more. The liquid phase was discarded and the remaining wet cell biomass was weighed. Dry cell weight (DCW) was determined using a previously-determined empirical formula (Formula (2)):

$$DCW = WCW * 0.3 \quad (2)$$

Data availability

All data generated or analysed during this study are included in this published article and Supplementary Materials. All additional data are found in the link: https://github.com/BigDataInSilicoBiologyGroup/Pichia_pastoris_metabolic_modeling. Scripts are found at https://github.com/BigDataInSilicoBiologyGroup/Pichia_pastoris_metabolic_modeling/tree/main/Scripts. All Supplementary Materials are found online: https://github.com/BigDataInSilicoBiologyGroup/Pichia_pastoris_metabolic_modeling/tree/main/Supplementary_materials.

Received: 30 March 2023; Accepted: 15 September 2023
Published online: 22 September 2023

References

1. Adam, D. How far will global population rise? Researchers can't agree. *Nature* **597**, 462–465. https://doi.org/10.1038/d41586-021-02522-6 (2021).
2. Cheng, M., McCarl, B. & Fei, C. Climate change and livestock production: A literature review. *Atmosphere* **13**, 140 (2022).
3. Kwasny, T., Dobernig, K. & Riefler, P. Towards reduced meat consumption: A systematic literature review of intervention effectiveness, 2001–2019. *Appetite* **168**, 105739 (2022).
4. Kyriakopoulou, K., Keppler, J. K. & van der Goot, A. J. Functionality of ingredients and additives in plant-based meat analogues. *Foods* **10**, 600 (2021).
5. Nezlek, J. B. & Forestell, C. A. Meat substitutes: Current status, potential benefits, and remaining challenges. *Curr. Opin. Food Sci.* **47**, 100890 (2022).
6. Toribio-Mateas, M. A., Bester, A. & Klimenko, N. Impact of plant-based meat alternatives on the gut microbiota of consumers: A real-world study. *Foods* **10**, 1–26 (2021).
7. Ramachandraiah, K. Potential development of sustainable 3d-printed meat analogues: A review. *Sustainability (Switzerland)* **13**, 1–20 (2021).
8. Simsa, R. *et al.* Extracellular heme proteins influence bovine myosatellite cell proliferation and the color of cell-based meat. *Foods* **8**, 521 (2019).
9. Hadi, J. & Brightwell, G. Safety of alternative proteins: Technological, environmental and regulatory aspects of cultured meat, plant-based meat, insect protein and single-cell protein. *Foods* **10**, 1226 (2021).
10. Igene, J. O., King, J. A., Pearson, A. M. & Gray, J. I. Influence of heme pigments, nitrite, and nonheme iron on development of warmed-over flavor (WOF) in cooked meat. *J. Agric. Food Chem.* **27**, 838–842 (1979).
11. Weickert, M. J., Pagratis, M., Glascock, C. B. & Blackmore, R. A mutation that improves soluble recombinant hemoglobin accumulation in Escherichia coli in heme excess. *Appl. Environ. Microbiol.* **65**, 640–647 (1999).
12. Villarreal, D. M. *et al.* Enhancement of recombinant hemoglobin production in Escherichia coli BL21(DE3) containing the Plesiomonas shigelloides heme transport system. *Appl. Environ. Microbiol.* **74**, 5854–5856 (2008).
13. Hoffman, S. J. *et al.* Expression of fully functional tetrameric human hemoglobin in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 8521–8525 (1990).
14. Hoyt, M. A. & Francisco, S. *Expression Constructs and Methods of Genetically Engineering Methylotrophic Yeast* Vol. 2 (Google Patents, 2020).
15. Spohner, S. C., Schaum, V., Quitmann, H. & Czermak, P. Kluyveromyces lactis: An emerging tool in biotechnology. *J. Biotechnol.* **222**, 104–116 (2016).
16. Shao, Y. *et al.* High-level secretory production of leghemoglobin in *Pichia pastoris* through enhanced globin expression and heme biosynthesis. *Biores. Technol.* **363**, 127884 (2022).
17. Ishchuka, O. P. *et al.* Genome-scale modeling drives 70-fold improvement of intracellular heme production in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* **119**, 1–9 (2022).
18. Filippo, M. D., Damiani, C. & Pescini, D. GPRuler: Metabolic gene-protein-reaction rules automatic reconstruction. *PLoS Comput. Biol.* **17**, 1–25 (2021).
19. Monk, J. M. *et al.* iML1515, a knowledgebase that computes *Escherichia coli* traits. *Nat. Biotechnol.* **35**, 904–908 (2017).
20. Lu, H. *et al.* A consensus *S. cerevisiae* metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. *Nat. Commun.* https://doi.org/10.1038/s41467-019-11581-3 (2019).
21. Stalidzans, E., Seiman, A., Peebo, K., Komasilovs, V. & Pentjuss, A. Model-based metabolism design: Constraints for kinetic and stoichiometric models. *Biochem. Soc. Trans.* **46**, 261–267 (2018).
22. Pentjuss, A. *et al.* Biotechnological potential of respiring *Zymomonas mobilis*: A stoichiometric analysis of its central metabolism. *J. Biotechnol.* **165**, 1–10 (2013).
23. Pentjuss, A. *et al.* Model-based biotechnological potential analysis of Kluyveromyces marxianus central metabolism. *J. Ind. Microbiol. Biotechnol.* **44**, 1177–1190 (2017).
24. Sajid, M., Stone, S. R. & Kaur, P. Transforming traditional nutrition paradigms with synthetic biology driven microbial production platforms. *Curr. Res. Biotechnol.* **3**, 260–268 (2021).
25. Komasilovs, V., Pentjuss, A., Elsts, A. & Stalidzans, E. Total enzyme activity constraint and homeostatic constraint impact on the optimization potential of a kinetic model. *BioSystems* **162**, 128–134 (2017).
26. Elsts, A., Pentjuss, A. & Stalidzans, E. SpaceScanner: COPASI wrapper for automated management of global stochastic optimization experiments. *Bioinformatics* **33**, 2966–2967 (2017).
27. Choi, S., Song, C. W., Shin, J. H. & Lee, S. Y. Biorefineries for the production of top building block chemicals and their derivatives. *Metab. Eng.* **28**, 223–239 (2015).
28. Kalnenieks, U. *et al.* Improvement of acetaldehyde production in *Zymomonas mobilis* by engineering of its aerobic metabolism. *Front. Microbiol.* **10**, 1–16 (2019).
29. Lopes, H. & Rocha, I. Genome-scale modeling of yeast: Chronology, applications and critical perspectives. *FEMS Yeast Res.* **17**, 1–14 (2017).
30. Kerkhoven, E. J., Lahtvee, P. J. & Nielsen, J. Applications of computational modeling in metabolic engineering of yeast. *FEMS Yeast Res.* https://doi.org/10.1111/1567-1364.12199 (2015).
31. Von Kamp, A. & Klamt, S. Growth-coupled overproduction is feasible for almost all metabolites in five major production organisms. *Nat. Commun.* **8**, 1–10 (2017).
32. Portnoy, V. A., Bezdan, D. & Zengler, K. Adaptive laboratory evolution-harnessing the power of biology for metabolic engineering. *Curr. Opin. Biotechnol.* **22**, 590–594 (2011).
33. Otero, J. M. *et al.* Industrial systems biology of *Saccharomyces cerevisiae* enables novel succinic acid cell factory. *PLoS ONE* **8**, 1–10 (2013).
34. Cregg, J. M., Cereghino, J. L., Shi, J. & Higgins, D. R. Recombinant protein expression in *Pichia pastoris*. *Appl. Biochem. Biotechnol. B Mol. Biotechnol.* **16**, 23–52 (2000).
35. Cereghino, G. P. L., Cereghino, J. L., Ilgen, C. & Cregg, J. M. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. *Curr. Opin. Biotechnol.* **13**, 329–332 (2002).
36. Macauley-Patrick, S., Fazenda, M. L., McNeil, B. & Harvey, L. M. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast* **22**, 249–270 (2005).
37. Li, P. *et al.* Expression of recombinant proteins in *Pichia pastoris*. *Appl. Biochem. Biotechnol.* **142**, 105–124 (2007).
38. Gurramkonda, C. *et al.* Simple high-cell density fed-batch technique for high-level recombinant protein production with *Pichia pastoris*: Application to intracellular production of Hepatitis B surface antigen. *Microb. Cell Fact.* **8**, 1–8 (2009).
39. Dagar, V. K. & Khasa, Y. P. Combined effect of gene dosage and process optimization strategies on high-level production of recombinant human interleukin-3 (hIL-3) in *Pichia pastoris* fed-batch culture. *Int. J. Biol. Macromol.* **108**, 999–1009 (2018).
40. Bolmanis, E., Grigs, O., Kazaks, A. & Galvanauskas, V. High-level production of recombinant HBcAg virus-like particles in a mathematically modelled *P. pastoris* GS115 Mut+ bioreactor process under controlled residual methanol concentration. *Bioprocess Biosyst. Eng.* **45**, 1447–1463 (2022).
41. Hansen, A., Choudhary, D. & Ajittvarma, P. *Rhizobium Biology and Biotechnology* Vol. 50 (Springer International Publishing, 2017).

42. Tomàs-Gamisans, M., Ferrer, P. & Albiol, J. Integration and validation of the genome-scale metabolic models of *Pichia pastoris*: A comprehensive update of protein glycosylation pathways, lipid and energy metabolism. *PLoS ONE* **11**, 1–24 (2016).
43. Heirendt, L. *et al.* Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat. Protocols* **14**, 639–702 (2019).
44. Looser, V. *et al.* Cultivation strategies to enhance productivity of *Pichia pastoris*: A review. *Biotechnol. Adv.* **33**, 1177–1193 (2014).
45. Feist, A. M. & Palsson, B. O. The biomass objective function. *Curr. Opin. Microbiol.* **13**, 344–349 (2010).
46. Schulz, C., Kumelj, T., Karlsen, E. & Almaas, E. Genome-scale metabolic modelling when changes in environmental conditions affect biomass composition. *PLoS Comput. Biol.* **17**, 1–22 (2021).
47. Nishimura, A. *et al.* Longevity regulation by proline oxidation in yeast. *Microorganisms* **9**, 1650 (2021).
48. Huynh, T. Y. L., Zareba, I., Baszanowska, W., Lewoniewska, S. & Palka, J. Understanding the role of key amino acids in regulation of proline dehydrogenase/proline oxidase (prodh/pox)-dependent apoptosis/autophagy as an approach to targeted cancer therapy. *Mol. Cell. Biochem.* **466**, 35–44 (2020).
49. Kitamura, S., Shimizu, H. & Toya, Y. Identification of a rate-limiting step in a metabolic pathway using the kinetic model and in vitro experiment. *J. Biosci. Bioeng.* **131**, 271–276 (2021).
50. Liu, H., Chen, Z., Zhang, J. & Liu, D. Elementary mode analysis for the rational design of efficient succinate conversion from glycerol by *Escherichia coli*. *J. Biomed. Biotechnol.* **2010**, 1–14 (2010).
51. Piper, M. D., Hong, S. P., Ball, G. E. & Dawes, I. W. Regulation of the balance of one-carbon metabolism in Saccharomyces cerevisiae. *J. Biol. Chem.* **275**, 30987–30995 (2000).
52. Jordá, T. & Puig, S. Regulation of Ergosterol Biosynthesis in Saccharomyces cerevisiae. *Genes* **11**, 795 (2020).
53. Protchenko, O. & Philpott, C. C. Regulation of intracellular heme levels by HMX1, a homologue of heme oxygenase, Saccharomyces cerevisiae. *J. Biol. Chem.* **278**, 36582–36587 (2003).
54. Looser, V. *et al.* Effects of glycerol supply and specific growth rate on methanol-free production of CALB by *P. pastoris*: functional characterisation of a novel promoter. *Appl. Microbiol. Biotechnol.* **101**, 3163–3176 (2017).
55. Tomàs-Gamisans, M., Ødum, A. S. R., Workman, M., Ferrer, P. & Albiol, J. Glycerol metabolism of *Pichia pastoris (Komagataella spp.)* characterised by 13 C-based metabolic flux analysis. *N. Biotechnol.* **50**, 52–59 (2019).
56. Ishchuk, O. P. *et al.* Genome-scale modeling drives 70-fold improvement of intracellular heme production in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* https://doi.org/10.1073/pnas.2108245119 (2022).
57. Sánchez, B. J. *et al.* Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Mol. Syst. Biol.* **13**, 935 (2017).
58. Lu, H. *et al.* A consensus *S. cerevisiae* metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. *Nat. Commun.* **10**, 3586 (2019).
59. Rebnegger, C. *et al.* In *Pichia pastoris*, growth rate regulates protein synthesis and secretion, mating and stress response. *Biotechnol. J.* **9**, 511–525 (2014).
60. Choi, H. S., Lee, S. Y., Kim, T. Y. & Woo, H. M. In silico identification of gene amplification targets for improvement of lycopene production. *Appl. Environ. Microbiol.* **76**, 3097–3105 (2010).
61. Orth, Jeffrey D., Ines Thiele, B. Ø. P. What is flux balance? *Nat. Biotechnol.* **28**, 245–248 (2010).
62. Invitrogen Corporation. *Pichia Fermentation Process Guidelines Overview Overview , continued*. *Progress in Botany* vol. 67 (2002).

Acknowledgements

This research was co-funded by European Regional Development Fund Project No. 1.1.1.1/21/A/044 "The development of an efficient pilot-scale leghemoglobine production technology, based on recombinant *Pichia pastoris* and *Kluyveromyces lactis* fed-batch fermentations (BioHeme)".

Author contributions

A.P. and J.L. performed genome-scale metabolic model development, E.B. and A.K. constructed the expression vector and selected clones, E.B., An.S., E.D., K.D. conceived the experiments, J.V. and Ar.S. conceived and coordinated the project. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-42865-w.

Correspondence and requests for materials should be addressed to A.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.