PNAS
www.pnas.org

Supplementary Information for

Yeast optimizes metal utilization based on metabolic network and enzyme kinetics

Yu Chen, Feiran Li, Jiwei Mao, Yun Chen, Jens Nielsen

Jens Nielsen
Email: nielsenj@chalmers.se

This PDF file includes:

Supplementary text
Figures S1 to S4
SI References

Other supplementary materials for this manuscript include the following:

Datasets S1 to S4

Supplementary Information Text

Construction of CofactorYeast. We constructed the model CofactorYeast that accounts for genome-scale metabolism, translation and cofactor binding processes, which can be divided into the following steps:

1. We expanded the genome-scale metabolic model of *Saccharomyces cerevisiae* Yeast8 (version 8.3.5)(https://github.com/SysBioChalmers/yeast-GEM) by adding several metabolic, transport and exchange reactions together with gene-protein-reaction associations (GPRs).
2. We split all reversible enzymatic reactions into forward and reverse reactions.
3. We also split reactions catalyzed by isozymes, resulting in multiple identical reactions with various isozymes being as GPRs.
4. We formulated translation reactions for all proteins in the model. The substrates of translation reactions are charged tRNA such as Ala-tRNA(Ala) while the products include uncharged tRNA such as tRNA(Ala) and the translated proteins. Note that the charged and uncharged tRNA are already metabolites in the original model Yeast8. To formulate translation reactions, we collected protein sequence data from the website (http://sgd-archive.yeastgenome.org/sequence/S288C_reference/genome_releases/S288C_reference_genome_R64-2-1_20150113.tgz). Note that we simplified the model by assuming that all proteins are translated in cytoplasm and that ribosomes do not participate in the process.
5. We collected enzyme cofactors. First, we collected cofactor information i.e., cofactor type and copy, for all proteins (not only metabolic proteins but also others) based on various databases including PDBe (https://www.ebi.ac.uk/pdbe/), BRENDA (https://www.brenda-enzymes.org/) and UniProt (https://www.uniprot.org/). Second, we manually added cofactor information from literatures. Third, we manually adjusted cofactor type and copy for homologue proteins so that they had identical cofactor information. The homologue proteins in yeast were downloaded from SGD website (https://www.yeastgenome.org/). Fourth, we excluded cofactors for some proteins since those cofactors are present as substrates/products in the reactions catalyzed by these proteins.
6. We formulated cofactor binding reactions. In this study we focused on eight metal ions and iron compounds as enzyme cofactors and therefore only formulated cofactor binding reactions for them. Note that the coefficients of cofactors were determined according to the collected cofactor copy data. As all proteins were assumed to be synthesized in cytoplasm, we also assumed that cofactor binding reactions were in cytoplasm. In addition, we assumed that iron and iron compounds such as iron-sulfur clusters and heme, which are produced in mitochondrion, can directly bind on proteins in cytoplasm.
7. We formulated enzyme formation reactions. To do so, we collected protein stoichiometry information from PDBe database (https://www.ebi.ac.uk/pdbe/) as well as Complex Portal website (https://www.ebi.ac.uk/complexportal/home).
8. We then formulated enzyme dilution reactions, representing the dilution of enzymes to daughter cells during cell division.
9. We re-formulated biomass equation of Yeast8. As we formulated dilution of all modeled proteins, accounting for 46% of total proteome by mass based on PAXdb database (https://pax-db.org/), we accordingly decreased the coefficient of protein in the biomass equation to 0.54, which was 1 in the original model Yeast8. In addition, we added unmodeled cofactor into the biomass equation to account for the cofactors binding onto unmodeled proteins, which could be estimated based on PAXdb abundances of the unmodeled proteins and cofactor copy numbers of those proteins.
10. We added a dummy protein in the model to fill up the total proteome in cases of low metabolic fluxes, which has average amino acid and cofactor composition.
11. To implement the parameter θ, we duplicated metabolic reactions catalyzed by cofactor-containing enzymes but marked the reaction IDs with "_withoutcofactor", and then formulated synthesis and dilution reactions for the enzyme without adding cofactors. This enables another parallel coupling constraint for the enzymes losing cofactors.
12. We imposed a fixed constraint of 0.2116 g/gCDW, i.e., 46%*0.46 g/gCDW, on modeled proteome, in which 46% is the fraction of modeled proteome mass in the total proteome

mass and 0.46 is the fraction of protein in biomass with the unit of g/gCDW. In addition, we imposed an upper bound of 0.1 g/gCDW on the mitochondrial proteome with the latter being approximately estimated based on published data (1–3).

13. The turnover rate is essential in coupling constraints. We downloaded a dataset of turnover rates of all organisms from BRENDA database, which excluded mutant enzymes and included only the maximal value in case that multiple values are available for a given substrate and organism. To assign turnover rate for a metabolic reaction, we used the EC number of the reaction to search in the dataset. Note that we selected the one with the highest confidence score (CS) when multiple values are available. We used the following criteria to determine CS. a) If the turnover rate was manually assigned then CS was 5; b) If both substrate and organism were matched in the dataset then CS was 4, and we assigned the maximal value if multiple values are available; c) If only organism was matched but not substrate then CS was 3, and we used the median of all turnover rates within the organism. d) If only substrate was matched but not organism then CS was 2, and we used the median of all turnover rates within matched substrate. e) If neither organism nor substrate was matched, we then used the median of all available values of within the EC number and determine CS to be 1; f) If no turnover rate was available for the EC number then CS was 0, and we assumed the turnover rate to be median of all assigned values. Note that the turnover rate should be adjusted based on the protein stoichiometry information, e.g., the turnover rate should time 2 for a dimer enzyme, and that the minimal turnover rate among subunits was selected for a complex when its subunits had various turnover rates. The turnover rates used in the model are available at https://github.com/SysBioChalmers/CofactorYeast/blob/master/kcat.xlsx.

Simulations of CofactorYeast. We performed different types of simulations with CofactorYeast, including a) growth on eight carbon sources, b) the Crabtree effect, c) growth on different carbon, nitrogen, phosphorus and sulfur sources, d) reduced availability of metal irons, nutrients and oxygen, and e) iron deficiency. Note that in some of the simulations, we should find the maximal growth rate, which is however an input of the linear programming as the growth rate has been integrated into coupling constraints. This can be addressed by a binary search workflow, which can search for the maximal growth rate that leads to a feasible solution. Note that we used Soplex 4.0.0 (https://soplex.zib.de/) to solve linear programming in this study. All the codes of simulations are available at https://github.com/SysBioChalmers/CofactorYeast.

In addition, we performed sensitivity analysis of iron-containing proteins and reduced cost analysis of amino acid uptake. Regarding the former, we performed the simulations by fixing the iron uptake rate at 50% of the reference value and using the θ value of 0.5, and meanwhile removed iron and iron compounds from cofactor binding reactions for iron-containing proteins one by one. Regarding the latter, we performed the simulations for both limited (50%) and unlimited (100%) iron uptake rate with the θ value of 0.5. In each simulation, we set the uptake rate of an amino acid at 0.01 mmol/gCDW/h and searched for the maximal growth rate. The reduced cost value can be calculated as $(\mu_i - \mu_{ref})/0.01$, in which μ_i is the simulated maximal growth rate with feeding the amino acid while μ_{ref} is the one without feeding any amino acid.

Simulations of *p*-HCA production. To investigate *in silico* production of *p*-HCA in this study, we expanded CofactorYeast to account for *p*-HCA biosynthesis and then performed simulations with the expanded model.

To expand CofactorYeast, we collected metabolic reactions and enzyme information for the *p*-HCA biosynthesis pathways based on the original study (4). We then formulated translation reactions and cofactor binding reactions for newly added enzymes based on protein sequence and cofactor information, which were collected from UniProt and PDBe databases as well as the original study (4). We also formulated reactions in the type of "_withoutcofactor" to implement the parameter θ. In addition, we assigned turnover rates for all reactions based on BRENDA and UniProt databases.

Next, we used the expanded model to simulate the effects of growth rate and iron uptake rate on *p*-HCA production. To this end, we maximized the *p*-HCA production rate for combinations of growth rates and upper bounds of the iron uptake rate. In all simulations of *p*-HCA production we set the θ value of 0.5. The iron usage of each protein can be estimated by iron copy per protein times the simulated protein concentration while iron usage of a pathway was just the sum of iron usage of all proteins involved in the pathway. In addition, the total iron usage can be calculated as iron uptake rate over growth rate.

Experimental validations of *p*-HCA production. We cultured a *p*-HCA producer strain under different levels of iron limitation conditions to test the effect of iron on *p*-HCA production. Strain, media, culture conditions and analytical methods are detailed below.

S. cerevisiae QL01 (4) was used as the *p*-HCA producer strain in this study, and the genotype is *MATa ura3-52 can1Δ::cas9-natNT2 TRP1 LEU2 HIS3 XII-2::(GPM1p- AtPAL2-FBA1t)+(TDH3p-AtC4H-CYC1t)+(tHXT7p-AtATR2- pYX212t)+(PGK1p-CYB5-ADH1t)*.

S. cerevisiae QL01 was cultivated in minimal medium (7.5 g L^{-1} $(NH_4)_2SO_4$, 14.4 g L^{-1} KH_2PO_4, 0.5 g L^{-1} $MgSO_4 \cdot 7H_2O$, and 20 g L^{-1} glucose), 2 mL L^{-1} trace metal (3.0 g L^{-1} $FeSO_4 \cdot 7H_2O$, 4.5 g L^{-1} $ZnSO_4 \cdot 7H_2O$, 4.5 g L^{-1} $CaCl_2 \cdot 2H_2O$, 0.84 g L^{-1} $MnCl_2 \cdot 2H_2O$, 0.3 g L^{-1} $CoCl_2 \cdot 6H_2O$, 0.3 g L^{-1} $CuSO_4 \cdot 5H_2O$, 0.4 g L^{-1} $Na_2MoO_4 \cdot 2H_2O$, 1.0 g L^{-1} H_3BO_3, 0.1 g L^{-1} KI, and 19.0 g L^{-1} $Na_2EDTA \cdot 2H_2O$), and 1mL L^{-1} vitamin solutions (0.05 g L^{-1} D-biotin, 1.0 g L^{-1} D-pantothenic acid hemicalcium salt, 1.0 g L^{-1} thiamin–HCl, 1.0 g L^{-1} pyridoxin–HCl, 1.0 g L^{-1} nicotinic acid, 0.2 g L^{-1} 4-aminobenzoic acid, and 25.0 g L^{-1} myo-inositol) (5) supplemented with 60 mg L^{-1} uracil.

Three biological replicates were inoculated in tubes with 2 mL minimal medium at 30 °C with 220 rpm agitation for 24 h. Then, the precultures were inoculated to the initial OD_{600} of 0.02 in 20 mL minimal medium containing 0 μM, 50 μM or 100 μM bathophenanthrolinedisulfonic acid disodium salt hydrate (Sigma-Aldrich, St. Louis, MO, USA) in the 100 mL unbaffled shake-flasks. The cells were cultured at 30 °C with 220 rpm agitation for 72 h. During the cultured process, samples were collected at 12 h and 20 h for OD_{600} measurement and residual glucose, *p*-HCA detection. 1 OD_{600} = 0.65 gCDW L^{-1} was used to estimate biomass concentration.

For residual glucose detection, the collected culture samples were centrifuged at 12,000 rpm for 10 min and filtered through 0.45 μm syringe filter. The supernatants were then quantified by high performance liquid chromatography (HPLC) (Thermo Fisher Scientific, CA, USA) equipped with an Aminex HPX-87G column (Bio-Rad, Hercules, CA, USA) and a UV and RI detector. 5 mM H_2SO_4 was used as the mobile phase and the column was kept at 45 °C with a flow rate of 0.6 mL min^{-1} for 35 min.

For *p*-HCA detection, the collected samples were thoroughly mixed with an equal volume of absolute ethanol (100% v/v) and centrifuged at 12,000 rpm for 10 min. The supernatants were used for *p*-HCA quantification on HPLC (Thermo Fisher Scientific, Waltham, MA, USA) equipped with a Discovery HS F5 150 mm x 46 mm column (particle size 5 μm) (Sigma-Aldrich, St. Louis, MO, USA). The HPLC analysis was carried out with 10 mM ammonium formate (pH 3.0, adjusted by formic acid) (Solvent A) and acetonitrile (Solvent B) as the eluents. The eluent flow rate was 1.5 mL min^{-1}. The gradient program started with 5% of solvent B (0–0.5 min), increased linearly from 5% to 60% (0.5–9.5 min), then the second linear increase from 60% to 100% (9.5–10.5 min) and maintained at 100% for 0.5 min (10.5–11 min), at last a linear decrease from 100% to 5% (11–11.5 min) and maintained at 5% for 1.5 min (11.5–13 min). *p*-HCA was detected by absorbance at 304 nm with a retention time of 6.3 min.

The glucose and *p*-HCA concentrations were calculated from the standard curves, and both glucose and *p*-HCA standards were purchased from Sigma-Aldrich.

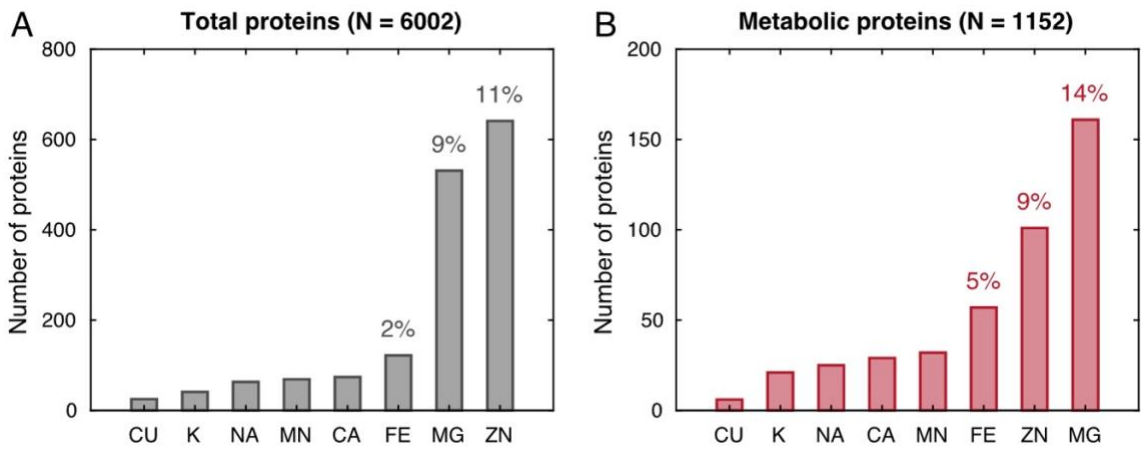

Fig. S1. Statistics of metal ions in yeast. (A) Distribution of metal ions in total proteins. The protein count was from https://www.ncbi.nlm.nih.gov/genome/15?genome_assembly_id=22535. (B) Distribution of metal ions in metabolic proteins. The number of metabolic proteins was from the model CofactorYeast.

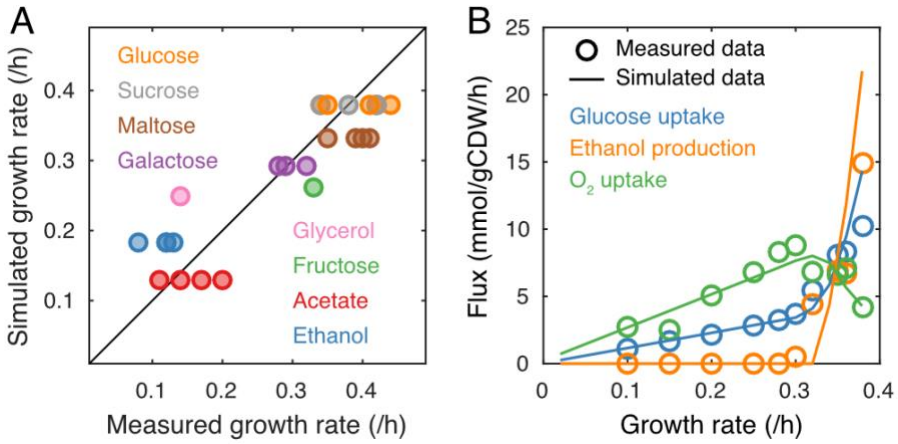

Fig. S2. Simulations of CofactorYeast as a proteome-constrained model. (A) Simulated growth rates on various carbon sources compared with measured data (6, 7). (B) Simulated glucose uptake rates, ethanol production rates and oxygen uptake rates with changing growth rates compared with measured data (8).

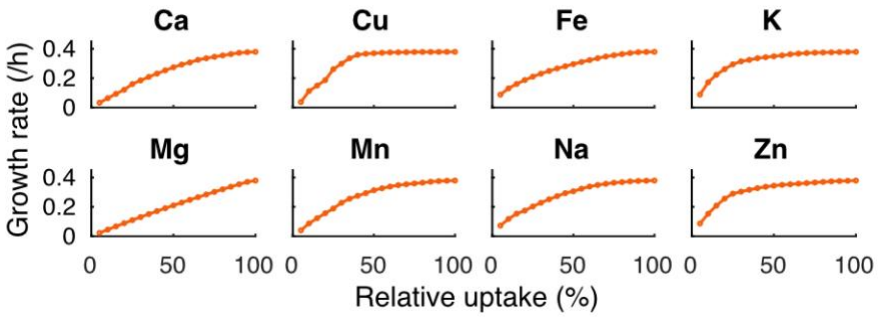

Fig. S3. Simulated growth rates with reduced uptake rates of metal ions.

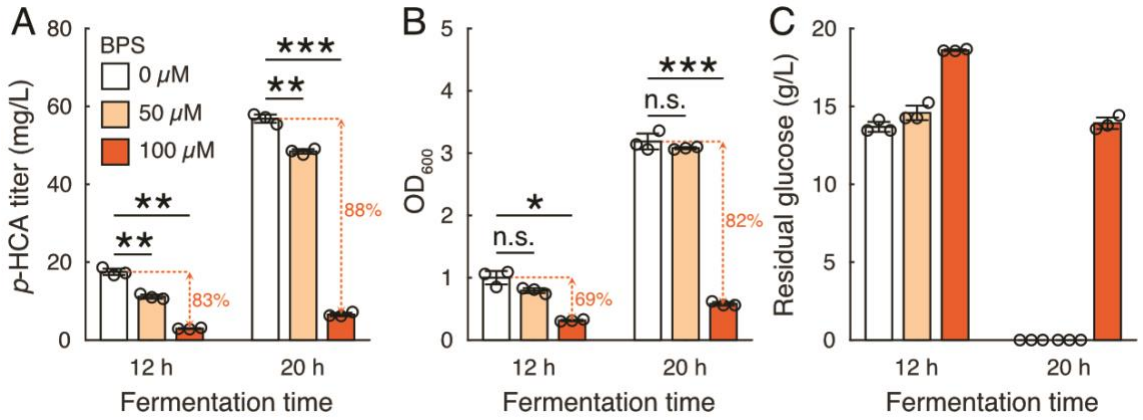

Fig. S4. Experiments of *p*-HCA production. (A) *p*-HCA titers, (B) OD_{600} and (C) residual glucose concentrations of samples at fermentation time of 12 h and 20 h upon different levels of iron limitation. Statistical analysis was performed by using Student's t test (two-tailed; two-sample unequal variance; n.s.: $p \geq 0.05$; *: $0.01 \leq p < 0.05$; **: $0.001 \leq p < 0.01$; ***: $p < 0.001$). All data represent the mean of n = 3 biologically independent samples and error bars show standard deviations.

Dataset S1 (separate file). Newly added metabolic reactions to Yeast8 model.

Dataset S2 (separate file). A collected dataset of cofactors in the yeast proteome.

Dataset S3 (separate file). Experimentally measured abundances of metal ions from different published studies.

Dataset S4 (separate file). Information to expand CofactorYeast to account for biosynthesis of *p*-HCA, including metabolic reactions, protein sequences, cofactor information and turnover rates.

SI References

1. Lahtvee P-J, et al. (2017) Absolute Quantification of Protein and mRNA Abundances Demonstrate Variability in Gene-Specific Translation Efficiency in Yeast. *Cell Syst* 4(5):495-504.e5.
2. Campbell K, et al. (2020) Building blocks are synthesized on demand during the yeast cell cycle. *Proc Natl Acad Sci*:201919535.
3. Yu R, et al. (2020) Nitrogen limitation reveals large reserves in metabolic and translational capacities of yeast. *Nat Commun* 11(1):1881.
4. Liu Q, et al. (2019) Rewiring carbon metabolism in yeast for high level production of aromatic chemicals. *Nat Commun* 10(1). doi:10.1038/s41467-019-12961-5.
5. Verduyn C, Postma E, Scheffers WA, Van Dijken JP (1992) Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* 8(7):501–517.
6. van Dijken J., et al. (2000) An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enzyme Microb Technol* 26(9–10):706–714.
7. Tyson CB, Lord PG, Wheals AE (1979) Dependency of size of *Saccharomyces cerevisiae* cells on growth rate. *J Bacteriol* 138(1):92–8.
8. van Hoek P, et al. (1998) Effects of pyruvate decarboxylase overproduction on flux distribution at the pyruvate branch point in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 64(6):2133–40.