bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Whole-cell modeling in yeast predicts compartment-specific proteome constraints that drive metabolic strategies

Ibrahim E. Elsemman[1,7,#], Angelica Rodrigues Prado[2,3,#], Pranas Grigaitis[2,#], Manuel Garcia Albornoz[4], Victoria Harman[5], Stephen Holman[5], Johan van Heerden[2], Frank J. Bruggeman[2], Mark M.M. Bisschops[3], Nikolaus Sonnenschein[1], Simon Hubbard[4], Rob Beynon[5], Pascale Daran-Lapujade[3], Jens Nielsen[1,6,*], and Bas Teusink[2,*]

[1] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK2800 Lyngby, Denmark

[2] Systems Biology Lab, Amsterdam Institute of Molecular and Life Sciences, Vrije Universiteit Amsterdam, The Netherlands

[3] Department of Industrial Microbiology, Technical University Delft, Delft, The Netherlands

[4] Division of Evolution & Genomic Sciences, University of Manchester, Manchester, UK

[5] Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

[6] Department of Biology and Biological Engineering, Chalmers University of Technology, SE41296 Gothenburg, Sweden

[7] Current address: Department of Mathematics, Faculty of Science, Assiut University, Assiut, Egypt

[#] These authors contributed equally to this work

[*] To whom correspondence should be addressed: nielsenj@chalmers.se (JN), b.teusink@vu.nl (BT)

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

When conditions change, unicellular organisms rewire their metabolism to sustain cell maintenance and cellular growth. Such rewiring may be understood as resource re-allocation under cellular constraints. Eukaryal cells contain metabolically active organelles such as mitochondria, competing for cytosolic space and resources, and the nature of the relevant cellular constraints remain to be determined for such cells. Here we developed a comprehensive metabolic model of the yeast cell, based on its full metabolic reaction network extended with protein synthesis and degradation reactions (16304 reactions in total). The model predicts metabolic fluxes and corresponding protein expression by constraining compartment-specific protein pools and maximising growth rate. Comparing model predictions with quantitative experimental data revealed that under glucose limitation, a mitochondrial constraint limits growth at the onset of ethanol formation - known as the Crabtree effect. Under sugar excess, however, a constraint on total cytosolic volume dictates overflow metabolism. Our comprehensive model thus identifies condition-dependent and compartment-specific constraints that can explain metabolic strategies and protein expression profiles from growth rate optimization, providing a framework to understand metabolic adaptation in eukaryal cells.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Macromolecular synthesis and energy conservation by metabolism underlies cellular maintenance, growth and fitness. Unicellular organisms such as yeasts generally display a great variety of metabolic strategies that lead to competitive fitness across conditions[1]. The associated reprogramming of metabolism between such metabolic strategies is of key interest in biotechnology and biomedical research.

One well-known example is "overflow" metabolism, in which under aerobic conditions not all substrate is fully oxidized but secreted as by-products. In cancer cells it is referred to as the Warburg effect: enhanced glycolytic activity with lactate as byproduct at the expense of respiration[2]. The same phenomenon is known as the Crabtree effect in *Saccharomyces cerevisiae* (Baker's yeast)[3]. At sugar limitation, in the presence of oxygen, yeast respires glucose completely to CO_2 for ATP generation; at sugar excess, it displays respirofermentative metabolism, where respiration is combined with ethanol formation (alcoholic fermentation). The extent to which these two metabolic strategies are used can be titrated in a glucose-limited chemostat: at a specific, "critical", dilution (=growth) rate, ethanol formation starts and increases linearly with growth rate[4]. Other microorganisms show similar behaviour[5]: for example, *E. coli* produces acetate at higher growth rates at the expense of respiration[6].

In the last decade, a theoretical framework has been developed that can explain why cells shift metabolic strategies upon environmental or gene-expression perturbations[5,7–10]. It is essentially based on the catalytic benefits of proteins, and their associated costs[11]. These costs comprise competition for building blocks, energy and synthesis machineries; and for space in cellular compartments. Two key features of this resource allocation paradigm can explain metabolic adaptations. First, cellular compartments are or can become "full", which is the case when they are fully occupied with (maximally) active proteins, and an increase in one protein has to come at the expense of another. This was postulated as a phenomenological rule based on experimental observations[12], but also follows naturally from growth-rate maximization[13]: at the maximum growth rate that is attainable at any condition, at least one compartment must be fully occupied and thus actively limit that growth rate. Second, cells allocate their limited resources

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

for protein synthesis according to their demands[14,15]. Consequently, fractions of needed proteins vary with growth rate within compartments whose protein content is bounded, and this can lead to "active proteome constraints" related to full compartments.

Within this framework, the onset of overflow metabolism was explained by the smaller protein cost of generating ATP through fermentation than respiratory pathways[6,7]; this becomes important under fast growth, when biosynthesis and ribosome demands are high and thus require large proteome fractions. Earlier work suggests that the proteome-constrained resource allocation paradigm, which was largely developed for *E. coli*, may also be a powerful perspective for regulation of eukaryal yeast metabolism, such as ribosome biosynthesis[16], and growth on different sugars[17]. However, a key feature of the metabolism of a eukaryal cell is the presence of metabolically active organelles, most prominently mitochondria. Each organelle introduces two new compartments (intra-organellar space and membrane), and how these compartments impacts adaptation of metabolism, and which compartments become limiting under different conditions, is an open question.

Moreover, despite the wealth of experimental data on *Saccharomyces cerevisiae*, a comprehensive, quantitative, data set in which growth rate is systematically varied and both fluxes and protein expression levels are measured, which are needed to validate resource allocation predictions, are still rare (see however, some recent studies[16,18]).

Here we generated such high-quality comprehensive data sets, and in parallel developed the most detailed and comprehensive, compartmentalized and quantitative model of metabolism and protein synthesis of yeast. The model can compute the costs and benefits of protein expression and translocation; It can be used to interpret or predict experimentally determined changes in growth rate, protein expression and metabolic fluxes as a result of growth rate optimization through resource allocation into different, compartmentalized, proteome fractions. Comparison of the model predictions with the data gives unprecedented insight into our physiological understanding of this important model organism.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Results

Construction of a comprehensive proteome-constrained yeast model

We extended an existing[19] metabolic genome-scale metabolic model of yeast (GEM), by coupling metabolic fluxes to the synthesis of the catalysing enzyme and added constraints on protein concentrations, expressed as protein fractions of the total proteome (Fig. 1a). We refer to the resulting model as proteome-constrained Yeast (pcYeast). Earlier GEM-based approaches exist that incorporate resource allocation, and for yeast, these considered constraints on enzyme activities and total protein content[17,20–23], whereas for *E. coli* there were added constraints and reactions associated with transcription and translation[9]. Others considered membrane-area constraints and limitations of protein allocation to specific pathways[8,24]. We combined all these extensions (see Supplementary Notes for detailed information) to make pcYeast: a next-generation yeast GEM and computable knowledge base that incorporates protein expression, translation, folding, translocation and degradation at genome-scale for a compartmentalized, eukaryal, organism. In our current model, we consider the protein compartments most relevant for central metabolism: plasma membrane, cytosol, mitochondrion, and mitochondrial membrane. Other cell compartments such as the nucleus or endoplasmic reticulum are not (yet) specified explicitly - but do occupy volume in the cytosol.

The cellular proteome was divided into metabolically active, ribosomal, and unspecified (UP) proteins. The UP fraction is cytosolic, has an average amino acid composition and is added to always maintain a constant protein density in the cytosol. It has a minimum expression level estimated from the experimental proteomics data (Supplementary Notes). The minimal UP fraction represents growth-rate independent structural, signalling and "household" proteins. Higher expression of UP than minimal represent both unspecified anticipatory proteins, or metabolic proteins that do not carry flux – including the unsaturated fraction of flux-carrying enzymes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Metabolic enzymes are assigned to a specific compartment, either cytosol, plasma membrane, mitochondrial matrix or inner mitochondrial membrane; Mitochondrial proteins require additional protein transport complexes[25]. For each protein, we comprehensively modelled synthesis and degradation processes, which are responsible for the largest fraction of cellular energy usage. Our model includes 1,523 proteins, whose life cycles are described by 16,304 reactions that include translation initiation, elongation and termination factors, ribosomal assembly factors, protein-specific folding by chaperones and degradation reactions, as well as 5'UTR-length dependent energetic costs for translation initiation (Table 1, Supplementary Notes).

We applied three classes of constraints that couple metabolic fluxes and peptides synthesis rates (Fig. 1b and Supplementary Notes for details). The *enzyme capacity constraint* sets the minimal enzyme synthesis rate required to achieve a certain metabolic flux. Thus, all metabolically-active proteins are modeled to work at their maximal rate and are minimally expressed; the unspecified protein was used to maintain protein density. In this way we prevent choices about unknown regulatory and kinetic mechanisms that may affect the activity of enzymes; rather we use the deviation between predicted minimal and measured actual protein expression levels to indicate such effects. The total enzyme synthesis rate is constrained by the abundance of ribosomes through a *ribosome capacity constraint*, for both cytosol and mitochondria. Finally, we added *compartment-specific constraints* on the proteome, for the cytosol, the plasma membrane, and the mitochondrial matrix and inner membrane, (Fig. 1b). The values for these constraints are based on independent literature data or were fitted to experimental data (as explicified in Supplementary Notes) and the values are either fixed or growth-rate dependent, depending on the nature of the constraint.

The steady-state metabolite balances, the enzyme synthesis and degradation balances, and the compartment-specific proteome constraints together specify a linear program with its fluxes as optimisation variables, provided the growth rate is treated as a parameter. We use a binary search algorithm to find the maximum growth rate where the linear programming

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

problem becomes just infeasible; the model returns all the flux values associated with the maximal feasible growth rate.

Calibrating the model against experimental data

We performed a series of experiments for collection of high-quality datasets of fluxes and protein levels, used either as model input or for comparison with model predictions. We used glucose-limited continuous cultures operated at dilution rates close to the critical dilution rate for ethanol formation, to capture proteome change upon the onset of overflow metabolism. Additionally, we varied the growth rate in pH-controlled batch experiments, either with different sugar quality or through translation inhibition. We measured fluxes, including O_2 and CO_2 fluxes (Supplementary Dataset 1), which combined with biomass measurements, allowed to estimate the so-called maintenance parameters, i.e. ATP usage that is not explicitly accounted for in the model (Supplementary Notes). Label-free proteome quantification allowed us to reliably estimate proteome fractions of around 3000 of the 6000 proteins (Supplementary Datasets 2, 3, and 4).

Parameters associated with translation strongly affected our model outcomes, and we used published quantitative proteomics data[16] to estimate parameters for protein translation, such as the elongation rate (Supplementary Notes). According with experimental reports we assumed a constant inactive fraction of ribosomes, and a fixed saturation of the actively translating ribosomes[16,27] and were hereby able to describe the growth-rate dependent ribosome mass fraction with the model (Fig. 1c). As evidence for correctly capturing the costs of protein synthesis, we correctly predicted the effect of over-expressing mCherry, an unneeded, "gratuitous" protein, on the specific growth rate (Fig. 1d).

The model predicts shifts in metabolic strategies

We subsequently used the model to analyse yeast's physiological response to different levels of glucose availability. Traditional Flux Balance Analysis computes continuous chemostat

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

cultures by minimizing glucose uptake rate at fixed growth (=dilution) rate[28]. Here we simulated glucose availability by varying the degree of saturation of the glucose transporter. We needed to constrain the maximal expression level of the glucose transport system, based on literature data (Supplementary Notes), as leaving expression free to occupy available membrane space led to unrealistically high expression levels and overestimation of growth rate at subsaturating glucose levels. We subsequently computed the maximal feasible growth rate and compared model predictions with published data[29], and from our glucose-limited chemostat cultures (growth rates between 0.2 – 0.34 h^{-1}). We also included our data from batch cultures on glucose (growth rates 0.37-0.39 h^{-1}) and on trehalose; Trehalose is a disaccharide of two glucose molecules, hydrolyzed extracelullarly[30], thus providing slow release of glucose that supports low growth rates.

The maximal feasible growth rate that the model predicted can be linked directly to the dilution rate in the chemostat, allowing comparison of model prediction and data (Fig. 2a). The (residual) glucose concentrations were calculated from documented (high) affinity of the transporters, which is close to 1 mM[31]. The resulting relationship between growth rate and residual glucose concentration fit experimental data very well (Fig. 2b), validating our expectation that we could ignore intracellular glucose[32]. Predicted biomass yield (Fig. 2c) and fluxes (Fig. 2d) corresponded well with the experimental data, as did the intracellular flux ratios from previously published ^{13}C-labeling flux analysis at three specific growth rates in glucose-limited chemostat cultures (Supplementary Figure 1). In particular, the model predicted a maximal oxygen consumption rate at dilution rates higher than 0.28 h^{-1}, at the onset of ethanol formation. Above 0.35 h^{-1}, this rate rapidly drops to the low level that is observed under glucose excess (batch) conditions. We conclude that the model can adequately predict the changes in metabolic fluxes when the growth rate is varied through the availability of glucose.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Changes in metabolic strategies are the result of proteome constraints

We used pcYeast to identify the active proteome constraints, *i.e.* the protein pools that limit growth rate, because, according to resource allocation theory, the number of active proteome constraints determines the maximal number of independent metabolic behaviors that are possible in optimal states[5,13]. For this we computed the occupancy of each protein pool: a pool that is fully occupied is indicative of an active constraint. At low growth rates, below 0.28 h^{-1}, the glucose transporter is the only proteome pool that is fully occupied (Fig 2e). With only glucose uptake as active constraint, pure respiration is the single optimal strategy. At the onset of ethanol formation, a second metabolic mode started to carry flux (for formal computation of these modes and the concomitant theory, see Supplementary Notes), and thus a second constraint must have become active. Indeed, at this growth rate the occupancy of the inner-mitochondrial membrane became maximal (Fig. 2e). Thus, the model suggests that under glucose-limited chemostat conditions, the onset of ethanol formation is caused by a limit of the mitochondrial membrane space, and hence the amount of proteins that yeast can maximally express in this compartment.

At a growth rate of 0.35 h^{-1} we found that the unspecified protein level reached its minimal value (Fig. 2e), equivalent to the cytosol being completely filled with maximally active proteins. Further growth rate increase requires higher ribosomes and biosynthetic protein fractions, which now has to come at the expense of the least proteome efficient pathway. The model confirmed earlier calculations[33] that respiration is less proteome efficient than fermentation (Supplementary Figure 2) and respiration is therefore replaced by fermentation. The model suggested therefore that at growth rate above 0.35 h^{-1} the second growth-limiting constraint was shifted from the mitochondrial proteome to the cytosolic proteome. Thus, the metabolic changes in the model, when growth rate and thus metabolic fluxes increase, are dictated by the filling up of different cellular compartments with active protein, unique for an eukaryal cell. The level of detail in our model to identify the condition-dependent, active,

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

protein-concentration constraints belonging to different compartments has so far not been provided by any other model.

Proteomics data validates model predictions

We subsequently measured protein levels with quantitative proteomics and compared them to the model predictions. We expected to underestimate most proteome fractions, because the model predict minimal protein levels required to support metabolic flux. Especially at lower growth rates where nutrient limitation is most severe, one can expect lower average enzyme saturation, and indeed we observed larger deviations between predicted minimal protein levels and measured protein fractions at low growth rates (Fig. 3a). The difference between the predicted minimal level and the data may be interpreted as a proxy for the average saturation of enzymes. We see an overall tendency that the saturation of enzymes increases with growth rate (Supplementary Figure 3). This is most prominent for the glycolytic pathway; also for amino acid biosynthesis, the protein expression is higher than expected based on metabolic activity, indicating also here a substantial undersaturation of the enzymes, as observed before for bacteria such as *E. coli*[34] and *L. lactis*[35]. We find similar patterns for other biosynthetic pathways, except for lipids (Supplementary Figure 4).

For mitochondrial proteins involved in the citric acid cycle and respiration, however, we found that predicted minimal proteome fractions were very close to the measured ones (Fig. 3a). Unless kcats of mitochondrial enzymes are systematically underestimated, this indicates that mitochondrial proteins work at higher average saturation than cytosolic proteins - and seemingly close to their maximal capacity. Regardless of absolute numbers, the saturation of the mitochondria seems rather constant, suggesting that yeast tunes the total amount of mitochondria, rather than make excess (but subsaturated) mitochondria, at least under these conditions. This may make sense, given the extra costs of mitochondrial components such as membranes, and for protein translocation of host-derived proteins during mitochondrial biogenesis, which competes for membrane space with respiratory proteins.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Upon closer inspection, we observed that at the onset of ethanol formation the total mitochondrial protein fraction started to decrease (Fig. 3b). The observed decay follows the theoretical dilution-by-growth kinetics if at that point the rate of mitochondrial biosynthesis has reached a maximum (Fig. 3b). Thus, the data suggests that the rate of mitochondrial biogenesis, rather than the maximal mitochondrial membrane area currently used by the model, may reach the host's maximal capacity at the onset of ethanol formation. When we zoom in on the mitochondrial proteome, we find that the mitochondrial ribosome fraction increased as a funtion of growth rate, and also other proteins re-allocated (Supplementary Fig. 5). Indeed, mitochondria are self-replicating entities abiding to the same resource allocation principles as the host, which even includes selection for fast replication - but obviously severely dictated by the proteins the host provides. More data related to the mitochondrial biosynthetic processes, such as mitochondrial ribosomal capacity and protein import machinery would be required to predict the maximal mitochondrial growth rate from first principles, which is outside the scope of this study. Nonetheless, the distinct changes of mitochondrial proteins at the critical dilution rate are consistent with the model prediction that a mitochondrial constraint is responsible for the onset of ethanol formation under glucose-limited conditions.

Constraints and fluxes under sugar excess conditions

We then varied growth rate (between 0.05 h^{-1} and 0.4 h^{-1}) by providing different sugars, i.e. trehalose, galactose, maltose and glucose during batch cultivation. Ethanol production was already observed on galactose, already at a growth rate of 0.16 h^{-1} so at a much lower growth rate than the critical growth rate of 0.28 h^{-1} under glucose-limited growth (Fig. 4a). Maltose showed intermediate growth rate and fluxes. Initial model simulations with a "naïve" model using the reported catalytic rates of the transporters and catabolic enzymes involved in galactose and maltose metabolism, however, resulted in predicted growth rates and fluxes not far from growth on glucose. This suggests that there are additional cost factors that were not

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

included in the model, and or that *Saccharomyces cerevisiae* is not as well adapted to these sugars.

We therefore used the model as data analysis tool to estimate possible changes in parameters that fit the observed growth rate and corresponding fluxes (see Supplementary Notes for details). Such parameter changes may be interpreted as costs for suboptimal metabolism of carbon sources other than glucose. The onset of ethanol formation at a growth rate of 0.16 h^{-1} required a combination of changes in both sugar uptake and the intracellular proteome (through the minimal UP fraction constraint): a lower sugar uptake capacity alone would be identical to lowering saturation of the transporter as was done for glucose (Fig 2), and pure respiration would have been found at 0.16 h^{-1}. Conversely, only an increase in minimal UP would have resulted in a proportional flux decrease that we also found with mCherry overexpression (or translation inhibition, Supplementary Figure 6), and more ethanol were to be found.

We had to decrease the maximal galactose uptake rate by a factor of 2.5 compared to glucose. Furthermore, an increase in minimal UP fraction was needed, to 0.49 g/g protein. To fit all fluxes optimally, we also required additional energetic costs (see Supplementary Notes), whose mechanistic underpinning remains to be explored but may be related to the reported toxicity of galactose intermediates[36]. Such a change in energetic costs were not needed to describe the data for growth on maltose: only a change in the maltose uptake rate and minimal UP fraction (of 0.34 g/g protein) were required to achieve good fit.

For maltose, a disaccharide of glucose, the reason for the required parameter changes is not clear. Only a maltose proton-symporter and a maltase protein distinguishes it from growth on glucose. The transport expression may be tightly regulated as very high maltose uptake rates can result in substrate-accelerated death[37]. For galactose, the toxicity of its intermediates[36] results in an evolutionary trade-off with growth on glucose[38]; on galactose yeast cells appear to be still prepared for growth on glucose, which may prevent them from optimal expression of proteins on galactose, as shown by expression titration experiments[39]. Indeed, laboratory

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

evolution experiments on galactose select mutations in Ras/cAMP signalling and adapted strains show increased growth rates and concomitant increased ethanol fluxes[38]. Interestingly, the direction of change points to the optimal behaviour predicted by the initial naïve model, suggesting that the pcYeast model may aid in predicting the direction of evolutionary change during laboratory evolution experiments (Supplementary Figure 7).

With the updated parameters, we identified for both sugars that the active constraints limiting growth were the sugar transport expression and the minimal UP fraction constraint (Figure 4d, Supplementary Notes, Supplementary Figure 8). These active constraints explains ethanol formation during growth on galactose even though the growth rate is lower than the critical dilution rate on glucose.

Proteomics data on sugar excess shows re-allocation of metabolic strategies

If growth rate is actively constrained by the cytosolic proteome under galactose, maltose and glucose excess conditions, it implies that all cytosolic proteins work at their maximum activity, and changes in flux must be brought about by changes in protein level. We therefore turned to proteomics again. Comparing the minimal levels of the model with experimental data, we find again that mitochondrial proteins for the TCA cycle and respiration are very similar to the predicted minimal levels required to sustain flux (Fig. 4c). Cytosolic proteins were underestimated - even at sugar excess conditions. (Note however that the expected maximal attainable activity is not likely at the maximal rate in the forward direction as product inhibition is inevitable in a chain of enzymes.)

More indicative of “a full cytosol” is that at the onset of ethanol formation (at galactose growth rate and higher) we find evidence for proportional relationships between protein and flux for high-flux carrying, pathway-grouped, proteins as a function of growth rate (Fig. 4c). This is observed even down to the individual protein level, as illustrated for glycolytic and respiratory proteins in Fig. 4d. This implies that under these conditions, enzyme saturation was constant (and maximal, we expect) and changes in flux could only be brought about by

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

corresponding changes in enzyme levels. This data illustrates how mitochondrial proteins are being traded in for glycolytic proteins needed for an enhanced fermentation and growth rate. It also confirms the model's prediction that the cytosolic proteome constraint is active during growth on these sugars.

Inhibition of translation highlights the role of environmental signaling in coordination of metabolism in yeast.

Finally we varied growth rate by translation inhibition by cyclohexamide under controlled glucose batch conditions, and again measured fluxes, growth rate and proteome profiles (Fig. 5a). Upon inhibition of translation, we found a decrease in growth rate and close to proportional decreases in glucose, ethanol and CO_2 fluxes, for both the model and the experimental data (Fig. 5b). Such behaviour is expected when one dominant constraint is active and its extent is varied (cf. glucose-limited fully respiratory growth, Fig. 2). In the case of glucose excess, the model suggested that the cytoplasmic volume was fully occupied with active proteins (minimal UP constraint was hit), and inhibition of translation required higher expression levels of ribosomes, taking away limited proteome space for growth-supporting activities.

However, experimental observations compromised this initial explanation. First, for oxygen the model also predicted a proportional increase with growth rate, but experimentally the fluxes did not change much as did the expression of enzymes involved in oxygen consumption, such as TCA cycle and oxidative phosphorylation (Fig. 5d). Moreover, the ribosomal proteome fraction increased much less with inhibition than the model predicted (Fig. 5c). Since translation inhibition in the model has the same effect as overexpression of a non-functional protein (Supplementary Figure 6), we followed the earlier observation that the inactive fraction of ribosomes could be recruited for translation, depending on the translational load[16], with only a small improvement (Supplementary Figure 9).

This suggested that either some constraint prevents the ribosomal fraction from increasing to the optimal levels predicted by the model, or the expression of ribosomes in yeast

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

is dominantly regulated by environmental nutrient signalling and less by internal cues. A dominant role of signalling in ribosomal biogenesis has been suggested before[16]. In yeast the TOR pathway appears to be the master regulator of ribosomal biosynthesis and assembly at steady-state growth[40,41]. Following the TOR-specific targets described by Kunkel[41], we find that key target proteins of this signalling pathway, including ribosomal auxiliary factors, had constant expression levels (Supplementary Figure 10 and Supplementary Table 1), supporting the dominant role of external rather than internal cues.

When we constrained ribosomal expression to the measured maximal response, ribosomal expression rapidly became the only active constraint in the model, and the proteome space that became available in the cytosol at the lower growth rates was used for increased respiration (Supplementary Figure 10). This is not observed experimentally, and our data suggest that respiration does not respond to internal cues either. In contrast the fluxes and expression of proteins involved in glycolysis and amino acid metabolism did decrease with growth rate (Fig. 5bd). This suggests that these pathways must be sensitive to internal feedback regulation, as is well known for amino acid metabolism[42]. Thus, the proportional fluxes we found for ethanol and glucose upon translation inhibition, are likely the result of control by demand[43], with lower demand at lower growth rate.

Discussion

In this work, we developed the most comprehensive model of a growing, compartmentalized, eukaryal cell to date. It includes all its known metabolic reactions, and details of the protein synthesis, degradation and transport machinery to express the enzymes. The key of our approach is the application of constraints on protein pools in the different compartments that have direct biochemical meaning and could be independently estimated from literature data. Our approach is unique in level of detail and in dealing with cellular compartmentation, in particular of the mitochondria. We furthermore generated a unique set of

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

high-quality quantitative data on both fluxes and the proteome under different, well-controlled, conditions. Through integration and comparison with the model, we provide deeper insight into the physiology of *Saccharomyces cerevisiae*.

First, we firmly established that metabolic growth strategies of yeast on glucose can be well understood from a proteome-constrained optimisation problem with growth rate as objective. Through our high resolution sampling around the critical dilution rate, we observed distinct changes in proteins exactly at the onset of ethanol formation in the glucose-limited chemostat. We also show that the active constraints that drive these changes can be different under different conditions such as batch growth on galactose - even if ethanol is made in both cases. Our approach to identify the active cellular constraints may resolve some of the discussion in current literature about *the* cause of overflow metabolism, not only in yeast but possibly also in other eukaryotes, including discussion about the Warburg effect in mammalian cells[44].

Second, the proteome constraints of the model are currently based on experimental observations, but further research could drill deeper into their origin. For example, why would the protein density in the cytosol be relatively constant; Does this balance diffusion rates with catalytic capacities[45]? Are the current morphological dimensions of a yeast cell optimal for growth rate? Recent work on selection for cell number showed that smaller cells can be readily selected for[46]. We also identified that the levels of glucose transport and that of mitochondria need to be constrained to describe the data. Why would yeast not express these components at higher levels? In the case of mitochondria, the proteomics data suggest that rather than a maximum mitochondrial membrane area and matrix volume, there is a maximal rate of mitochondrial biogenesis. Can we calculate this rate from first principles? One could imagine that an upper limit for mitochondrial “growth rate” exists if all but eight metabolic proteins need to be transported over the same membrane that must also harbour the full machinery for oxidative phosphorylation. Moreover, we focused on mitochondrial protein content, and ignored details on morphology, lipid synthesis, or possible assembly costs. Thus, a next version of the

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

model will need to address the mitochondrial transport, biosynthesis and morphology in much more detail.

In the case of glucose transport, the model suggested that further increase in glucose transporters beyond wild type expression did not increase growth rate substantially and would likely be invisible for evolution. At maximal saturation of the transporter, glucose transport expression was (just) no longer an active constraint in our model (Fig. 2e). Thus, it appears as if yeast expresses just enough glucose transporters to maximise its growth rate under glucose excess – as found in bacteria[47]. Expressing higher transport levels at lower glucose levels would enhance growth rate but may not pay off if this state is a transient towards glucose starvation, or could be outright dangerous if suddenly glucose would become available[37]. The expression level of the hexose transporters may thus have evolved to be an adaptation to dynamic environments[48]. Long-term evolution experiments in glucose-limited chemostats indeed show gene duplications of high-affinity glucose transporters[49], showing that growth limitation, and hence selection pressure, is on glucose transport under these conditions.

Third, in the case of nutrient uptake limitation, there appears to be "excess" proteome space that could be filled with anticipatory proteins or heterologous enzymes at no cost in fitness. Even though the composition of such excess proteome space cannot be predicted with our model, we were able to predict metabolic fluxes very well: in this nutrient-limited regime metabolic efficiency (ATP per glucose), not proteome efficiency (ATP per protein), determines the best growth rate strategy. This explains why Flux Balance Analysis applied to only the metabolic network has been so successful, but only under nutrient-limited conditions.

Finally, we found linear or even proportional relationships between growth rate and flux, and between flux and enzyme levels in a sugar excess (batch culture) regime. In terms of regulation analysis[50], such a regime is characterised by hierarchical regulation with absence of metabolic regulation, that is, all changes in flux are brought about by changes in enzyme levels, not their degree of saturation. For glycolysis and amino acid metabolism, the average saturation, estimated as the ratio of the predicted minimal enzyme level to the expressed enzyme level, at

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

maximal growth rate is around 0.5, incidently the level predicted as theoretical optimum for specific reaction rate[51]. In contrast, when growth is limited by glucose availability, the degree of saturation varies and the model suggests a mixture of hierarchical and metabolic regulation, as previously observed in chemostats as well[52].

To conclude, we present a mechanistic, compartmentalized, model of an eukaryal organism in full details, which can act as a valuable, computable, knowledge base. We show how it can be used to compute protein costs and identify active growth-limiting constraints, and how it can be combined with quantitative flux and proteomics data to provide unprecedented insight into cellular physiology. Finally, we show that also in eukaryal cells, metabolic strategies can be understood on the basis of growth rate optimisation under nutrient and proteome constraints. What remains to be understood is how the cell's signalling and regulatory networks manage to implement these (optimal) proteome allocation strategies.

Methods

Model development

The full description of the pcYeast model is provided as Supplementary Notes. The model codes are available per request to the authors and will be published on GitHub upon acceptance of this manuscript.

Strains and shake flask cultivation

The strain used for this study was *Saccharomyces cerevisiae* strain CEN.PK 113-7D[53]. The stocks used for the experiments were grown in 500 mL shake flask containing 100 mL of YPD medium (10 g L^{-1} of Bacto yeast extract, 20 g L^{-1} of peptone and 20 g L^{-1} of D-glucose). The culture was grown up to early stationary phase and 1 mL aliquots were stored in 20% (v/v) of glycerol at -80 °C. For chemostats, pre-cultures were grown in 500 mL shake flasks containing 100 mL of synthetic medium, the pH was set to 6.0 with 2M KOH and the medium was supplemented with

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

20 g L^{-1} of D-glucose[54]. Shake flasks with medium were inoculated with the 1 mL frozen stocks of the strain and the cultivations were performed in an orbital shaker at 200 rpm at 30 °C. Pre-cultures for batches with translation inhibitors were performed using a similar approach, whereas for batches with different carbon sources the pre-cultures were made with the respective carbon sources instead of D-glucose.

Chemostat cultivations

Chemostat cultivations were performed in 2 L bioreactors (Applikon, Schiedam, The Netherlands) with a working volume of 1.0 L, the dilution rates used in this study were 0.2, 0.23, 0.27, 0.3, 0.32 and 0.34 h^{-1} in two independent replicate cultures. Growth rates were controlled by modifying the inflow rate on each experiment. Synthetic medium according to Verduyn[54] supplemented with 7.5 g L^{-1} of glucose and 0.25 g L^{-1} Pluronic 6100 PE antifoaming agent was supplied to the bioreactor from a 20 L continuously mixed reservoir vessel. Cultures were sparged with dried air at a flow rate of 700 mL min^{-1} and stirred at 800 rpm. The pH of the cultures was maintained at 5.0 by automatic addition of 2 M KOH. If, after at least six volume changes, the cultures dry cell weight concentration and carbon dioxide production ratediffered less than 2% over two consecutive volume changes the cultures were considered to be in steady state. For cultures with dilution rates of 0.27, 0.3, 0.32 and 0.34 h^{-1}, cultures were first maintained at a dilution rate of 0.2 h^{-1} for 15 hours (3 volume changes) prior to increasing the specific dilution rate to said values.

Batch cultivations with different carbon sources

Batch cultivations were performed using synthetic medium[54], the medium was supplemented with 20 g L^{-1} final concentrations of the carbon sources, either D-trehalose, D-galactose, D-maltose or D-glucose (Sigma Aldrich). The bioreactors were inoculated with 100 mL of yeast shake flask cultures, exponentially growing on the specific carbon source.The final OD_{660} of all pre-cultures was 4. Cultivations were performed at 30 °C, the pH was kept at 5.0 by automatic

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

addition of 2M KOH. The working volume of the bioreactors was 1.4 L in 2 L bioreactors (Applikon, Schiedam, The Netherlands). The cultures were stirred at 8000 rpm and sparged with a flow rate of 700 mL min^{-1} of dried air. Oxygen levels were kept above 40% of the initial saturation level as measured with Clark electrode (Mettler Toledo, Greifensee, Switzerland).

Batch cultivations with the translation inhibitor cycloheximide

Batch cultivations with the translation inhibitor cycloheximide were performed as for the batches with different carbon sources, except that all the batch cultures ran on 20 g L^{-1} of D-glucose and were supplemented with different concentrations of cycloheximide with the aim of reaching specific growth rates. In total five growth rates were studied, being 0.06, 0.12, 0.2, 0.32 and 0.41 h^{-1} (adding respective cycloheximide concentrations of 228.96, 124.51, 52.15, 25.99 and 0 µg L^{-1}).

Analytical methods

Cultures dry weight was measured by filtering 20 mL of culture, the sample was filtered in pre-dried and pre-weight membrane filters with a pore size of 0.45 µm (Gelman Science), the filter was washed with demineralized water, subsequently it was dried in a microwave (20 min, 350 W) and the final weight was measured as described previously.

For the measurement of organic acids and residual carbon source concentrations, supernatants of the cultures were used. For carbon-limited chemostat cultures, the samples were directly quenched with cold steel beads and filtered[55], whereas samples from batch cultures were centrifuged (5 min at 16.000× g). The supernatants were analysed by high-performance chromatography analysis on an Agilent 1100 HPLC (Agilent Technologies) equipped with an Aminex HPX-87H ion-exchange column (BioRad, Veenendaal, The Netherlands), operated with 5 mM H_2SO_4 as the mobile phase at a flow rate of 0.6 mL min^{-1} and at 60 °C. Detection was according to a dual-wavelength absorbance detector (Agilent G1314A) and a refractive-index detector (Agilent G1362A), as described previously.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The exhaust gas from batch cultures was cooled down with a condenser (2°C) and dried with a PermaPure Dryer (model MD 110-8P-4; Inacom Instruments, Veenendaal, the Netherlands) before online analysis of carbon dioxide and oxygen with a Rosemount NGA 2000 Analyser (Baar, Switzerland).

Glycogen and trehalose assays

1 mL of culture was taken from the chemostats and directly added to 5 mL of cold methanol (-40 °C). The sample was mixed and centrifuged (4400× *g*, -20 °C for 5 minutes), the supernatant was discarded, and the pellet was washed in 5 mL of cold methanol (-40 °C), and pellets were stored at -80 °C until further processing. Subsequently, the pellets were resuspended in 0.25 M Na_2CO_3 and processed as described previously[56,57]. D-glucose released from trehalose and glycogen were measured with a D-glucose assay kit (K-GLUC Megazyme), two biological replicates and three technical replicates were analysed per condition.

RNA determination

For RNA determination, 1-2 mL of broth was transferred to a filter (pore size of 0.45 μm, Gelman Science), after which the filter was washed with cold TCA 5 %. The cells were resuspended in 3 mL of TCA 5% and centrifuged for 15 minutes at 4 °C at 4000 rpm. The supernatant was removed and the pellet was stored at -20 °C. Finally, samples were processed as described by Popolo et al., 1982. Two biological replicates and three technical replicates were analysed per condition.

Protein determination

For the batches with CHX, culture volumes corresponding to 50 mg of DCW were centrifuged, washed twice with cold demineralized sterile waterand divided into two aliquots of 5 mL. 2 mL of the aliquot (containing 10 mg DW) was mixed with 1 mL of 3 M NaOH and incubated at 100 °C for 10 minutes. The final mix was diluted and processed following the copper-sulfate based

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

method as described previously[58]. The absorbance of the supernatant was measured at 510 nm, for calibration lyophilized bovine serum albumin (A2153, Sigma Aldrich) was used. Two biological replicates and 3 technical replicates were analysed per condition.

Proteomics sample preprocessing

Aliquots of 20 mL of culture from chemostats and batches with different carbon sources were centrifuged (4000 rpm 4 °C, 10 minutes) and washed two times, the final pellet was flash frozen in liquid nitrogen and stored at -80 °C. Two biological replicates and two technical replicates were analysed per condition.

Frozen cell pellets were thawed on ice before transfer to Precellys® Lysing Kit 2 ml screw cap vials with 0.5mm glass beads (Bertin Instruments, France). Lysis was performed in 250 µl lysis buffer, 50 mM ammonium bicarbonate with cOmplete protease inhibitor cocktail (ROCHE, Switzerland), using a Minilys Personal Tissue Homogenizer (Bertin Instruments, France), at maximum speed for 15 cycles of 30 seconds with a one-minute rest on ice between each cycle. Lysed material was centrifuged for 10 minutes 13,000× *g* at 4°C, the supernatant fraction was removed and retained. Fresh lysis buffer (250 µl) was added to the insoluble material, which was resuspended before extraction from the vial via a small hole inserted into the vial base. Soluble and insoluble fractions were recombined and the total final volume recorded. Protein concentration was determined using Pierce™ Coomassie Plus Bradford Assay Kit (ThermoFisher Scientific, UK).

Protein (100 µg) from each sample was treated with 0.05 % (w/v) RapiGest™ SF surfactant (Waters, UK) at 80 °C for 10 minutes, reduced with 4 mM dithiothreitol (Melford Laboratories Ltd., UK) at 60 °C for 10 minutes and subsequently alkylated with 14 mM iodoacetamide (SIGMA, UK) at room temperature for 30 minutes. Proteins were digested with 2 µg Trypsin Gold, Mass Spectrometry Grade (Promega, US) at 37 °C for 4 hours before a top-up of a further 2 µg trypsin and incubation at 37 °C overnight. Digests were acidified by addition of trifluoroacetic acid (Greyhound Chromatography and Allied Chemicals, UK) to a final concentration of 0.5 %

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(v/v) and incubated at 37 °C for 45 minutes before centrifugation at 13,000× *g* (4°C) to remove insoluble non-peptidic material.

Proteomics analytics

The sample running order was randomised using a random number generator (Random.org). Samples were analysed using an UltiMate™ 3000 RSLCnano system (ThermoFisher Scientific) coupled to a Q Exactive™ HF Hybrid Quadrupole-Orbitrap™ Mass Spectrometer. Protein digests (1 ug of each) were loaded onto a trapping column (Acclaim PepMap 100 C18, 75 µm x 2 cm, 3 µm packing material, 100 Å) using 0.1 % (v/v) trifluoroacetic acid, 2 % (v/v) acetonitrile in water at a flow rate of 12 µL min-1 for 7 min.

The peptides were eluted onto the analytical column (EASY-Spray PepMap RSLC C18, 75 µm x 50 cm, 2 µm packing material, 100 Å) at 40°C using a linear gradient of 120 minute shallow gradient rising from 8 % (v/v) acetonitrile/0.1 % (v/v) formic acid (Fisher Scientific, UK) to 30 % (v/v) acetonitrile/0.1 % (v/v) formic acid at a flow rate of 300 nL min^{-1}. The column was then washed at 1 % A : 99 % B for 8 min, and re-equilibrated to starting conditions. The nano-liquid chromatograph was operated under the control of Dionex Chromatography MS Link 2.14.

The nano-electrospray ionisation source was operated in positive polarity under the control of QExactive HF Tune (version 2.5.0.2042), with a spray voltage of 2.1 kV and a capillary temperature of 250°C. The mass spectrometer was operated in data-dependent acquisition mode. Full MS survey scans between m/z 300-2000 were acquired at a mass resolution of 60,000 (full width at half maximum at m/z 200). For MS, the automatic gain control target was set to 3e^6, and the maximum injection time was 100 ms. The 16 most intense precursor ions with charge states of 2-5 were selected for MS/MS with an isolation window of 2 m/z units. Product ion spectra were recorded between m/z 200-2000 at a mass resolution of 30,000 (full width at half maximum at m/z 200). For MS/MS, the automatic gain control target was set to 1e^5, and the maximum injection time was 45 ms. Higher-energy collisional dissociation was

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

performed to fragment the selected precursor ions using a normalised collision energy of 30 %. Dynamic exclusion was set to 30 s.

Proteomics data analysis

The resulting raw data files generated by XCalibur (version 3.1) were processed using MaxQuant software (version 1.6.0.16)[59]. The search parameters were set as follows: label free experiment with default settings; cleaving enzyme trypsin with 2 missed cleavages; Orbitrap instrument with default parameters; variable modifications: oxidation (M) and Acetyl (protein N-term); first search as default; in global parameters, the software was directed to the FASTA file; for advanced identification "Match between runs" was checked; for protein quantification we only used unique, unmodified peptides. All other MaxQuant settings were kept as default. The false discovery rate (FDR) for accepted peptide spectrum matches and protein matches was set to 1%. The CEN.PK113-7D Yeast FASTA file was downloaded from the *Saccharomyces* Genome Database (SGD) (https://downloads.yeastgenome.org/sequence/strains/CEN.PK/CEN.PK113-7D/CEN.PK113-7D_Delft_2012_AEHG00000000/).

The resulting MaxQuant output was then analysed using the MSstats package (version 3.5.6)[60] in the R environment (version 3.3.3) to obtain differential expression fold changes with associated *p* values, along with normalized LFQ and intensity values as described previously[61].

Acknowledgements

This work was supported by NWO (NWO ERA-IB-2, project No 053.80.722 to BT and PDL) and Biotechnology and Biological Sciences Research Council (BB/M025748/1 to S. Hubbard, BB/M025756/1 to RB). PG and BT also acknowledge support by Marie Skłodowska-Curie Actions ITN "SynCrop" (grant agreement No 764591). We thank SURFsara for the HPC resources through access to the Lisa Compute Cluster.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Author Contributions

Conceptualization, funding acquisition and supervision: BT, JN, PDL, RB, S. Hubbard; experimental data collection: ARP, VH, S. Holman; experimental data analysis: ARP, PG, MGA, MMMB; computational modeling: IEE, PG; formal analysis: IEE, PG, JvH, FJB, NS, JN, PDL, BT; writing – original draft: BT; writing – editing: IEE, PG, FJB, JN, BT. All authors have read and approved the manuscript.

References

1. Blank, L., Lehmbeck, F. & Sauer, U. Metabolic-flux and network analysis in fourteen hemiascomycetous yeasts. *FEMS Yeast Res.* **5**, 545–558 (2005).
2. Pfeiffer, T. & Morley, A. An evolutionary perspective on the Crabtree effect. *Front. Mol. Biosci.* **1**, 1–6 (2014).
3. De Deken, R. H. The Crabtree effect: a regulatory system in yeast. *J. Gen. Microbiol.* **44**, 149–156 (1966).
4. Van Hoek, P., Van Dijken, J. P. & Pronk, J. T. Effect of specific growth rate on fermentative capacity of baker's yeast. *Appl. Environ. Microbiol.* **64**, 4226–4233 (1998).
5. de Groot, D. H., van Boxtel, C., Planqué, R., Bruggeman, F. J. & Teusink, B. The number of active metabolic pathways is bounded by the number of cellular constraints at maximal metabolic rates. *PLOS Comput. Biol.* **15**, e1006858 (2019).
6. Basan, M. *et al.* Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature* **528**, 99–104 (2015).
7. Molenaar, D., van Berlo, R., de Ridder, D. & Teusink, B. Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol. Syst. Biol.* **5**, 323 (2009).
8. Mori, M., Hwa, T., Martin, O. C., De Martino, A. & Marinari, E. Constrained Allocation Flux Balance Analysis. *PLoS Comput. Biol.* **12**, 1–24 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

9. O'Brien, E. J., Lerman, J. A., Chang, R. L., Hyduke, D. R. & Palsson, B. Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol. Syst. Biol.* **9**, (2013).
10. Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of Cell Growth and Gene Expression: Origins and Consequences. *Science* **330**, 1099–1102 (2010).
11. Wortel, M. T., Bosdriesz, E., Teusink, B. & Bruggeman, F. J. Evolutionary pressures on microbial metabolic strategies in the chemostat. *Sci. Rep.* **6**, 29503 (2016).
12. Scott, M. & Hwa, T. Bacterial growth laws and their applications. *Curr. Opin. Biotechnol.* **22**, 559–565 (2011).
13. Wortel, M. T., Peters, H., Hulshof, J., Teusink, B. & Bruggeman, F. J. Metabolic states with maximal specific rate carry flux through an elementary flux mode. *FEBS J.* **281**, 1547–1555 (2014).
14. Berkhout, J. *et al.* How biochemical constraints of cellular growth shape evolutionary adaptations in metabolism. *Genetics* **194**, 505–512 (2013).
15. Bosdriesz, E., Molenaar, D., Teusink, B. & Bruggeman, F. J. How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization. *FEBS J.* **282**, 2029–2044 (2015).
16. Metzl-Raz, E. *et al.* Principles of cellular resource allocation revealed by condition-dependent proteome profiling. *eLife* **6**, e28034 (2017).
17. 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).
18. Björkeroth, J. *et al.* Proteome reallocation from amino acid biosynthesis to ribosomes enables yeast to grow faster in rich media. *Proc. Natl. Acad. Sci.* **117**, 21804–21812 (2020).
19. Sánchez, B. J., Feiran Li, Hongzhong Lu, Kerkhoven, E. & Nielsen, J. *SysBioChalmers/yeast-GEM: yeast 7.6.0*. (Zenodo, 2016). doi:10.5281/ZENODO.1495468.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

20. Beg, Q. K. *et al.* Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity. *Proc. Natl. Acad. Sci.* **104**, 12663–12668 (2007).
21. Adadi, R., Volkmer, B., Milo, R., Heinemann, M. & Shlomi, T. Prediction of microbial growth rate versus biomass yield by a metabolic network with kinetic parameters. *PLoS Comput. Biol.* **8**, (2012).
22. Labhsetwar, P. *et al.* Population FBA predicts metabolic phenotypes in yeast. *PLOS Comput. Biol.* **13**, e1005728 (2017).
23. Nilsson, A. & Nielsen, J. Metabolic Trade-offs in Yeast are Caused by F1F0-ATP synthase. *Sci. Rep.* **6**, 22264 (2016).
24. Zhuang, K., Vemuri, G. N. & Mahadevan, R. Economics of membrane occupancy and respiro-fermentation. *Mol. Syst. Biol.* **7**, 1–9 (2011).
25. Okamoto, K. The protein import motor of mitochondria: a targeted molecular ratchet driving unfolding and translocation. *EMBO J.* **21**, 3659–3671 (2002).
26. Kafri, M., Metzl-Raz, E., Jona, G. & Barkai, N. The Cost of Protein Production. *Cell Rep.* **14**, 22–31 (2016).
27. Bruggeman, F. J., Planqué, R., Molenaar, D. & Teusink, B. Searching for principles of microbial physiology. *FEMS Microbiol. Rev.* **44**, 821–844 (2020).
28. Famili, I., Forster, J., Nielsen, J. & Palsson, B. O. Saccharomyces cerevisiae phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. *Proc. Natl. Acad. Sci.* **100**, 13134–13139 (2003).
29. Canelas, A. B., Ras, C., ten Pierick, A., van Gulik, W. M. & Heijnen, J. J. An in vivo data-driven framework for classification and quantification of enzyme kinetics and determination of apparent thermodynamic data. *Metab. Eng.* **13**, 294–306 (2011).
30. Jules, M., Guillou, V., François, J. & Parrou, J.-L. Two distinct pathways for trehalose assimilation in the yeast Saccharomyces cerevisiae. *Appl. Environ. Microbiol.* **70**, 2771–2778 (2004).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

31. Diderich, J. A. *et al.* Glucose uptake kinetics and transcription of HXT genes in chemostat cultures of Saccharomyces cerevisiae. *J. Biol. Chem.* **274**, 15350–15359 (1999).
32. Teusink, B., Diderich, J. A., Westerhoff, H. V., van Dam, K. & Walsh, M. C. Intracellular Glucose Concentration in Derepressed Yeast Cells Consuming Glucose Is High Enough To Reduce the Glucose Transport Rate by 50%. *J. Bacteriol.* **180**, 556–562 (1998).
33. Chen, Y. & Nielsen, J. Energy metabolism controls phenotypes by protein efficiency and allocation. *Proc. Natl. Acad. Sci.* **116**, 17592–17597 (2019).
34. O'Brien, E. J., Utrilla, J. & Palsson, B. O. Quantification and Classification of E. coli Proteome Utilization and Unused Protein Costs across Environments. *PLOS Comput. Biol.* **12**, e1004998 (2016).
35. Goel, A. *et al.* Protein costs do not explain evolution of metabolic strategies and regulation of ribosomal content: does protein investment explain an anaerobic bacterial Crabtree effect?: Protein costs and evolution of metabolic strategies. *Mol. Microbiol.* **97**, 77–92 (2015).
36. de Jongh, W. A. *et al.* The roles of galactitol, galactose-1-phosphate, and phosphoglucomutase in galactose-induced toxicity in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **101**, 317–326 (2008).
37. Postma, E., Verduyn, C., Kuiper, A., Scheffers, W. A. & Van Dijken, J. P. Substrate-accelerated death ofSaccharomyces cerevisiae CBS 8066 under maltose stress. *Yeast* **6**, 149–158 (1990).
38. Hong, K.-K., Vongsangnak, W., Vemuri, G. N. & Nielsen, J. Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 12179–12184 (2011).
39. Keren, L. *et al.* Massively Parallel Interrogation of the Effects of Gene Expression Levels on Fitness. *Cell* **166**, 1282-1294.e18 (2016).
40. Chaker-Margot, M. Assembly of the small ribosomal subunit in yeast: mechanism and regulation. *RNA* **24**, 881–891 (2018).
41. Kunkel, J., Luo, X. & Capaldi, A. P. Integrated TORC1 and PKA signaling control the temporal activation of glucose-induced gene expression in yeast. *Nat. Commun.* **10**, 3558 (2019).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

42. Chubukov, V., Gerosa, L., Kochanowski, K. & Sauer, U. Coordination of microbial metabolism. *Nat. Rev. Microbiol.* **12**, 327–340 (2014).

43. Hofmeyr, J. S. & Cornish-Bowden, A. Regulating the cellular economy of supply and demand. *FEBS Lett.* **476**, 47–51 (2000).

44. Vander Heiden, M. G. & DeBerardinis, R. J. Understanding the Intersections between Metabolism and Cancer Biology. *Cell* **168**, 657–669 (2017).

45. Dill, K. A., Ghosh, K. & Schmit, J. D. Physical limits of cells and proteomes. *Proc. Natl. Acad. Sci.* **108**, 17876–17882 (2011).

46. van Tatenhove-Pel, R. J. *et al.* Serial propagation in water-in-oil emulsions selects for Saccharomyces cerevisiae strains with a reduced cell size or an increased biomass yield on glucose. *Metab. Eng.* **64**, 1–14 (2021).

47. Jensen, P. R., Michelsen, O. & Westerhoff, H. V. Control analysis of the dependence of Escherichia coli physiology on the H(+)-ATPase. *Proc. Natl. Acad. Sci.* **90**, 8068–8072 (1993).

48. Buziol, S. *et al.* Dynamic response of the expression of hxt1, hxt5 and hxt7 transport proteins in Saccharomyces cerevisiae to perturbations in the extracellular glucose concentration. *J. Biotechnol.* **134**, 203–210 (2008).

49. Gresham, D. *et al.* The Repertoire and Dynamics of Evolutionary Adaptations to Controlled Nutrient-Limited Environments in Yeast. *PLoS Genet.* **4**, e1000303 (2008).

50. ter Kuile, B. H. & Westerhoff, H. V. Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. *FEBS Lett.* **500**, 169–171 (2001).

51. Cornish-Bowden, A. The effect of natural selection on enzymic catalysis. *J. Mol. Biol.* **101**, 1–9 (1976).

52. Daran-Lapujade, P. *et al.* The fluxes through glycolytic enzymes in Saccharomyces cerevisiae are predominantly regulated at posttranscriptional levels. *Proc. Natl. Acad. Sci.* **104**, 15753–15758 (2007).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

53. Nijkamp, J. F. *et al.* De novo sequencing, assembly and analysis of the genome of the laboratory strain Saccharomyces cerevisiae CEN.PK113-7D, a model for modern industrial biotechnology. *Microb. Cell Factories* **11**, 36 (2012).
54. Verduyn, C., Postma, E., Scheffers, W. A. & Van Dijken, J. P. Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* **8**, 501–517 (1992).
55. Mashego, M. R., van Gulik, W. M., Vinke, J. L. & Heijnen, J. J. Critical evaluation of sampling techniques for residual glucose determination in carbon-limited chemostat culture ofSaccharomyces cerevisiae. *Biotechnol. Bioeng.* **83**, 395–399 (2003).
56. Parrou, J. L., Teste, M.-A. & François, J. Effects of various types of stress on the metabolism of reserve carbohydrates in Saccharomyces cerevisiae: genetic evidence for a stress-induced recycling of glycogen and trehalose. *Microbiology* **143**, 1891–1900 (1997).
57. Boer, V., Daran, J., Almering, M., Dewinde, J. & Pronk, J. Contribution of the transcriptional regulator Leu3p to physiology and gene expression in nitrogen- and carbon-limited chemostat cultures. *FEMS Yeast Res.* **5**, 885–897 (2005).
58. Verduyn, C., Postma, E., Scheffers, W. A. & van Dijken, J. P. Physiology of Saccharomyces Cerevisiae in Anaerobic Glucose-Limited Chemostat Culturesx. *J. Gen. Microbiol.* **136**, 395–403 (1990).
59. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
60. Choi, M. *et al.* MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* **30**, 2524–2526 (2014).
61. Garcia-Albornoz, M. *et al.* A proteome-integrated, carbon source dependent genetic regulatory network in *Saccharomyces cerevisiae*. *Mol. Omics* **16**, 59–72 (2020).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

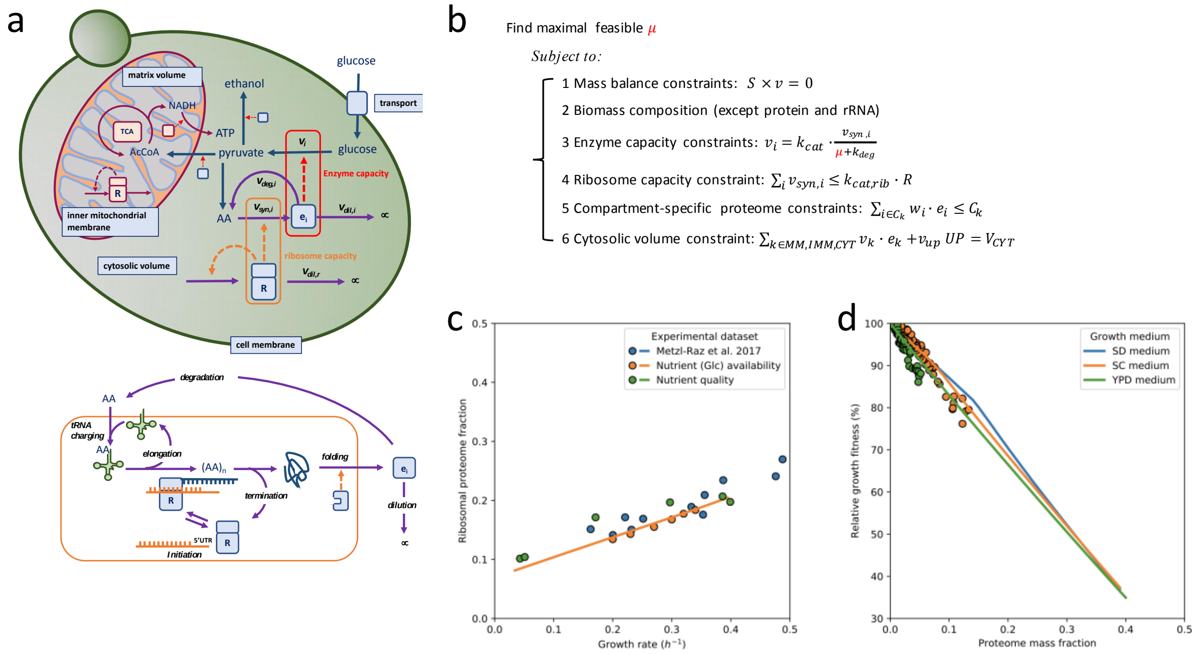


Fig. 1 pcYeast model formulation and calibration of protein synthesis parameters. **a.** A schematic overview of reactions in the model, their interdependence and constraints. Metabolic reactions v_i are proportional to enzyme concentrations e_i that are synthesized at rate $v_{syn,i}$ by the ribosomes R. Each protein can be degraded with rate $v_{deg,i} = k_{deg} \cdot e_i$ or diluted by growth rate $v_{dil,i} = \mu \cdot e_i$. Compartment-specific constraints are indicated in the light-blue boxes. **b.** Optimisation problem with the key constraints, including 1) steady-state mass balances; 2) production of biomass components such as DNA, lipids, cell wall and polysaccharides. Proteins and tRNA are excluded as their synthesis rates are optimisation variables 3) enzyme capacity constraints that couple metabolic flux to catalytic rate $k_{cat,i}$ and the enzyme level, whose value at steady state is determined by its synthesis rate, rates of enzyme degradation, and dilution by growth. Note we use equalities and hence enzymes work at their maximal rate and minimal required protein levels are computed; 4) ribosome capacity that defines an upper bound for protein synthesis rate; 5) compartment-specific proteome constraints that define the maximal concentration of proteins that can be contained in that compartment, with w_i the specific volume or area of protein i; 6) a cytosolic protein density constraint that has the same function as that of proteome constraints, but whose equality forces the cell to fill up any vacant proteome space with unspecified protein UP with average amino acid composition. **c.** Growth rate was

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

varied through sugar type (trehalose, galactose, maltose, glucose) or glucose concentration, and ribosomal protein fraction was determined by proteomics (which was consistent with literature data, also plotted). The translation rate was calibrated on that data, as detailed in Supplementary Notes. **d.** Impact of mCherry protein overexpression on growth rate. Symbols show experimental data[26], solid lines show model predictions based on glucose minimal (SD) medium or rich SC/YPD media.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

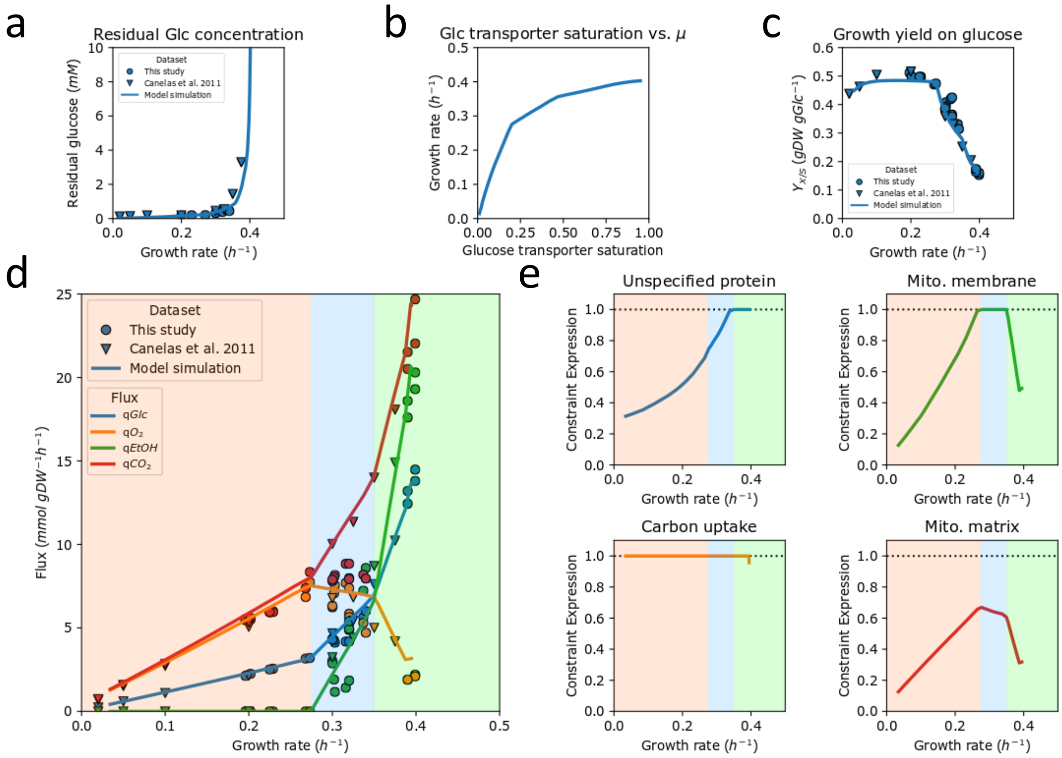
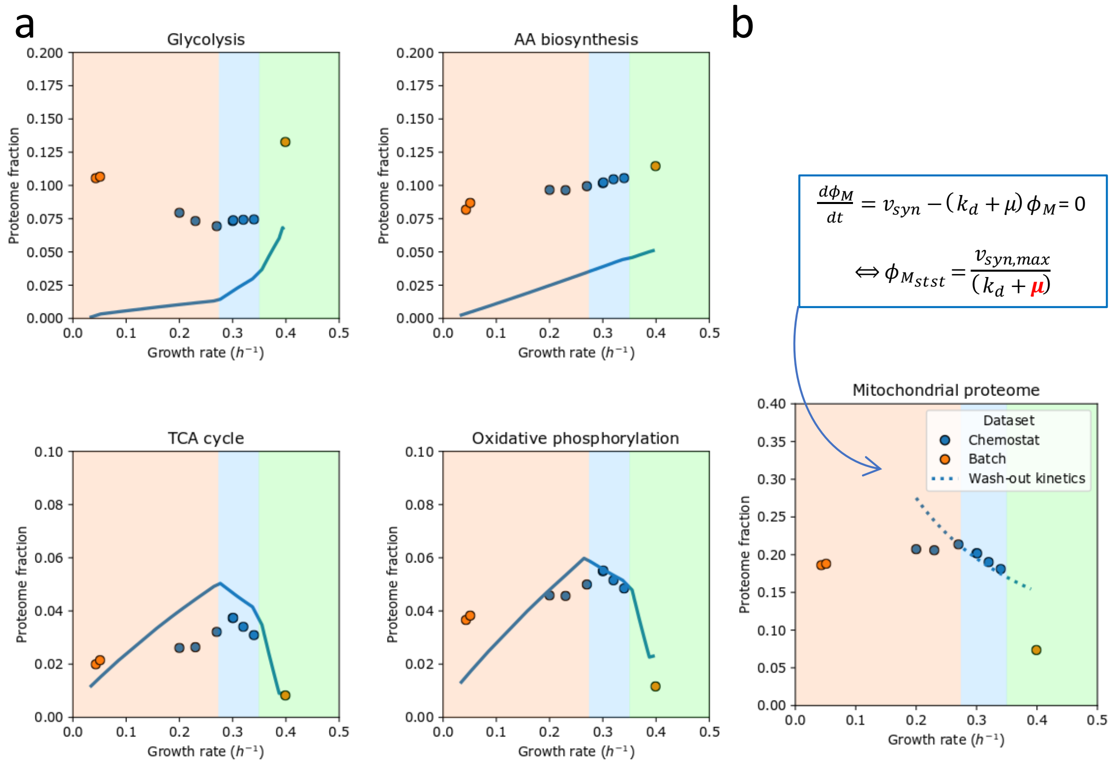


Fig. 2 Predicted and measured physiological response of *S. cerevisiae* CEN.PK as a function of glucose availability a. Measured (symbols) and predicted (line) residual glucose concentrations as a function of growth rate. The latter was calculated based on published affinity for glucose and assuming negligible intracellular glucose under these conditions. **b.** Maximal feasible growth rates of the model as a function of the glucose transporter saturation. **c.** measured (symbols) and predicted biomass yield on glucose. **d.** Experimental fluxes from glucose-limited chemostats at different dilution rates (circles) and from two batch experiments (triangles): excess trehalose (which mimicks glucose limitation at low dilution rate[30]) and excess glucose at the highest growth rate. The lines are model predictions; background colors indicate regimes with different active constraints; **e**. Computed proteome occupancy of different constrained protein pools. A fraction of 1 means that the compartment is full with metabolically actively proteins and constrains the growth rate at that condition.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



815

816 **Fig. 3 Proteomics data of selected pathways as a function of glucose availability**. Blue symbols are

817 glucose-limited chemostat data; orange symbols are controlled batch experiments with excess

818 trehalose (lowest growth rate) or glucose (highest growth rate) **a.** Comparison of predicted minimal

819 proteome fractions to sustain growth with the experimentally determined proteome fraction for four

820 pathways. The ratio between the two represents an estimate of the saturation level of the constituent

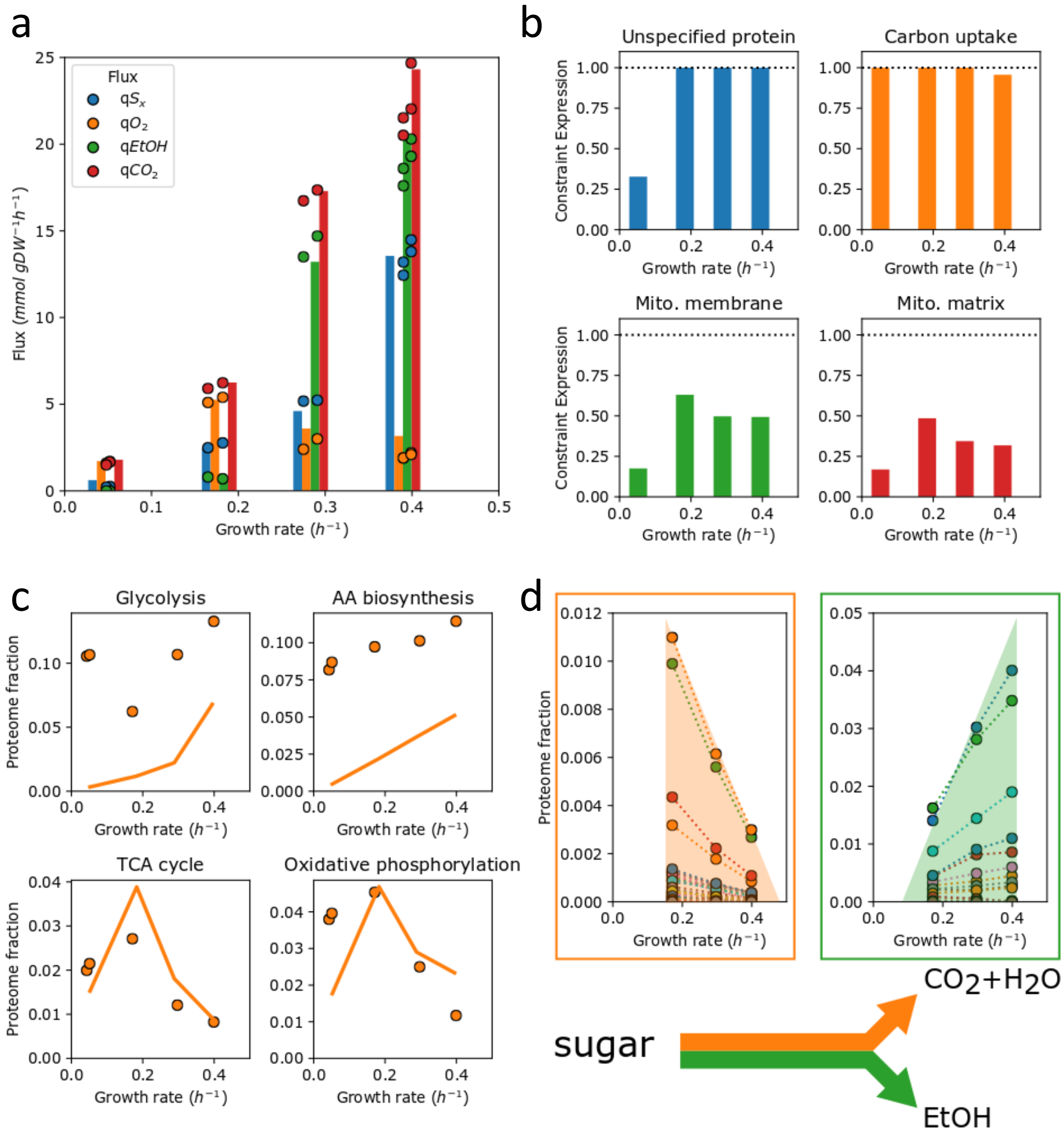
821 enzymes. Lines represent the model; experimental data are symbols. **b**. Decay of steady-state

822 mitochondrial protein fraction with growth rate at onset of ethanol formation suggests a maximal rate

823 of mitochondrial biosynthesis $v_{syn,max}$.

824

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



825

826 **Fig. 4 Model predictions, fluxes and protein levels plotted as a function of growth rate during hexose**

827 **sugar excess conditions (in the order: trehalose, galactose, maltose, glucose) a**. Fluxes of sugar

828 consumption, oxygen consumption and ethanol production. Circles are experimental data, bar plots

829 indicate model predictions (of both the growth rate and fluxes); **b**. Predicted active constraints under

830 the different sugar excess conditions as predicted by the mode (see legend of Fig. 2 for details). **c**.

831 Comparison of predicted minimally needed proteome fractions with experimentally determined ones

832 suggests differences in saturation level between pathways. Lines represent the model, experimental

833 data are circles; **d**. Linearity of the expression of individual enzymes in glycolysis (right) and respiration

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(left) with growth rate suggests trading in of respiratory protein for fermentative protein. The respiratory proteins converge at 0.474 ± 0.0002 h^{-1}.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

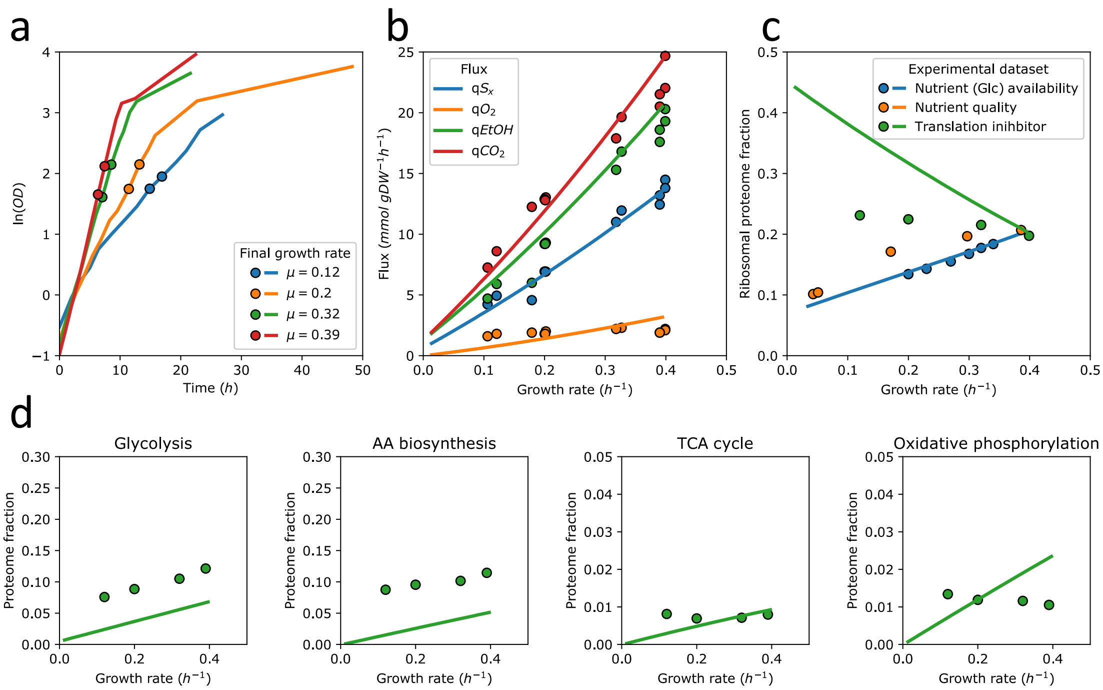


Fig. 5 The effect of translation inhibition by cyclohexamide on growth rate, fluxes and proteome fractions in controlled aerobic batch fermentations on glucose. a. Dependency of culture optical density (OD) on the time post-inoculation to the medium supplemented with cycloheximide. Lines are values of consecutive OD measurements, points represent the times when cultures were sampled. **b**. Main catabolic fluxes as a function of the growth rate. **c**. Ribosomal proteome fractions. Data from Fig. 1c are included for comparison. **d.** Proteome fractions measured for key metabolic pathways, and the minimal proteome fractions predicted by pcYeast. For b-d, lines are model predictions; symbols are experimental data points.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 1. Statistics of the pcYeast model.

Process/Compartment	**# of reactions**	**# of proteins**
Total	**24422**	**1520**
Metabolic network	**5774**	**913**
from Yeast7.6	5738	909
manually added metabolic reactions	36	4
Cytoplasm	**2349**	**778**
Plasma membrane	**529**	**114**
Mitochondria	**1089**	**272**
Endomembrane system	**2127**	**133**
Metabolic complex formation, disassembly, dilution	**2787**	**-**
tRNA turnover and modification	**2194**	**56**
Protein synthesis and turnover	**13312**	**403**
Cytoplasmic translation	1512	138
Mitochondrial translation	8	89
Protein folding	1515	31
Protein degradation	1607	42
Protein misfolding, refolding	6061	73
Protein transport	1324	30
Protein dilution by growth	1285	-
Formation of macromolecular complexes	**355**	**196**

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.11.448029; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 2. Changes to the parameters for simulating sugar excess conditions. NGAM is non-growth related ATP maintenance.

Growth condition	Unit	Glucose (naïve)	Galactose	Maltose
Maximal hexose transporter area	$\mu m^2/cell$	7.5	3.0	3.5
Carbon-related NGAM	$mmol/gDW/h$	0.0	3.0	0.0
Minimal UP fraction	$g\ UP/g\ protein$	0.245	0.49	0.34