

¹ **Elucidating yeast glycolytic dynamics at steady state**
² **growth and glucose pulse through kinetic metabolic**
³ **modeling**

⁴ **David Lao-Martil¹, Joep P.J. Schmitz², Bas Teusink³, and Natal A.W. van Riel^{1,4,*}**

⁵ ¹Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, Noord-Brabant, 5612AE,
⁶ The Netherlands

⁷ ²DSM Biotechnology Center, Delft, Zuid-Holland, 2613AX, The Netherlands

⁸ ³Systems Biology Lab, Vrije Universiteit Amsterdam, Amsterdam, Noord-Holland, 1081HZ, The Netherlands

⁹ ⁴Amsterdam University Medical Center, University of Amsterdam, Amsterdam, Noord-Holland, 1105AZ, The
¹⁰ Netherlands

¹¹ *Correspondent author: n.a.w.v.riel@tue.nl

¹² **Running title**

¹³ Glycolytic dynamics: steady state and glucose pulses

¹⁴ **Abbreviations**

¹⁵ [E]], enzyme concentration; ADP, adenosine diphosphate; ALD, aldolase; AMP, adenosine monophosphate; ATP, adenoside triphosphate; CCM, central carbon metabolism; CFD, computational fluid dynamics; ¹⁷ CO_2 , carbon dioxide; *E.coli*, *Escherichia coli*; ENO, enolase; F6P, fructose 6-phosphate; FF, feast famine; ¹⁸ G6P, glucose 6-phosphate; GP, glucose pulse; GSM, genome-scale model; K_{cat} , catalytic constant; ¹⁹ K_m , m_{ATP} , ATP maintenance rate; michaelis constant; NAD, nicotinamide-adenine-dinucleotide; O_2 , ²⁰ oxygen; ODE, ordinary differential equation; *P.chrysogenum*, *Penicillium chrysogenum*; P/O ratio, phosphate/oxygen ratio; PDC, pyruvate decarboxylase; PEP, phosphoenolpyruvate; PFK, phosphofructokinase; ²² PGI, phosphoglucoisomerase; PPP, pentose phosphate pathway; PTM, post-translational modifications; ²³ q_{ATP} , ATP production rate; q_{CO_2} , carbon dioxide transport rate; q_{O_2} , oxygen transport rate; RQ, respiratory ²⁴ quotient; *S.cerevisiae*, *Saccharomyces cerevisiae*; SS, steady state; sum_{AXP} , sum of ATP, ADP and AMP; ²⁵ TCA cycle, Tricarboxylic acid cycle; TPS1, trehalose phosphate synthase 1; V_{max} , maximum reaction rate; ²⁶ VVUQ, verification, validation and uncertainty quantification.

²⁷ **Keywords**

²⁸ Glycolysis, Growing Cell, Kinetic Modeling, Parameter Estimation, *Saccharomyces cerevisiae*

29 **Abstract**

30 Microbial cell factories face changing environments during industrial fermentations. Kinetic metabolic
31 models enable the simulation of the dynamic metabolic response to these perturbations, but their develop-
32 ment is challenging due to model complexity and experimental data requirements. An example of this
33 is the well-established microbial cell factory *Saccharomyces cerevisiae*, for which no consensus kinetic
34 model of central metabolism has been developed and implemented in industry. Here, we aim to bring
35 the academic and industrial communities closer to this consensus model. We developed a physiology
36 informed kinetic model of yeast glycolysis connected to central carbon metabolism by including the effect
37 of anabolic reactions precursors, mitochondria and the trehalose cycle. To parametrize such a large model,
38 a parameter estimation pipeline was developed, consisting of a divide and conquer approach, supplemented
39 with regularization and global optimization. Additionally, we show how this first mechanistic description
40 of a growing yeast cell captures experimental dynamics at different growth rates and under a strong glucose
41 perturbation, is robust to parametric uncertainty and explains the contribution of the different pathways in
42 the network. Such a comprehensive model could not have been developed without using steady state and
43 glucose perturbation data sets. The resulting metabolic reconstruction and parameter estimation pipeline
44 can be applied in the future to study other industrially-relevant scenarios.

45 **Highlights**

- 46 • Physiology informed kinetic model of yeast glycolysis connected to central carbon metabolism.
- 47 • The metabolic reconstruction reproduces both steady state and glucose perturbation data and explains
48 the contribution of different metabolic pathways to the glycolytic response.
- 49 • A new parameter estimation pipeline is developed consisting of a divide and conquer approach,
50 regularization and global optimization.
- 51 • This model is a tool that can be used for rational metabolic engineering.

52 1 Introduction

53 *Saccharomyces cerevisiae*, commonly known as baker's yeast, is a prominent cell factory for the biotechnology industry¹. In large-scale fermentations, dynamic gradients expose yeast cells to rapid nutrient changes in their extracellular environment, which in turn will impact intracellular metabolic regulation^{2,3}.

54 To understand this dynamic stress response process, mechanistic models representing intracellular carbon metabolism kinetics have been developed^{4–6}. However, upscaling these models from lab to industry has proven challenging and restricted to simplified mechanistic growth models with little information on intracellular dynamics^{7,8}.

55 This could be partly explained because models lacked important physiological information. In the fields of pharmacology and medicine, validation of physiological realism is a necessary step in development of credible mechanistic model⁹. To develop physiologically representative yeast models, several aspects are to be considered. Current state of the art models study glycolysis in isolation^{4–6}, but should consider the complex cellular context it interacts with, for example by representing the trehalose and tricarboxylic acid (TCA) cycles.

56 The situation regarding biomass synthesis is an important example. Whereas maximization of biomass synthesis is the physiological function to which genome scale models (GSM) are optimized¹⁰, it has been neglected in yeast kinetic networks, resulting in models that cannot represent a growing cell and thus poorly represent an important physiological state that comes with importable demands¹¹. Considering sink reactions for glycolytic intermediates is of upmost importance, as they were shown to be central in achieving the steady state¹².

57 Moreover, variables alien to carbon flux regulate glycolysis as well. For instance, oxygen concentration determines if respiration or fermentation is performed¹³ and is routinely quantified in both lab and industrial scale bioreactors^{3,14}, but has not been used to constrain dynamic models. Furthermore, cofactors NAD and ATP are required to carry out multiple glycolytic reactions, but their availability cannot be taken for granted. ATP homeostasis is challenged under strong dynamic perturbations, where the inosine salvage pathway is used as transient store¹⁵. In addition, ATP required for maintenance changes under different growth conditions¹⁶. Still, a yeast model considering the effect of these variables does not exist.

58 Recently published data and tools make implementation of yeast models in a physiological context more feasible. The first perturbations studies in yeast were restricted to a single glucose pulse perturbation of 1 g L⁻¹¹⁷, but experimental quantification has been extended to larger perturbations that result in profound

intracellular dynamics, including possible nonlinear effects of stress response^{6,15}. Furthermore, data on different growth steady states (SS)^{4,14} and industrially relevant feast/famine (FF) regimes¹⁸ has also been generated. New experimental variables have been quantified, such as most metabolic species in central carbon metabolism (CCM)¹⁴, dynamic flux profiles¹⁹, and proteome composition^{16,20}. These data can now be used to quantify model parameters within a physiological range. Previous work showed that kinetic constants measured *in vitro* in isolated enzymes assays might not be representative of *in vivo* behavior^{12,21}. To overcome this issue, model parameters can be directly estimated to fit *in vivo* metabolomic and fluxomic data²¹. Few works estimated part of the yeast kinetic model parameter set in this way^{5,22,23}, but the data available has notably increased since then.

To address the abovementioned challenges, here we conceived a kinetic model of yeast glycolysis and trehalose cycle that considers the physiological effect of growth rate, gas exchange, ATP synthesis and maintenance, by developing a state-of-the-art approach for parameter identification. We show how this first mechanistic description of a growing yeast cell captures experimental dynamics at different growth rates and under a strong glucose perturbation, is robust to parametric uncertainty and explains the contribution of the different pathways in the network. Our work suggests that combining multiple types of data and computational methods, complex but physiologically representative and robust models can be achieved, and simultaneously points at specific locations in glycolytic models where regulation mechanisms are missing.

100 **2 Results**

101 **2.1 An integrated, physiology informed kinetic model of yeast glycolysis**

102 Existing yeast glycolysis models studied this pathway in isolation^{5,6}. To transit from pathways in isolation,
103 to pathways embedded in a growing cell, inclusion of physiological variables informative of the bioreactor
104 environment is crucial²⁴. The required data to implement these variables is now available for *S. cerevisiae*.
105 Carbon flux distribution in central metabolism and secretion of O_2 and CO_2 were determined at different
106 dilution rates in chemostat¹⁴. Additionally, growth and non-growth -associated maintenance (GAM and
107 NGAM, respectively) were determined and extensively implemented in GSMS and can be used to estimate
108 the ATPase activity^{25,26}.

109 These variables have been included in the model developed in this work. Carbon metabolism includes
110 glycolysis, trehalose cycle, glycerol branch and sink reactions to biomass and other by products (Fig 1).
111 Production of CO_2 (q_{CO_2}) consisted of the sum of what is released in the pyruvate decarboxylase (PDC)
112 reaction and all the carbon moles that are processed in the mitochondria. The uptake of oxygen (q_{O_2}) was
113 estimated using the q_{CO_2} and the experimental respiratory quotient (RQ) value. Finally, the synthesis of
114 ATP (q_{ATP}) consisted of glycolytic production and mitochondrial contribution, estimated considering q_{O_2}
115 and P/O ratio²⁷. The maintenance consumption of ATP (m_{ATP}) was estimated using the GAM and NGAM
116 determined in GSMS²⁶.

117 The model consists of a series of nonlinear ordinary differential equations (ODE), where each mass balance
118 describes a metabolite concentration, that changes over time according to reaction rates. Reactions that
119 cannot be described by a single enzyme, such as the rates in which glycolytic intermediates are taken up
120 to form biomass, are described by phenomenological expressions that closely resemble experimental data.
121 The change in enzyme concentrations at different dilution rates found in²⁸ is adjusted for each enzyme.
122 Model parameters were estimated to fit experimental data (see later sections). For further details on
123 model implementation, parameters and kinetic expressions, see the materials and methods and supporting
124 information.

125 **2.2 The model reproduced physiological properties of a growing cell in a bioreactor**

126 The resulting model predicted experimental dynamics, differentiating respiratory and respirofermentative
127 metabolism. At growth rates below 0.2 h^{-1} , carbon uptake was directed to the sink reactions, which
128 account for biomass synthesis and TCA cycle (Fig 2B). Above this threshold, metabolism gradually

shifted to ethanol production which eventually became the main carbon product. In line with this, q_{CO_2} closely resembled q_{O_2} at low growth rates but became predominant above 0.2 h^{-1} , discerning respiratory from respirofermentative metabolism (Fig 2C). Likewise, ATP was produced by respiratory metabolism at low growth rates and partially fermentative at high growth rates (Fig 2D). Maintenance reaction rate was close to a GAM of 40 mmol of ATP per g of biomass during the respiratory state but increased above this level as fermentation became predominant. A mismatch occurred at 0.2 h^{-1} , when carbon flux was notable underestimated. We will discuss this later.

In conclusion, the agreement between model simulations and experimental data suggests that the model can work within the physiologically relevant region studied here. The implementation of the growth rate dependency seems essential. For instance, if only the ATP produced by glycolysis is considered, the overall maintenance needs cannot be matched at the different growth rates. This opens the possibility for kinetic models to be used in new setups, such as different growth rates or respiration/fermentation regimes, by using variables such as gas exchange rate as constraints, rather than as validation.

2.3 Parameter estimation: A problem decomposition approach supplemented by regularization, parameter space sampling and cost function balancing

Quantification of the parameter set was performed by optimizing the model fit to multiple experimental setups and data types^{6, 14, 28}. Nonetheless, the high number of unknown kinetic parameters made this task challenging due to multiple local optimum and ill-conditioning. To deal with non-identifiability problems in large kinetic models, multiple optimization methods have been proposed and evaluated²⁹. For instance, the so-called divide and conquer method aimed to analyze undetermined biochemical networks despite parameter uncertainty³⁰. This problem decomposition approach splits the model in smaller parts where parameters are estimated and later assembled in a complete network.

To develop the *S. cerevisiae* model in this study, this problem decomposition method was adapted and supplemented with global optimization and regularization to improve convergence, akin to the approach described in³¹. Global optimization algorithms meet the danger presented by local minima by starting parameter estimation from a set of initial parameter guesses³². Regularization methods reduce ill-conditioning by introducing prior knowledge on the parameters³³. Additionally, to control that overfitting did not occur, weighting factors were used to leverage the components of the cost function.

The parameter estimation pipeline is illustrated for the well-studied enzyme enolase (ENO) in (Fig 3). In the first stage, the divide and conquer problem decomposition approach was applied to glycolytic enzymes.

159 As suggested in³⁰, steady state data was used^{14,28}. A regularization factor was added in the cost function,
160 so that parameters resembled prior knowledge, provided that the model error was not increased (Fig
161 3A). The benefits of global optimization and regularization can be appreciated for parameters $V_{max,ENO}$
162 and $K_{M,F6P,PFK}$ in Fig 3B and 3C, respectively. A parameter for phosphofructokinase (PFK) is shown
163 since this is the most complex enzyme in the pipeline, and results differed according to enzyme kinetics
164 complexity. While for ENO most initial parameter samples ended in the same parameter value as the
165 literature one, the regularization factor brought estimated paraemters closer to literature values (Fig 3B).
166 For PFK, global sampling revealed the existence of parameter dependencies (Fig 3C), and convergence
167 towards literature values was achieved upon regularization. Since enzymatic kinetic constants vary within
168 a large range, the parameters estimated were transformed from linear to logarithmic scale, as this has
169 shown to be more reliable, also with global optimization methods^{29,34}.

170 In the second stage, the developed models were assembled in a complete glycolysis model with trehalose
171 cycle and cofactor metabolism. The weighting factors in the cost function were balanced ad hoc, until
172 model fit for multiple data could not be further improved. This resulted in pareto fronts, such as the one
173 for ENO (Fig 3D), where the fit for phosphoenolpyruvate (PEP) concentration in steady state and the 110
174 Mm glucose pulse was balanced (Fig 3E and 3F, respectively). See the supplementary information for an
175 overview of the divide and conquer parameter estimation results, enzyme by enzyme.

176 2.4 Combining different data types was central for physiological system identification

177 Determining a model parameter set, or system identification, is a data-intensive process³⁵. To identify the
178 model properties, excitation experiments such as impulse or step response and time variant and invariant
179 data are used. Considering different experimental data types when building a kinetic metabolic model
180 contributes to an accurate physiological description of the system³⁶. In the case of yeast fermentations,
181 early studies focused on the transient metabolic response to a glucose pulse¹⁷, but implementations have
182 broadened to study the role of the trehalose cycle during higher residual glucose concentrations⁶ and
183 steady state dynamics were analyzed at different growth rates¹⁴. In the process, data types have also
184 increased from describing metabolomics to fluxomics and even enzymatic constraints^{20,28}.

185 Therefore, the benefits of combining data types were examined for the developed model, in the areas
186 highlighted in (Fig 4A). Simulations were compared when changing enzyme concentrations were
187 considered as shown in²⁸ (data displayed in the supporting information), to when they were fixed at the
188 dilution rate value of 0.1 h^{-1} (Fig 4B). This analysis showed that inclusion of proteome changes improved

189 the predicted fluxes substantially, especially at higher dilution rates, as the glycolytic flux is increased.
190 This is supported by the relevance attributed to the proteome at different growth rates using GSMS^{20,37}.
191 Furthermore, parameters were estimated to fit a single or combination of experimental variables, to show
192 the benefits of using multiple data types when fitting the model. For instance, reduced confidence intervals
193 were obtained for TPS1 kinetic constants if trehalose concentrations were used in combination with glucose
194 6-phosphate (G6P) (Fig 4C), and similar results were obtained combining metabolomic and fluxomic,
195 or steady state and glucose perturbation data (Fig 4D and 4E, respectively). This tendency highlights
196 how the identified system becomes more accurate, on top of more representative of cell dynamics. This is
197 relevant, given that kinetic yeast CCM models have often been only developed with a single type of data,
198 most usually a glucose perturbation for which only some metabolites were measured.

199 **2.5 Ensemble modelling suggested that the model was robust to parameter heterogeneity**

200 Kinetic metabolic models are generally designed as deterministic descriptions of an average cell in the
201 bioreactor, represented by a unique parameter set³⁵. Nonetheless, heterogeneity between cells of the
202 same population or between strains implies a certain degree of metabolic diversity^{6,38,39}. These Bayesian
203 dynamics can be represented by means of ensemble modelling, in which parameters are sampled from a
204 variability distribution and are often used to account for uncertainty^{40–42}. In this work, a Monte Carlo
205 approach was taken to assess how robust were the modelled glucose perturbation response and steady state
206 dynamics. The models were generated by adding a 10% random noise to the parameter set .

207 Model simulations were found to be notably robust within the conditions tested. Predictions showed
208 certain quantitative but little qualitative deviation (Fig 5). Steady state concentrations and fluxes were
209 especially robust, as it has been found for other kinetic networks^{40,43,44}. Variability was more pronounced
210 during the glucose perturbation, which could be explained by the notably higher extracellular glucose
211 concentrations (110 mM) experienced during the strong perturbation, when compared to the steady state
212 range (0.13-3.30 mM).

213 **2.6 Problem decomposition revealed a mismatch between literature and *in vivo* parame- 214 ters**

215 A concern in the literature is the accuracy of the estimated enzymatic kinetic constants. Previous work
216 showed that kinetic constants measured *in vitro* in isolated enzymes assays might not be representative
217 of *in vivo* behavior^{12,21}. This was improved by standardizing *in vivo*-like assay conditions, but only
218 some glycolytic parameters were updated^{5,45}. In this work, the abovementioned problem decomposition

approach was used to assess this experimental bias and to adjust the parameters to fit the data. Simulating the experimental steady state reaction rates with the models generated in the problem decomposition step revealed a notable bias if the parameter values found in the literature were used, even if *in vivo*-like experimental parameters were considered (Fig 6). These fits were improved using the parameter estimation pipeline described in (Fig 3). In line with²¹, this highlights the need of cross validating experimental data generated in individual enzyme assays with *in vivo* data and suggest that even with improved media conditions, an impactful bias exists. See the supplementary information for a complete comparison between the estimated and the literature parameter set.

2.7 Large deviations from literature parameters suggested missing knowledge around PFK kinetics

When the estimated parameters were compared to the literature, most deviated by less than 0.5-magnitude order from the literature value. Nonetheless, the kinetic constants which deviated the most belonged to PFK kinetics or the two enzymes found before and after this reaction in glycolysis, phosphoglucose isomerase (PGI) or aldolase (ALD) (Table 1). Furthermore, the model fit agreed for most experimental variables, except for the SS concentrations of metabolites G6P and fructose 6-phosphate (F6P), for which simulations disagreed both qualitatively and quantitatively. This could be solved if PFK kinetics parameters were changed in a growth rate dependent manner, for instance, by adding a 'fudge factor' to the maximum reaction rate (see supplementary information), which could suggest that there is a missing regulation on PFK kinetics . Even though well studied, the kinetics of this enzyme are known to be notable complex^{12,46}. For instance, the concentrations of F26bP, a modulator of PFK was not present in the data and a constant value was assumed in our simulations.

2.8 Dilution rate dependent glucose transport kinetics

Transmembrane glucose transport kinetics are known to change with dilution rate. For instance, changing affinities had to be considered to model different dilution rates⁴⁷. Thereafter, multiple transporters (HXT1-7) were found to be notably expressed at different growth rates⁴⁸ . The kinetics of these transporters were found to be divergent and isoenzyme dependent, both in terms of affinity constants (K_m) and maximum reaction rates (V_{max})⁴⁹⁻⁵¹ . Therefore, we investigated whether a unique parameter set could explain the entire data, or dilution rate dependency had to be considered on the kinetic constants.

The simulations with the unique parameter set fit the data to a notable degree, but with certain deviation (Fig 7A). Uptake was overestimated at dilution rates below 0.2 h⁻¹, resulting in fermentative excess where

residual glucose concentrations remained under 0.20 mM (Fig 7B)⁴⁷. Then, this fit could be improved by allowing HXT V_{max} to change in a dilution rate dependent manner (Fig 7A). As expected, V_{max} was found to decrease and increase at low and high dilution rates, respectively (Fig 7C). Nonetheless, the change at low growth rates could also be explained by a decreased enzyme affinity (see supplementary information) or a combination of both V_{max} and K_{cat} changes, in line with⁴⁸. Interestingly, a data point that could not be fit was the dilution rate 0.2 h⁻¹, given the low residual glucose concentration. As suggested in⁴⁷, this could be caused by intracellular changes on the driving force.

2.9 Glycolysis was regulated by other pathways in central carbon metabolism

Glycolysis is a pathway that does not act in isolation but that is influenced by other metabolic routes in CCM. Therefore, here we investigated how the different pathways in the model regulate glycolysis. Flux towards the different glycolytic intermediate sinks was dilution rate dependent. At low growth rates, most carbon entered the sink reactions, while fermentation was predominant at higher growth rates (Fig 8A). Nonetheless, the model did overestimate fermentative flux at lower growth rate, due to small deviations of the already low glucose transport flux. Furthermore, the effect of knock outs resulted in expected dynamics. Simulating the lack of oxygen by knocking out mitochondrial reactions resulted in an increase in fermentation and less NADH mitochondrial respiration lowered glycolytic intake, since the network compensated by over-activation of the glycerol branch (Fig 8B and 8C). Then, knocking out the trehalose cycle during glucose perturbation resulted in an imbalance between upper and lower glycolysis reaction rates, in line with⁶ (Fig 8D) and a knock out of the inosine salvage pathway meant that no ATP paradox would be observed (Fig 8E).

2.10 Missing regulation could explain glycolytic response under a 20 g L⁻¹ glucose pulse

At this point, we investigated how our model could simultaneously explain the response to a strong glucose perturbation of a 20 g L⁻¹ glucose pulse experiment⁶, the moderate perturbations of 0.4 g L⁻¹ from FF experiments¹⁹, and the steady state data from¹⁴.

We identified two type of dynamics depending on if PFK and ALD reaction rates balanced each other (Fig 9A). By analyzing the experimental reaction rate data, we could see that during the moderate glucose perturbations in¹⁹, the glycolytic response of PFK and ALD occurred hand in hand, resulting in a transient increase of ATP (Fig 9B). Meanwhile, during the stronger perturbation in⁶ the response of ALD was slower (Fig 9C), leading to more ATP being consumed in upper glycolysis than recycled in lower glycolysis, and a consequent temporary decrease in ATP.

279 To understand what differed between the two conditions, we modelled both and found that they could
280 not be simultaneously understood with a single parameter set. A pareto analysis was performed to find
281 out which parameters needed to change to fit one and the other dataset. This resulted in the pareto front
282 described in (Fig 9D). If the steady state fructose 1,6-bisPhosphate (FBP) concentrations were reproduced,
283 ALD flux rapidly increased in response to the glucose perturbation. Another solution along the Pareto
284 front was obtained for which the increase in ALD flux was slower , resembling experimental data, but then,
285 the fit for FBP was not achieved (Fig 9E and 9F, respectively). Interestingly, this slower response resulted
286 from apparent substrate affinity constants in upper glycolysis strongly deviating from the literature value
287 and de facto caused enzymes to require a higher buildup of substrate before carrying out their reaction.
288 Altogether , this could imply that a mechanism is missing that regulates glycolytic response under strong
289 perturbations by, for instance, transiently decreasing ALD reaction rates. Since parameter values are closer
290 to the literature when the SS data is fit, which also results in a fast ALD response during FF, this seems to
291 suggest that our model is best suited to fit physiology from SS and FF regime.

292 **3 Discussion**

293 How cells respond and adapt to environmental changes while balancing internal needs prevails as a
294 fundamental question in biology. Mechanistic models based on enzyme kinetics are important to cope with
295 the complexity of the underlying molecular circuitry. For these models to be useful, however, they require
296 a proper intracellular embedding of central carbon metabolism in the context of the larger metabolic
297 system and should not be studies in isolation.

298 Hence, a physiology informed kinetic model of this pathway was developed. It contained detailed
299 descriptions of the enzymatic mechanisms that compose glycolysis, trehalose cycle and glycerol branch and
300 extended this knowledge with coarse-grained and growth rate-dependent phenomenological descriptions
301 of glycolytic intermediate demand, gas exchange, mitochondrial activity and ATPase activity (Fig 1).
302 This more was then used to characterize metabolic dynamics under both a 110 mM glucose pulse and at
303 different growth rates at steady state (Fig 2). Smaller models exist for *Escherichia coli* and *Penicillium*
304 *chrysogenum* metabolic models [52,53](#), but to our knowledge this is the first time a kinetic CCM model
305 can realistically describe a growing *S. cerevisiae* cell. Therefore, new circumstances are now open for
306 exploration with a unique model. For instance, by constraining the model to anaerobic conditions or by
307 studying the shift between respiratory and respirofermentative metabolism and glucose perturbations at
308 different growth rates [4,54,55](#).

309 One of the core developments of this work is a parameter estimation pipeline suited to identify complex
310 systems and using large data sets (Fig 3). This approach made use of different resources: (i) divide
311 and conquer model decomposition resulted in smaller scale systems easier to identify that the full
312 system, (ii) global sampling determined whether multiple local minima existed, (iii) regularization
313 improved convergence when parameters were still underdetermined and (iv) leveraging the weights of
314 the components of the cost function to avoid overfitting. As a result, this method increases parameter
315 identifiability, and we believe that it is well-suited for implementation in other cell factory models.

316 Compared to other state of the art parameter estimation methods [5,29,31,56,57](#), this pipeline makes use of the
317 divide and conquer approach, combined with regularization for parameters where physiological values are
318 known. Nonetheless, limitations of this approach are the lack of reliable parameter estimates confidence
319 intervals, given that the regularization bias added is affected by experimental noise, and the need for
320 leveraging the weights of these data in the cost function.

321 In order to develop this complex and physiologically representative model, multiple data types needed to

322 be combined. A collection of datasets containing metabolomic, fluxomic and proteomic data at steady
323 state and during glucose perturbations was considered to make the complete model identifiable^{6,14,28}.
324 Considering this data was required and increased the accuracy of the parameter estimates (Fig 4). We
325 therefore suggest making use of growingly available proteomics data to constrain the parameter solution
326 space, as it is the case for GSM^{20,37,58}.

327 Ensemble modelling suggested that the network dynamics are robust to parameter heterogeneity. A
328 modelling approach based on Monte Carlo sampling of parameter values was used to represent the
329 heterogeneity in a cell population. This showed how dynamics and steady states were resilient to
330 parameter value deviations (Fig 5). Assays such as this start with the premise that phenotypes arise
331 from model structure and are performed to represent population heterogeneity⁵⁹⁻⁶¹. Therefore, given
332 the heterogeneity in the industrial fermenter³, we propose that this ensemble modelling perspective is a
333 more physiologically realistic representation than a single parameter set. Nonetheless, one limitation of
334 this approach is the range to which parameters are disturbed, given that cell-specific measurements in a
335 population are very limited⁶².

336 Furthermore, we advocate to estimate parameters and supplement with literature knowledge for determina-
337 tion of *in vivo* enzymatic kinetic constants. Enzymatic kinetic constants were traditionally estimated *in*
338 *vitro* for isolated enzymes^{4,5,12}. However, using these to simulate experimental data showed how a notable
339 bias still exists (Fig 6). Hence, we propose to re-estimate parameters to fit the *in vivo* metabolomic and
340 fluxomic data, in line with²¹, and use the *in vitro* determined constants as reference to which the cost
341 function is regularized. The resulting improved data fit with generally little deviation from the literature
342 parameters implies that, even though *in vitro* studies are good estimators, they need adjustment to be
343 implemented in a full-scale model.

344 When estimated parameters widely deviated from the literature values or when steady state data was not
345 properly fit, we identified uncertainty in our model that reflected possible unknown biology. The estimated
346 parameters that deviated the most from the literature value were from PFK or surrounding enzymes PGI
347 and ALD (Table 1). Together with the fact that steady state concentrations of G6P and F6P are only fit in
348 the model decomposition simulations (Fig 6) and not in the complete model simulations (Fig 5), this hints
349 at uncertainty surrounding PFK reaction kinetics, not covered in the complex regulation already considered
350 for this reaction^{12,46}. Furthermore, another area where steady state data is not perfectly fit is the HXT
351 reaction. HXT kinetics could be explained by changes in enzyme concentration and isoenzyme-specific
352 affinity constants, in line with^{48,51}. Nonetheless, the consistent lack of fit at 0.2 h⁻¹ points at a missing

353 mechanism acting intracellularly⁴⁷.

354 Finally, missing regulation could explain the transient imbalance occurring between PFK and ALD upon
355 strong glucose perturbation (Fig 9). When explaining this imbalance with the same model structure, we
356 found that a subset of parameters had to be adjusted, leading to a lower ALD reaction rate. Nonetheless, in
357 the industrial bioreactor, prolonged exposure to glucose concentrations higher than 100 mM as in⁶ is rare⁶³.
358 A more realistic representation is the FF setup, in which this imbalance did not take place under moderate
359 perturbations in FF experiments. This could hint at influence by glucose level¹⁸ and suggest that there is a
360 missing phenomenon in the model, which becomes active upon large glucose perturbations. We believe
361 that this could be due to either a regulation cascade, such as post translation modifications or cAMP/PKA
362 pathway^{64,65}, given that cAMP buildup was detected during perturbation with high residual concentration
363 but not for low^{18,66}. Consequently, the synthesis of fructose-2,6-bisphosphate by the activated PKA
364 pathway could be regulating could be regulating PFK activity. In addition, the pH decay observed during
365 the glucose perturbation⁶ could also be relevant, given the sensitivity of glycolytic enzyme constants to
366 this variable^{67,68}. Nonetheless, to further proof these hypothesis, more experimental testing would be
367 needed.

368 Here, we presented a model that describes yeast glycolysis as part of the more comprehensive central
369 carbon metabolism, and not in isolation. This model was built using data from the metabolome, fluxome
370 and proteome; to which extent gene regulation influenced the metabolic dynamics here studied remains
371 an open question. At this point, gradients in the industrial bioreactor can be considered thanks to the
372 transport rates and yields included in the model^{3,69}, bringing it closer to industrial applications. We believe
373 that complex intracellular models like the hereby presented could soon be linked to bioreactor dynamics
374 simulations, an area that has been so far restricted to simplified cell models^{8,53}.

375 **4 Conclusions**

376 Academia and industry seek to explain how the metabolism of microbial cell factories responds to changing
377 environments. Kinetic metabolic models are great tools to answer this question, but their development
378 is hampered by their complex nature. Here, we set to develop such a model for the microorganism
379 *Saccharomyces cerevisiae*. By embedding central carbon metabolism in the context of the larger metabolic
380 system, a more realistic and physiology informed model was developed. To properly parameterize this
381 large model, data sets from steady state conditions and glucose perturbations were necessary. Furthermore,
382 to tackle the complexity issue, a new parameter estimation pipeline is developed consisting of a divide
383 and conquer approach, regularization and global optimization. Our work provides a tool that can be used
384 for rational metabolic engineering of *S. cerevisiae*. This model can now be expanded to explore new
385 experimental conditions. Nonetheless, due to how relevant and conserved central metabolism is, this work
386 could also be used as a chassis to study other microbial cell factories.

387 **5 Materials and methods**

388 **5.1 Experimental data sets used to develop the model**

389 Three experimental datasets were used in model development: (i) steady state metabolic concentrations
390 and fluxes at different chemostat dilution rates ($0.025 - 0.375 \text{ h}^{-1}$)¹⁴, (ii) concentrations and fluxes during
391 a single glucose perturbation of 20 g L^{-1} (van Heerden et al., 2014) and (iii) glycolytic enzymes activity
392 also at several chemostat dilution rates²⁸. Data in^{6,14} was obtained with the haploid yeast *Saccharomyces*
393 *cerevisiae* CEN PK 113-7D strain, while²⁸ used the strain DS28911.

394 **5.2 Model description**

395 The kinetic metabolic model in this work included individual enzyme reaction descriptions for glycolysis,
396 glycerol branch and trehalose cycle, whose kinetics were taken from^{6, 70} and⁷¹, respectively. The
397 reaction of UDP-glucose phosphorylase in the trehalose cycle was lumped due to lack of experimental
398 data. Cofactor metabolism reactions and sinks for anabolic precursors were lumped, and the sink reactions
399 and inosine salvage pathway were adapted from⁵² and¹⁵, respectively.

400 **5.3 System of ordinary differential equations and reaction rate equations**

401 The model consisted of a series of ordinary differential equations representing the mass balances for
402 each metabolite modelled in the cytosol, except glucose and inorganic phosphate, also modelled in the
403 extracellular space and vacuole. Moiety conservations were used for the total sum of adenosine and
404 nicotinamide adenine nucleotides (ATP + ADP + AMP and NAD + NADH, respectively) as in⁵. The ATP
405 + ADP + AMP moiety conservation was not considered during the single glucose perturbation response,
406 where the inosine salvage pathway was included as a pool. When available, the dilution rate-dependent
407 protein activity change²⁸ was considered by adjusting the V_{max} .

408 Enzymatic reaction rate kinetics were obtained from previous works, where reversible Michaelis-Menten
409 kinetics dominated^{6,70,71}. Exceptions to this were hill type kinetics for pyruvate kinase and decarboxylase⁴,
410 phosphofructokinase and facilitated diffusion for membrane transport¹². Allosteric regulation acted both
411 by activation and (competitive) inhibition. Hydrolysis reactions were modelled as irreversible. Sink
412 reactions were modelled by phenomenological expressions that closely resembled experimental data in¹⁴.
413 Reaction rates were expressed in (mM s^{-1}). For a detailed description, see supplementary information.

414 **5.4 Simulation setup**

415 Simulations were performed in agreement with the experimental setup. To confirm simulation stability,
416 the model was first simulated for 3000 seconds using the experimental concentrations at 0.1 h^{-1} dilution
417 rate. Then, the steady states at different dilution rates were modelled in parallel simulations where the
418 residual glucose concentration was changed to the value in¹⁴, for 3000 seconds. The trehalose cycle was
419 not modelled in steady state. Anabolic sink reaction rates, ATP maintenance and the mitochondrial activity
420 were adjusted in a dilution rate-dependent manner as well. For the single glucose perturbation, residual
421 glucose was increased to 110 mM⁶ and the inosine salvage pathway was made active in a step manner.
422 This simulation lasted 340 seconds. Matlab version 9.3.0.713579 (R2017b) was used. A summary of
423 the differences between simulating steady states at different growth rates of the 110 mM single glucose
424 perturbation can be seen in Table 2. Detailed explanation on these can be found in the supplementary
425 information.

426 **5.5 Literature parameter values**

427 Initial estimates of kinetic constants were obtained from the literature. For glycolytic enzymes,⁴ and⁵
428 were the most recent estimates of V_{max} and K_m . If a parameter was not available, it was taken from¹².
429 Glycerol branch and trehalose cycle parameters were retrieved from⁷¹ and⁷⁰, respectively. For the specific
430 parameter values used, see supplementary materials.

431 **5.6 In vivo parameter estimation and cost function development**

432 Parameters in the model were estimated to fit the in vivo experimental data. The lsqnonlin solver in the
433 Matlab optimization toolbox (version 9.3.0.713579, R2017b), which uses an interior reflective Newton
434 method⁷², was used to minimize the error between experimental and simulated data. To overcome the
435 effects of ill conditioning and parameter dependencies in this large parameter set³¹, a model decomposition
436 approach, also known as divide-and-conquer, was used³⁰. This was first implemented to individual
437 reactions in glycolysis. Once parameters were estimated for them, parameters were estimated for other
438 pathways. The global solution space was explored by means of multi-start deterministic local searches²⁹.
439 This model fit was supplemented with L1-type regularization^{57,73}, where kinetic constants were biased to
440 resemble the experimental measurement, as long as data fit was adequate. Steady state data¹⁴ was fit first
441 and, afterwards, single glucose perturbation⁶.

⁴⁴² **Author contributions**

⁴⁴³ **David Lao-Martil:** Conceptualization; Validation; Investigation; Methodology; Writing - originak draft;
⁴⁴⁴ Writing - review and editing. **Joep Schmitz:** Conceptualization; Investigation; Methodology; Writing
⁴⁴⁵ - review and editing. **Bas Teusink:** Conceptualization; Writing - review and editing. **Natal van Riel:**
⁴⁴⁶ Conceptualization; Writing; Methodology - review and editing.

⁴⁴⁷ **Acknowledgements**

⁴⁴⁸ We thank our colleagues from the Yeast 3M team, S.Aljoscha Wahl, Koen J.A. Verhagen, Laura Guilherme
⁴⁴⁹ Luzia, Evelina Tutucci and Johann van Heerden, for the insightful discussions.

⁴⁵⁰ **Data and model availability statement**

⁴⁵¹ This work was developed in MATLAB and model implementations are also available in Python language
⁴⁵² and SBML format. The data used and model developed in this work are available in the github repository
⁴⁵³ from the author (github.com/DavidLaoM/y3m1_ss_gp) and the CBio group (github.com/Computational-Biology-TUE).
⁴⁵⁴

⁴⁵⁵ **Competing interests statement**

⁴⁵⁶ The authors declare no conflict of interest.

⁴⁵⁷ **Funding information**

⁴⁵⁸ Dutch Research Council (NWO). Project number: 737.016.00.

459 **References**

- 460 1. Nielsen, J., Larsson, C., van Maris, A., and Pronk, J. (2013) Metabolic engineering of yeast for
461 production of fuels and chemicals. *Current opinion in biotechnology*, **24**(3), 398–404.
- 462 2. Enfors, S.-O., Jahic, M., Rozkov, A., Xu, B., Hecker, M., Jürgen, B., Krüger, E., Schweder, T., Hamer,
463 G., O'beirne, D., et al. (2001) Physiological responses to mixing in large scale bioreactors. *Journal of
464 biotechnology*, **85**(2), 175–185.
- 465 3. Haringa, C., Tang, W., Deshmukh, A. T., Xia, J., Reuss, M., Heijnen, J. J., Mudde, R. F., and Noorman,
466 H. J. (2016) Euler-Lagrange computational fluid dynamics for (bio) reactor scale down: an analysis of
467 organism lifelines. *Engineering in life sciences*, **16**(7), 652–663.
- 468 4. van Eunen, K., Kiewiet, J. A., Westerhoff, H. V., and Bakker, B. M. (2012) Testing biochemistry revis-
469 ited: how in vivo metabolism can be understood from in vitro enzyme kinetics. *PLoS computational
470 biology*, **8**(4), e1002483.
- 471 5. Smallbone, K., Messiha, H. L., Carroll, K. M., Winder, C. L., Malys, N., Dunn, W. B., Murabito, E.,
472 Swainston, N., Dada, J. O., Khan, F., et al. (2013) A model of yeast glycolysis based on a consistent
473 kinetic characterisation of all its enzymes. *FEBS letters*, **587**(17), 2832–2841.
- 474 6. van Heerden, J. H., Wortel, M. T., Bruggeman, F. J., Heijnen, J. J., Bollen, Y. J., Planqué, R., Hulshof,
475 J., O'Toole, T. G., Wahl, S. A., and Teusink, B. (2014) Lost in transition: start-up of glycolysis yields
476 subpopulations of nongrowing cells. *Science*, **343**(6174), 1245114.
- 477 7. Wang, G., Haringa, C., Noorman, H., Chu, J., and Zhuang, Y. (2020) Developing a computational
478 framework to advance bioprocess scale-up. *Trends in Biotechnology*, **38**(8), 846–856.
- 479 8. Sarkizi Shams Hajian, C., Haringa, C., Noorman, H., and Takors, R. (2020) Predicting by-product
480 gradients of baker's yeast production at industrial scale: A practical simulation approach. *Processes*,
481 **8**(12), 1554.
- 482 9. Pruett, W. A., Clemmer, J. S., and Hester, R. L. (2020) Physiological Modeling and Simula-
483 tion—Validation, Credibility, and Application. *Annual review of biomedical engineering*, **22**, 185–206.
- 484 10. Yasemi, M. and Jolicoeur, M. (2021) Modelling cell metabolism: a review on constraint-based
485 steady-state and kinetic approaches. *Processes*, **9**(2), 322.

- 486 11. van Heerden, J. H., Bruggeman, F. J., and Teusink, B. (2015) Multi-tasking of biosynthetic and
487 energetic functions of glycolysis explained by supply and demand logic. *BioEssays*, **37**(1), 34–45.
- 488 12. Teusink, B., Passarge, J., Reijenga, C. A., Esgalhado, E., Van der Weijden, C. C., Schepper, M.,
489 Walsh, M. C., Bakker, B. M., Van Dam, K., Westerhoff, H. V., et al. (2000) Can yeast glycolysis be
490 understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. *European*
491 *Journal of Biochemistry*, **267**(17), 5313–5329.
- 492 13. Otterstedt, K., Larsson, C., Bill, R. M., Ståhlberg, A., Boles, E., Hohmann, S., and Gustafsson, L.
493 (2004) Switching the mode of metabolism in the yeast *Saccharomyces cerevisiae*. *EMBO reports*,
494 **5**(5), 532–537.
- 495 14. Canelas, A. B., Ras, C., ten Pierick, A., van Gulik, W. M., and Heijnen, J. J. (2011) An in vivo
496 data-driven framework for classification and quantification of enzyme kinetics and determination of
497 apparent thermodynamic data. *Metabolic engineering*, **13**(3), 294–306.
- 498 15. Walther, T., Novo, M., Rössger, K., Létisse, F., Loret, M.-O., Portais, J.-C., and François, J.-M. (2010)
499 Control of ATP homeostasis during the respiro-fermentative transition in yeast. *Molecular systems*
500 *biology*, **6**(1), 344.
- 501 16. Chen, Y. and Nielsen, J. (2019) Energy metabolism controls phenotypes by protein efficiency and
502 allocation. *Proceedings of the National Academy of Sciences*, **116**(35), 17592–17597.
- 503 17. Theobald, U., Mailinger, W., Baltes, M., Rizzi, M., and Reuss, M. (1997) In vivo analysis of
504 metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. *Biotechnology and*
505 *bioengineering*, **55**(2), 305–316.
- 506 18. Suarez-Mendez, C. A., Sousa, A., Heijnen, J. J., and Wahl, A. (2014) Fast “feast/famine” cycles for
507 studying microbial physiology under dynamic conditions: a case study with *Saccharomyces cerevisiae*.
508 *Metabolites*, **4**(2), 347–372.
- 509 19. Suarez-Mendez, C., Ras, C., and Wahl, S. (2017) Metabolic adjustment upon repetitive substrate
510 perturbations using dynamic ¹³C-tracing in yeast. *Microbial Cell Factories*, **16**(1), 1–14.
- 511 20. Elsemman, I. E., Rodriguez Prado, A., Grigaitis, P., Garcia Albornoz, M., Harman, V., Holman, S. W.,
512 van Heerden, J., Bruggeman, F. J., Bisschops, M. M., Sonnenschein, N., et al. (2022) Whole-cell
513 modeling in yeast predicts compartment-specific proteome constraints that drive metabolic strategies.
514 *Nature communications*, **13**(1), 1–12.

- 515 21. Davidi, D. and Milo, R. (2017) Lessons on enzyme kinetics from quantitative proteomics. *Current*
516 *opinion in biotechnology*, **46**, 81–89.
- 517 22. Pritchard, L. and Kell, D. B. (2002) Schemes of flux control in a model of *Saccharomyces cerevisiae*
518 glycolysis. *European journal of biochemistry*, **269**(16), 3894–3904.
- 519 23. Van Riel, N., Giuseppin, M., TerSchure, E., and Verrips, C. (1998) A Structured, Minimal parameter
520 Model of the Central Nitrogen Metabolism in *Saccharomyces cerevisiae*: The Prediction of the
521 Behaviour of Mutants. *Journal of theoretical biology*, **191**(4), 397–414.
- 522 24. Lao-Martil, D., Verhagen, K. J., Schmitz, J. P., Teusink, B., Wahl, S. A., and van Riel, N. A.
523 (2022) Kinetic Modeling of *Saccharomyces cerevisiae* Central Carbon Metabolism: Achievements,
524 Limitations, and Opportunities. *Metabolites*, **12**(1), 74.
- 525 25. Beeftink, H., Van der Heijden, R., and Heijnen, J. (1990) Maintenance requirements: energy supply
526 from simultaneous endogenous respiration and substrate consumption. *FEMS Microbiology Ecology*,
527 **6**(3), 203–209.
- 528 26. Lu, H., Li, F., Sánchez, B. J., Zhu, Z., Li, G., Domenzain, I., Marcišauskas, S., Anton, P. M., Lappa,
529 D., Lieven, C., et al. (2019) A consensus *S. cerevisiae* metabolic model Yeast8 and its ecosystem for
530 comprehensively probing cellular metabolism. *Nature communications*, **10**(1), 1–13.
- 531 27. Verduyn, C., Stouthamer, A. H., Scheffers, W. A., and van Dijken, J. P. (1991) A theoretical evaluation
532 of growth yields of yeasts. *Antonie Van Leeuwenhoek*, **59**(1), 49–63.
- 533 28. van Hoek, P., van Dijken, J. P., and Pronk, J. T. (2000) Regulation of fermentative capacity and levels
534 of glycolytic enzymes in chemostat cultures of *Saccharomyces cerevisiae*. *Enzyme and microbial*
535 *technology*, **26**(9-10), 724–736.
- 536 29. Villaverde, A. F., Fröhlich, F., Weindl, D., Hasenauer, J., and Banga, J. R. (2019) Benchmarking
537 optimization methods for parameter estimation in large kinetic models. *Bioinformatics*, **35**(5), 830–
538 838.
- 539 30. Kotte, O. and Heinemann, M. (2009) A divide-and-conquer approach to analyze underdetermined
540 biochemical models. *Bioinformatics*, **25**(4), 519–525.
- 541 31. Gábor, A. and Banga, J. R. (2015) Robust and efficient parameter estimation in dynamic models of
542 biological systems. *BMC systems biology*, **9**(1), 1–25.

- 543 32. Moles, C. G., Mendes, P., and Banga, J. R. (2003) Parameter estimation in biochemical pathways: a
544 comparison of global optimization methods. *Genome research*, **13**(11), 2467–2474.
- 545 33. Engl, H. W., Flamm, C., Kügler, P., Lu, J., Müller, S., and Schuster, P. (2009) Inverse problems in
546 systems biology. *Inverse Problems*, **25**(12), 123014.
- 547 34. Kreutz, C. (2016) New concepts for evaluating the performance of computational methods. *IFAC-
548 PapersOnLine*, **49**(26), 63–70.
- 549 35. Almquist, J., Cvijovic, M., Hatzimanikatis, V., Nielsen, J., and Jirstrand, M. (2014) Kinetic models in
550 industrial biotechnology—improving cell factory performance. *Metabolic engineering*, **24**, 38–60.
- 551 36. Peskov, K., Mogilevskaya, E., and Demin, O. (2012) Kinetic modelling of central carbon metabolism
552 in *Escherichia coli*. *The FEBS journal*, **279**(18), 3374–3385.
- 553 37. Sánchez, B. J., Zhang, C., Nilsson, A., Lahtvee, P.-J., Kerkhoven, E. J., and Nielsen, J. (2017)
554 Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating
555 enzymatic constraints. *Molecular systems biology*, **13**(8), 935.
- 556 38. Nidelet, T., Brial, P., Camarasa, C., and Dequin, S. (2016) Diversity of flux distribution in central
557 carbon metabolism of *S. cerevisiae* strains from diverse environments. *Microbial cell factories*, **15**(1),
558 1–13.
- 559 39. Rodrigues, C. I. S., Wahl, A., and Gombert, A. K. (2021) Aerobic growth physiology of *Saccha-*
560 *romyces cerevisiae* on sucrose is strain-dependent. *FEMS yeast research*, **21**(3), foab021.
- 561 40. Tan, Y., Rivera, J. G. L., Contador, C. A., Asenjo, J. A., and Liao, J. C. (2011) Reducing the allowable
562 kinetic space by constructing ensemble of dynamic models with the same steady-state flux. *Metabolic
563 engineering*, **13**(1), 60–75.
- 564 41. Jia, G., Stephanopoulos, G., and Gunawan, R. (2012) Ensemble kinetic modeling of metabolic
565 networks from dynamic metabolic profiles. *Metabolites*, **2**(4), 891–912.
- 566 42. Tiemann, C. A., Vanlier, J., Oosterveer, M. H., Groen, A. K., Hilbers, P. A., and van Riel, N. A. (2013)
567 Parameter trajectory analysis to identify treatment effects of pharmacological interventions. *PLoS
568 computational biology*, **9**(8), e1003166.
- 569 43. Tran, L. M., Rizk, M. L., and Liao, J. C. (2008) Ensemble modeling of metabolic networks. *Biophysical
570 journal*, **95**(12), 5606–5617.

- 571 44. Christodoulou, D., Link, H., Fuhrer, T., Kochanowski, K., Gerosa, L., and Sauer, U. (2018) Reserve
572 flux capacity in the pentose phosphate pathway enables Escherichia coli's rapid response to oxidative
573 stress. *Cell systems*, **6**(5), 569–578.
- 574 45. Van Eunen, K., Bouwman, J., Daran-Lapujade, P., Postmus, J., Canelas, A. B., Mensonides, F. I.,
575 Orij, R., Tuzun, I., Van Den Brink, J., Smits, G. J., et al. (2010) Measuring enzyme activities under
576 standardized in vivo-like conditions for systems biology. *The FEBS journal*, **277**(3), 749–760.
- 577 46. Gustavsson, A.-K., van Niekerk, D. D., Adiels, C. B., Kooi, B., Goksör, M., and Snoep, J. L. (2014)
578 Allosteric regulation of phosphofructokinase controls the emergence of glycolytic oscillations in
579 isolated yeast cells. *The FEBS journal*, **281**(12), 2784–2793.
- 580 47. Postma, E., Alexander Scheffers, W., and Van Dijken, J. P. (1989) Kinetics of growth and glucose
581 transport in glucose-limited chemostat cultures of *Saccharomyces cerevisiae* CBS 8066. *Yeast*, **5**(3),
582 159–165.
- 583 48. Diderich, J. A., Schepper, M., van Hoek, P., Luttik, M. A., van Dijken, J. P., Pronk, J. T., Klaassen,
584 P., Boelens, H. F., de Mattos, M. J. T., van Dam, K., et al. (1999) Glucose uptake kinetics and
585 transcription of HXTGenes in chemostat cultures of *Saccharomyces cerevisiae*. *Journal of Biological
586 Chemistry*, **274**(22), 15350–15359.
- 587 49. Bosdriesz, E., Wortel, M. T., Haanstra, J. R., Wagner, M. J., De La Torre Cortés, P., and Teusink, B.
588 (2018) Low affinity uniporter carrier proteins can increase net substrate uptake rate by reducing efflux.
589 *Scientific reports*, **8**(1), 1–9.
- 590 50. Maier, A., Völker, B., Boles, E., and Fuhrmann, G. F. (2002) Characterisation of glucose transport in
591 *Saccharomyces cerevisiae* with plasma membrane vesicles (countertransport) and intact cells (initial
592 uptake) with single Hxt1, Hxt2, Hxt3, Hxt4, Hxt6, Hxt7 or Gal2 transporters. *FEMS yeast research*,
593 **2**(4), 539–550.
- 594 51. Reifenberger, E., Boles, E., and Ciriacy, M. (1997) Kinetic characterization of individual hexose
595 transporters of *Saccharomyces cerevisiae* and their relation to the triggering mechanisms of glucose
596 repression. *European journal of biochemistry*, **245**(2), 324–333.
- 597 52. Chassagnole, C., Noisommit-Rizzi, N., Schmid, J. W., Mauch, K., and Reuss, M. (2002) Dynamic
598 modeling of the central carbon metabolism of *Escherichia coli*. *Biotechnology and bioengineering*,
599 **79**(1), 53–73.

- 600 53. Tang, W., Deshmukh, A. T., Haringa, C., Wang, G., van Gulik, W., van Winden, W., Reuss, M.,
601 Heijnen, J. J., Xia, J., Chu, J., et al. (2017) A 9-pool metabolic structured kinetic model describing days
602 to seconds dynamics of growth and product formation by *Penicillium chrysogenum*. *Biotechnology*
603 and *Bioengineering*, **114**(8), 1733–1743.
- 604 54. Wiebe, M. G., Rintala, E., Tamminen, A., Simolin, H., Salusjärvi, L., Toivari, M., Kokkonen, J. T.,
605 Kiuru, J., Ketola, R. A., Jouhten, P., et al. (2008) Central carbon metabolism of *Saccharomyces*
606 cerevisiae in anaerobic, oxygen-limited and fully aerobic steady-state conditions and following a shift
607 to anaerobic conditions. *FEMS Yeast research*, **8**(1), 140–154.
- 608 55. Ishtar Snoek, I. and Yde Steensma, H. (2007) Factors involved in anaerobic growth of *Saccharomyces*
609 cerevisiae. *Yeast*, **24**(1), 1–10.
- 610 56. Raue, A., Steiert, B., Schelker, M., Kreutz, C., Maiwald, T., Hass, H., Vanlier, J., Tönsing, C.,
611 Adlung, L., Engesser, R., et al. (2015) Data2Dynamics: a modeling environment tailored to parameter
612 estimation in dynamical systems. *Bioinformatics*, **31**(21), 3558–3560.
- 613 57. Steiert, B., Timmer, J., and Kreutz, C. (2016) L 1 regularization facilitates detection of cell type-
614 specific parameters in dynamical systems. *Bioinformatics*, **32**(17), i718–i726.
- 615 58. Chan, S. H., Cai, J., Wang, L., Simons-Senftle, M. N., and Maranas, C. D. (2017) Standardizing
616 biomass reactions and ensuring complete mass balance in genome-scale metabolic models. *Bioinfor-*
617 *matics*, **33**(22), 3603–3609.
- 618 59. Gunawardena, J. (2010) Models in systems biology: the parameter problem and the meanings of
619 robustness. *Elements of computational systems biology*, **1**, 21–47.
- 620 60. Oguz, C., Watson, L. T., Baumann, W. T., and Tyson, J. J. (2017) Predicting network modules of cell
621 cycle regulators using relative protein abundance statistics. *BMC systems biology*, **11**(1), 1–24.
- 622 61. Nijhout, H. F., Best, J. A., and Reed, M. C. (2019) Systems biology of robustness and homeostatic
623 mechanisms. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, **11**(3), e1440.
- 624 62. Newman, J. R., Ghaemmaghami, S., Ihmels, J., Breslow, D. K., Noble, M., DeRisi, J. L., and
625 Weissman, J. S. (2006) Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of
626 biological noise. *Nature*, **441**(7095), 840–846.

- 627 63. Haringa, C., Deshmukh, A. T., Mudde, R. F., and Noorman, H. J. (2017) Euler-Lagrange analy-
628 sis towards representative down-scaling of a 22 m³ aerobic *S. cerevisiae* fermentation. *Chemical*
629 *Engineering Science*, **170**, 653–669.
- 630 64. Tripodi, F., Nicastro, R., Reghellin, V., and Coccetti, P. (2015) Post-translational modifications
631 on yeast carbon metabolism: regulatory mechanisms beyond transcriptional control. *Biochimica et*
632 *Biophysica Acta (BBA)-General Subjects*, **1850**(4), 620–627.
- 633 65. Tamaki, H. (2007) Glucose-stimulated cAMP-protein kinase A pathway in yeast *Saccharomyces*
634 *cerevisiae*. *Journal of bioscience and bioengineering*, **104**(4), 245–250.
- 635 66. Botman, D., O'Toole, T. G., Goedhart, J., Bruggeman, F. J., van Heerden, J. H., and Teusink, B.
636 (2019) A yeast FRET biosensor enlightens cAMP signalling. *bioRxiv*, p. 831354.
- 637 67. Van Leemputte, F., Vanthienen, W., Wijnants, S., Van Zeebroeck, G., and Thevelein, J. M. (2020)
638 Aberrant Intracellular pH Regulation Limiting Glyceraldehyde-3-Phosphate Dehydrogenase Activity
639 in the Glucose-Sensitive Yeast *tps1* Δ Mutant. *Mbio*, **11**(5), e02199–20.
- 640 68. Luzia, L., Lao-Martil, D., Savakis, P., van Heerden, J., van Riel, N., and Teusink, B. (2022) pH
641 dependencies of glycolytic enzymes of yeast under in vivo-like assay conditions. *The FEBS Journal*,.
- 642 69. Nadal-Rey, G., McClure, D. D., Kavanagh, J. M., Cornelissen, S., Fletcher, D. F., and Gernaey, K. V.
643 (2021) Understanding gradients in industrial bioreactors. *Biotechnology Advances*, **46**, 107660.
- 644 70. Smallbone, K., Malys, N., Messiha, H. L., Wishart, J. A., and Simeonidis, E. (2011) Building a kinetic
645 model of trehalose biosynthesis in *Saccharomyces cerevisiae*. *Methods in Enzymology*, **500**, 355–370.
- 646 71. Cronwright, G. R., Rohwer, J. M., and Prior, B. A. (2002) Metabolic control analysis of glycerol
647 synthesis in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, **68**(9), 4448–4456.
- 648 72. Coleman, T. F. and Li, Y. (1996) A reflective Newton method for minimizing a quadratic function
649 subject to bounds on some of the variables. *SIAM Journal on Optimization*, **6**(4), 1040–1058.
- 650 73. Dolejsch, P., Hass, H., and Timmer, J. (2019) Extensions of l1 regularization increase detection
651 specificity for cell-type specific parameters in dynamic models. *BMC bioinformatics*, **20**(1), 1–13.

652 **Figure and Table Legends**

Table 1. Largest deviations between estimated and literature values are found in PGI, PFK and ALD enzymes. Only deviations above 0.75-magnitude order change from the literature value are displayed.

Parameter	Enzyme	Magnitude order change	Reference
$K_{m,DHAP}$	ALD	-1.90	12
$K_{m,F6P}$	PGI	1.37	5
K_{AMP}	PFK	-0.98	12
$K_{m,G6P}$	PGI	0.89	5
$K_{m,ATP}$	PFK	0.88	12
$K_{i,FBP}$	PFK	0.77	12
K_{cat}	PGI	-0.76	4

Table 2. Summary of main differences in the simulation setup of the steady state and the single glucose perturbation.

Variable	Steady state	Single Glucose Perturbation
Trehalose cycle	Inactive. Its steady state flux is included in the G6P-sink reaction	Active.
Inosine salvage pathway	Inactive.	Active.
sumAXP (= ATP + ADP + AMP)	Changes in the total sum of ATP + ADP + AMP are considered by setting the initial concentrations of the adenosine nucleotides to the experimental concentration at the beginning of each steady state simulation.	The inosine salvage pathway acts as a sink of adenosine nucleotides.
Enzyme concentration	A ratio between the experimental activity at a given dilution rate and the value at 0.1h^{-1} is calculated and then multiplied to the V_{max} .	This ratio equals 1.
Sink reactions	A phenomenological expression is derived to make these reaction rates resemble experimental data.	The kinetics at 0.1 h^{-1} is used.
ATPase activity	ATPase is adjusted in a growth rate dependent manner, using GAM and NGAM.	The reaction constant is estimated to fit the data.
mitoATP	The growth rate-dependent mitochondrial activity is considered using the PYR sink reaction. This is implemented by changes in the reaction constant.	The reaction constant is estimated to fit the data.
mitoN-ADH	The growth rate-dependent mitochondrial activity is considered using the PYR sink reaction. This is implemented by changes in the reaction constant.	The reaction constant is estimated to fit the data.

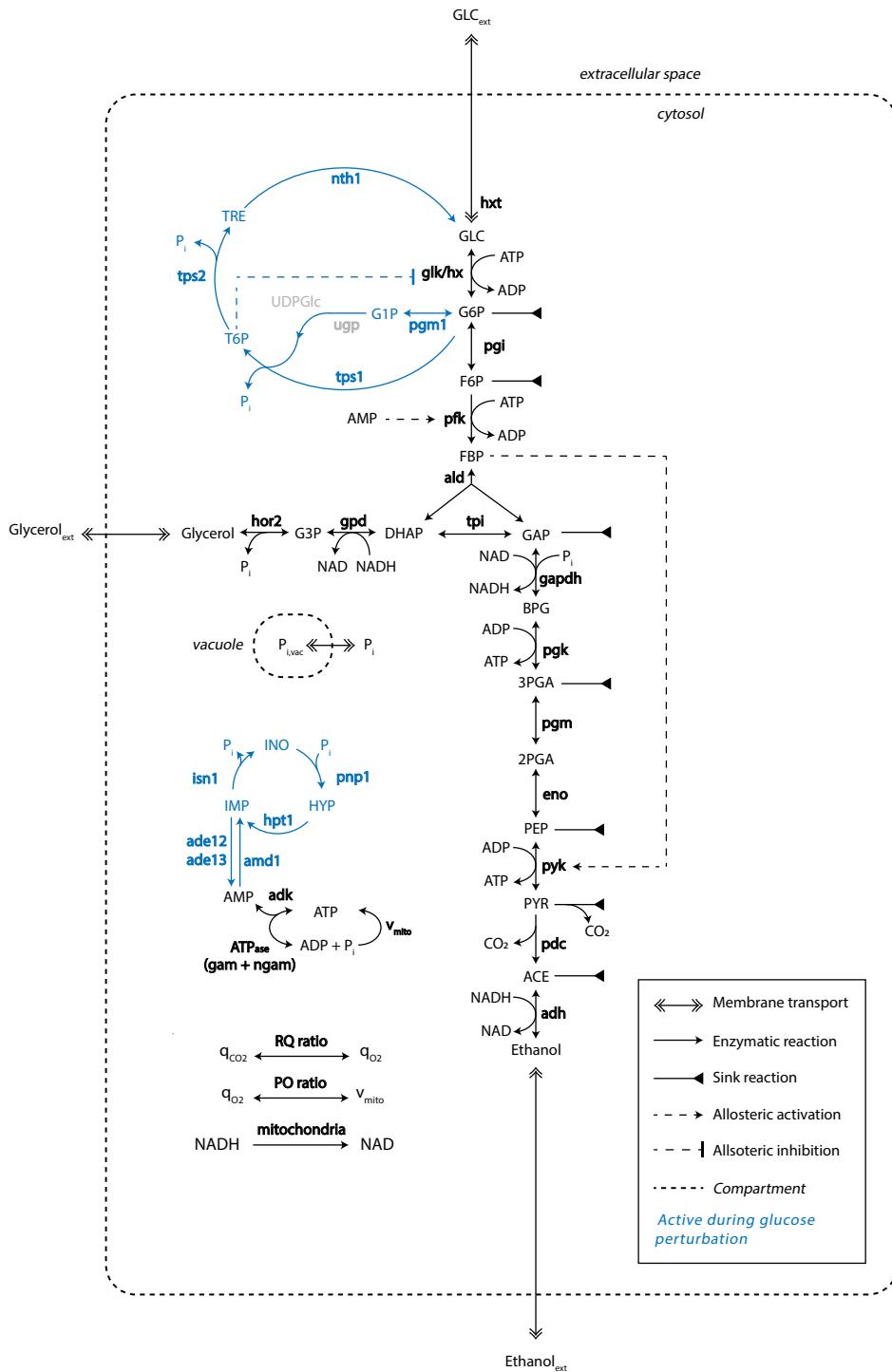


Figure 1. An updated central carbon metabolism *S. cerevisiae* model considering variables of physiological relevance. Black colored reactions are active for all simulations, blue only during the glucose perturbation and gray are lumped. Metabolites are connected by the reactions in the model (continuous lines), which are catalysed by enzymes (bold). Allosteric regulation is shown by semi-continuous arrows and compartments by semi-continuous lines.

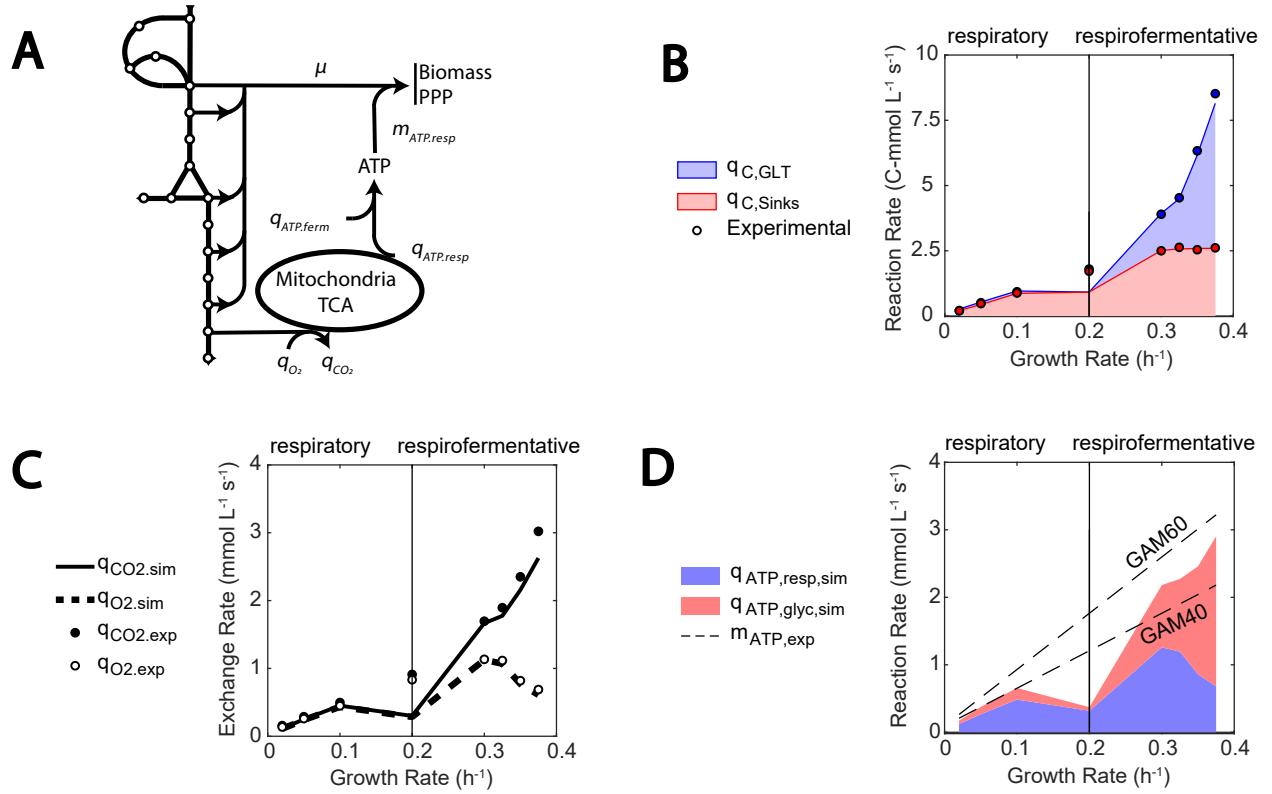


Figure 2. The model reproduces physiological properties of a growing cell in a bioreactor. (A)

Simplified model scheme highlighting the role of sinks for anabolic reaction precursors, PPP and mitochondria (g6p, f6p, gap, pep, pyr, ace). (B) Carbon flux entering the sink reactions (red) compared to the total carbon uptake (blue). (C) Simulated exchange rate of CO_2 (continuous line) and O_2 (semi-continuous line). Experimental data shown as bullet points. (D) ATP produced by respiration (blue area) and fermentation (red area). Semi continuous lines point at the theoretical values calculated using the NGAM and GAM (40 and 60 mmol g DW^{-1}) found in genome scale representations²⁶.

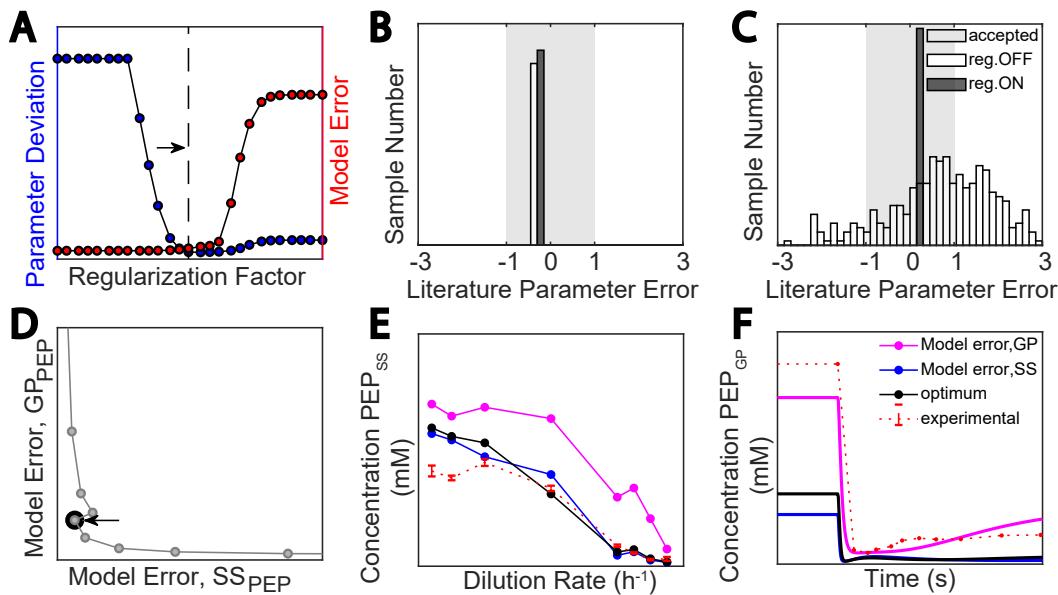


Figure 3. A problem decomposition approach supplemented by regularization, parameter sampling space and weighting data types. (A) Regularization approach: the regularization factor (x-axis), literature parameter deviation (blue, left y-axis) and model error (red, right y-axis) (B and C) Parameter estimation case for an identifiable (Enolase, K_{cat}) and weakly identifiable parameter (Phosphofructokinase, K_{F6P}), respectively. Histogram plot showing the estimated parameter deviation from the literature value (x-axis) and the number of initial parameter samples in the y-axis. Non-regularized and regularized cases in white and dark bars, respectively. The shaded area covers the range of one order of magnitude of deviation from the literature parameter value. (D) Pareto front commonly obtained when fitting two types of data simultaneously: case of PEP concentration. Model error for steady state and glucose perturbation concentrations of PEP (x-axis and y-axis, respectively). (E and F) Model error, PEP steady state and glucose perturbation concentrations, respectively.

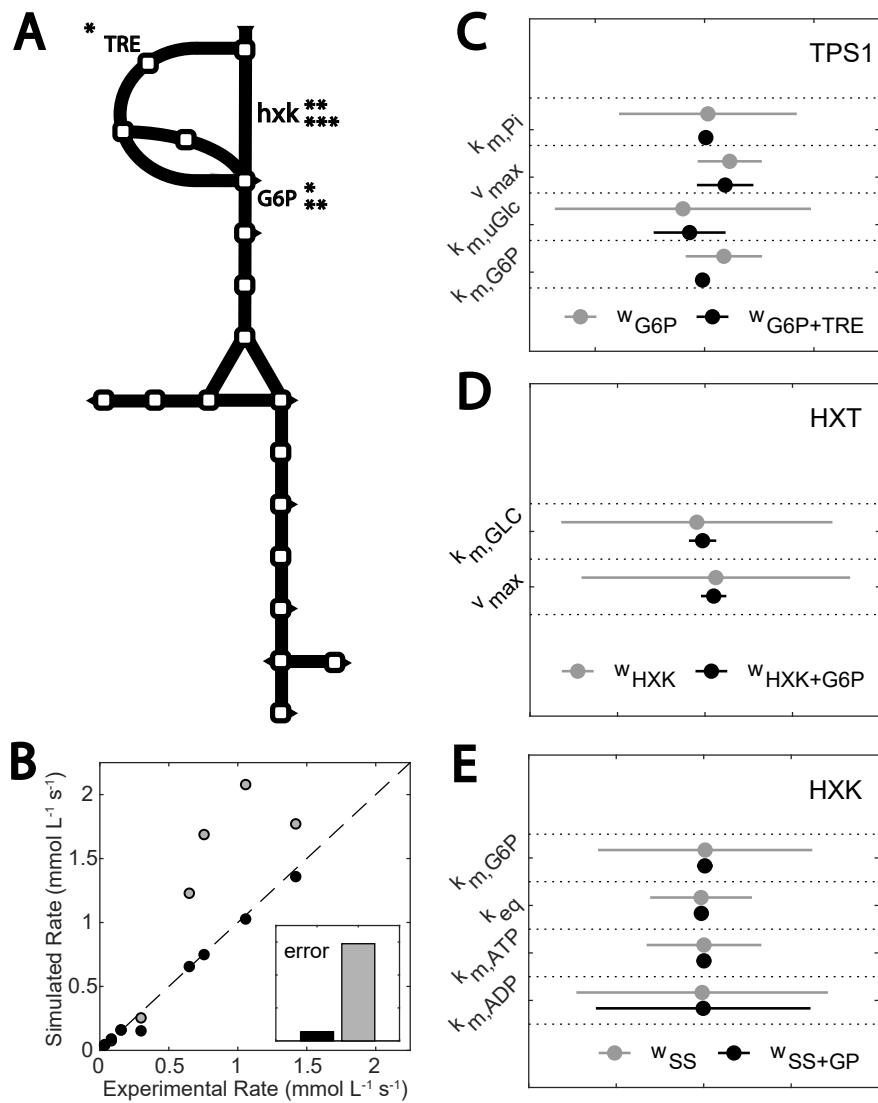
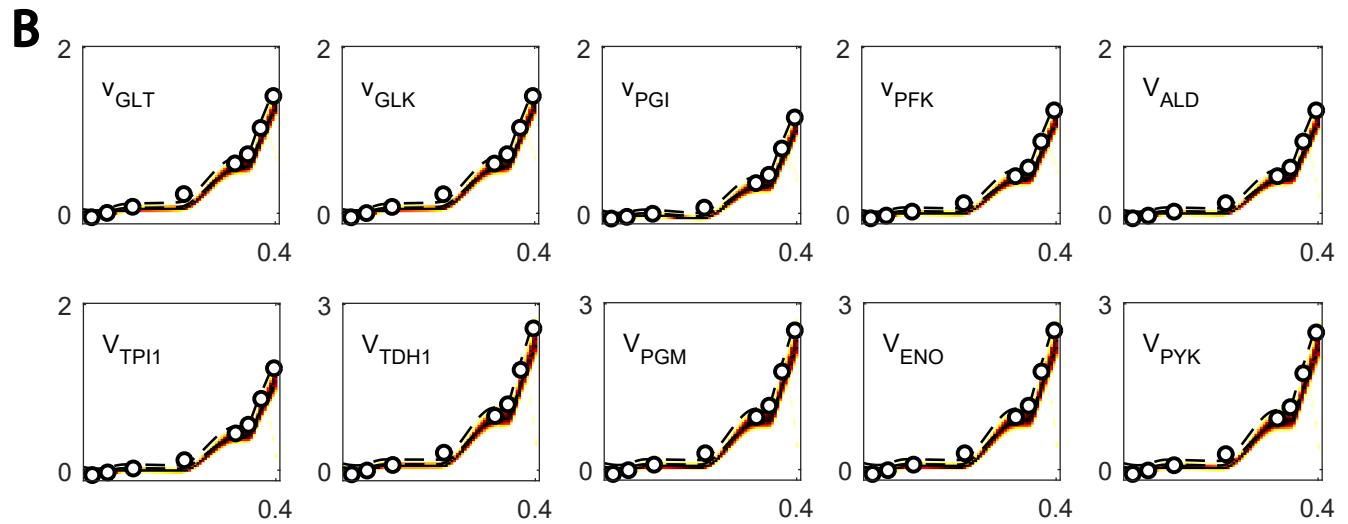
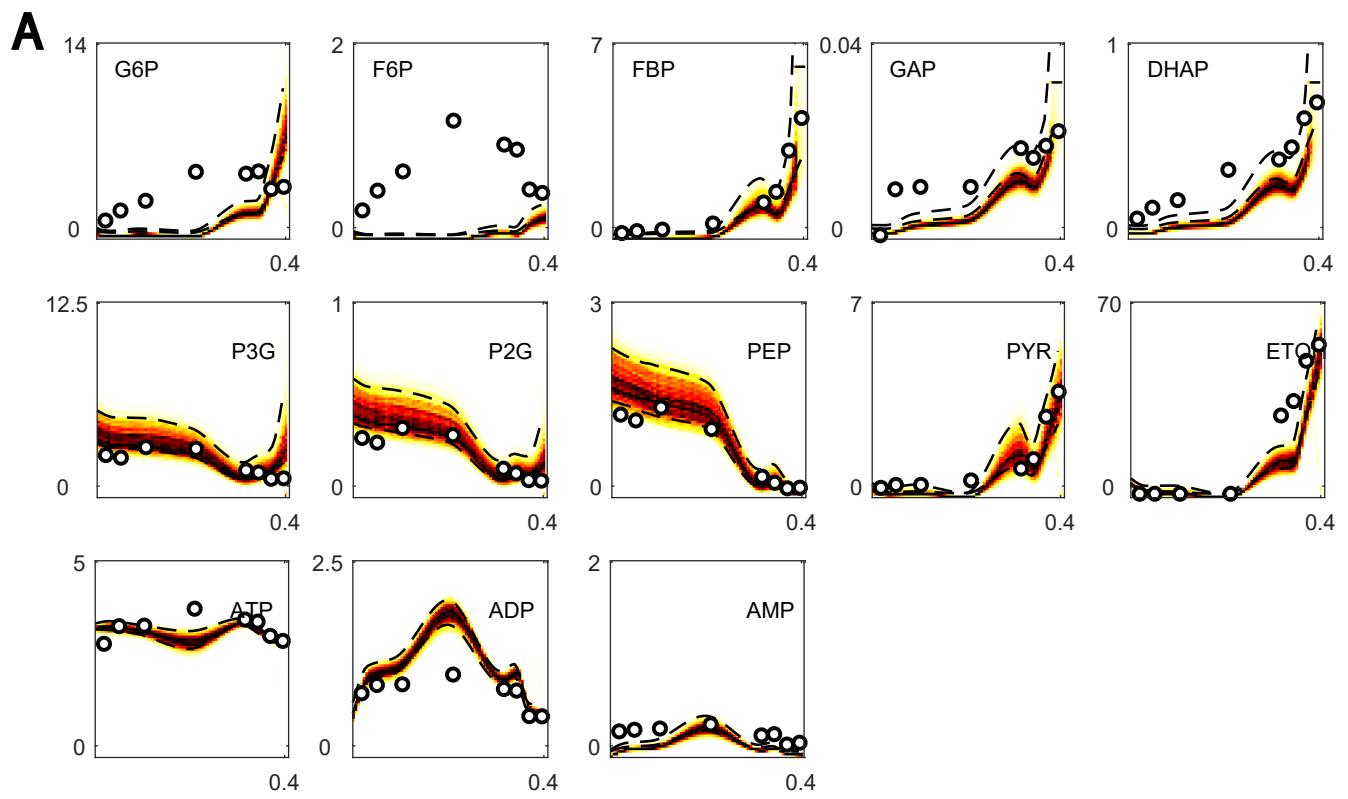
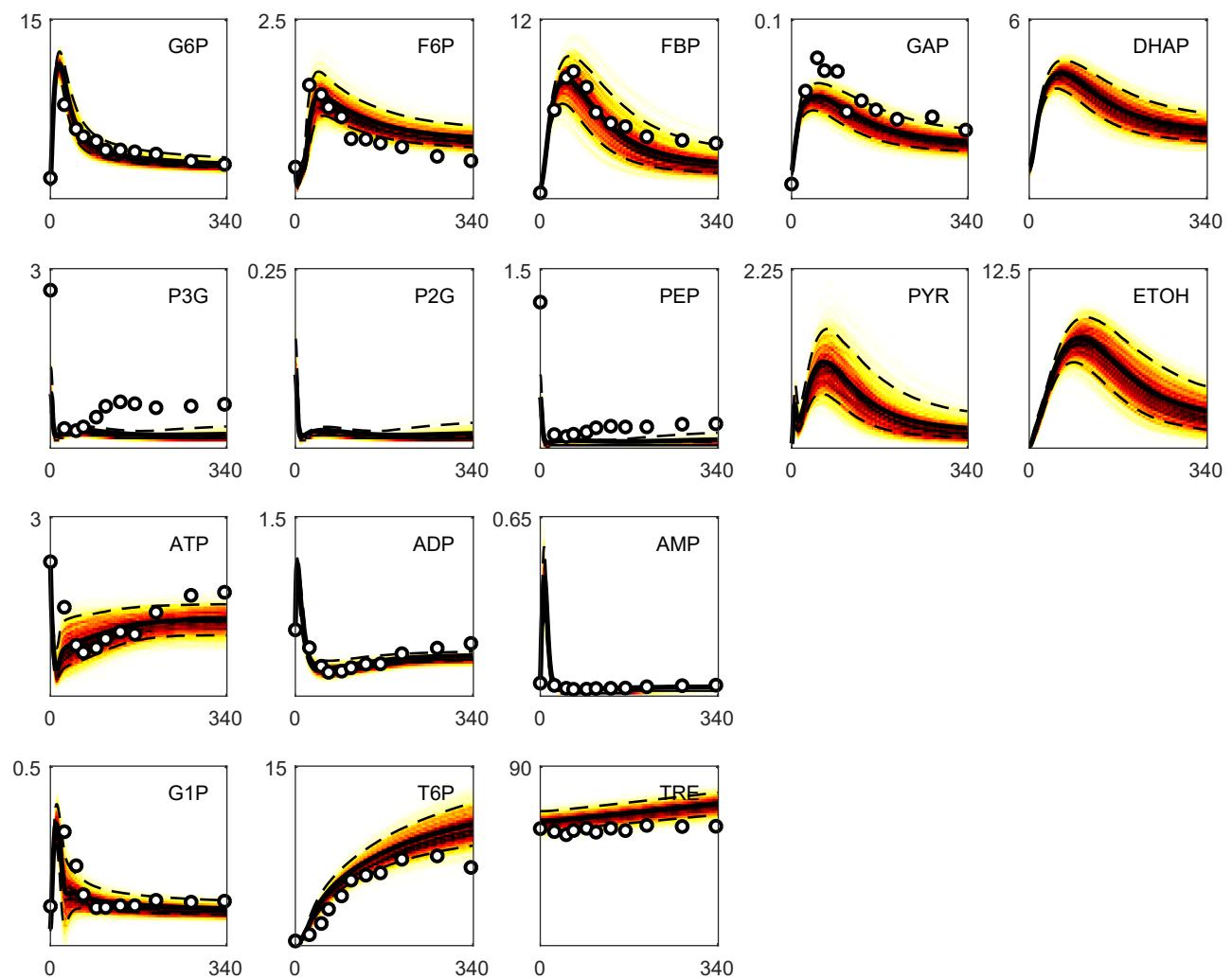


Figure 4. Contribution of different data sets to model identification. (A) Simplified model diagram: variables included in the TPS1 (*), HXT (**) and HXK (***) parameter estimation examples. (B) Simulated HXT reaction rates plotted against experimental rates. Including dilution rate dependency in enzyme concentrations (black) or keeping them constant at the experimental value found for 0.1 h⁻¹. Confidence intervals obtained for the parameter estimates when: (C) G6P or the combination G6P+trehalose was used in TPS1 parameter estimation, (D) reaction rates (HXK) or the combination reaction rate and metabolite concentration (G6P) was used in HXT parameter estimation and (E) SS or the combination SS+GP data was included in HXK parameter estimation.



C

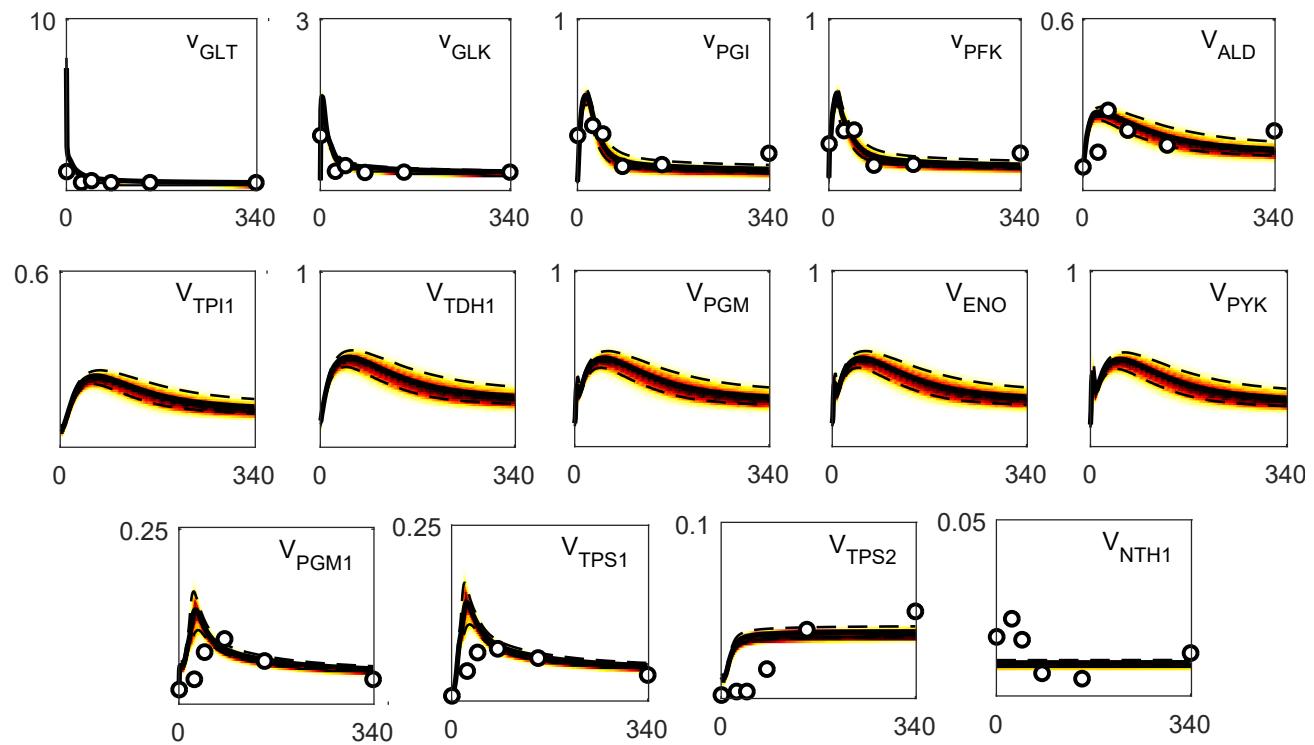
D

Figure 5. The model was robust to parameter heterogeneity. Simulations with the ensemble of models and the experimental data. Random noise was added to generate up to 1000 models in the ensemble. This noise was sampled from a random distribution within 10% of the value of each parameter. The model ensemble is shown as heat maps, where darker regions indicate higher simulation agreement. Experimental data is shown as bullet points. The semi-continuous lines indicate the region where 90% of simulations were found. (A) Steady state concentrations (mM) and dilution rate (h^{-1}) in the Y. and X-axes, respectively. (B) Steady state reaction rates (mM s^{-1}) and dilution rate (h^{-1}) in the Y. and X-axes, respectively. (C) Glucose perturbation concentrations (mM) over time (s) in the Y. and X-axes, respectively. (D) Glucose perturbation reaction rates (mM s^{-1}) over time (s) in the Y. and X-axes, respectively. To see the simulations of each metabolite and reaction rate in the model, see supplementary materials.

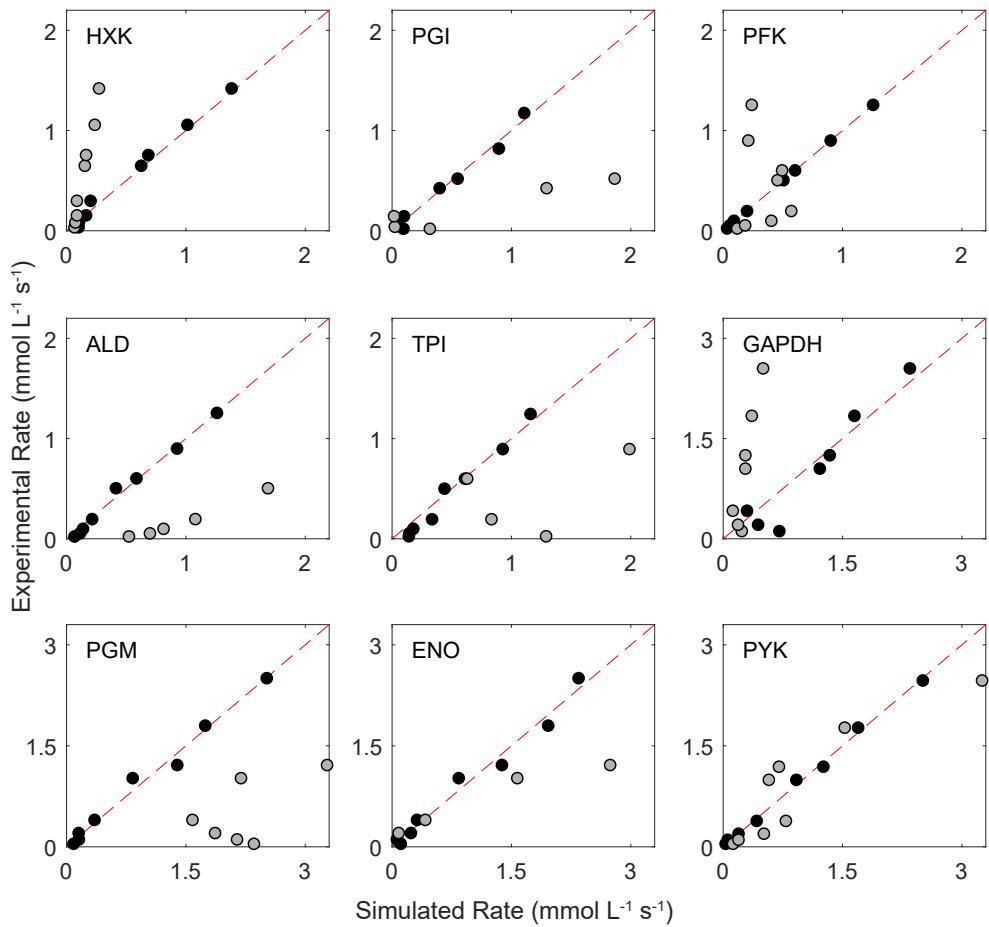


Figure 6. Problem decomposition reveals a mismatch between literature and *in vivo* parameters.

Experimental reaction rates plotted against simulated reaction rates. Gray dots show the model fit using literature parameters. Black dots show the fit with the developed models. The red dashed line shows the location of a perfect fit. Each plot shows a different enzyme in glycolysis. Upper and lower glycolysis in top and bottom row, respectively.

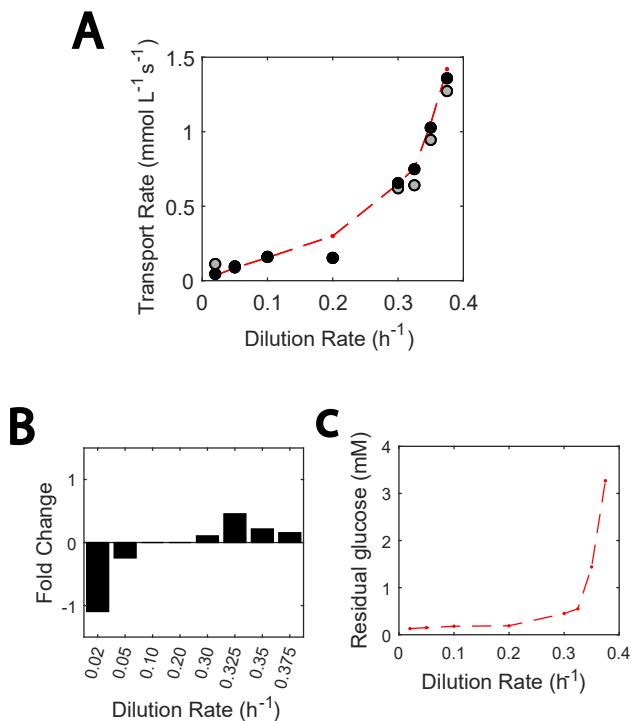


Figure 7. Adjustments in the HXT V_{max} can describe most growth rate-dependent transporter kinetics. (A) Reaction rates plotted against dilution rates. Gray markers show the model fit if a single set of kinetic constants for the hexose transporter is used. Black markers highlight how this fit is improved if the maximum reaction rate is fit for every growth rate. Experimental data point are shown in red markers, united by the semi-continuous red line. (B) Residual glucose concentration (mM) plotted against dilution rate, and (C) Fold change required for V_{max} at each dilution rate.

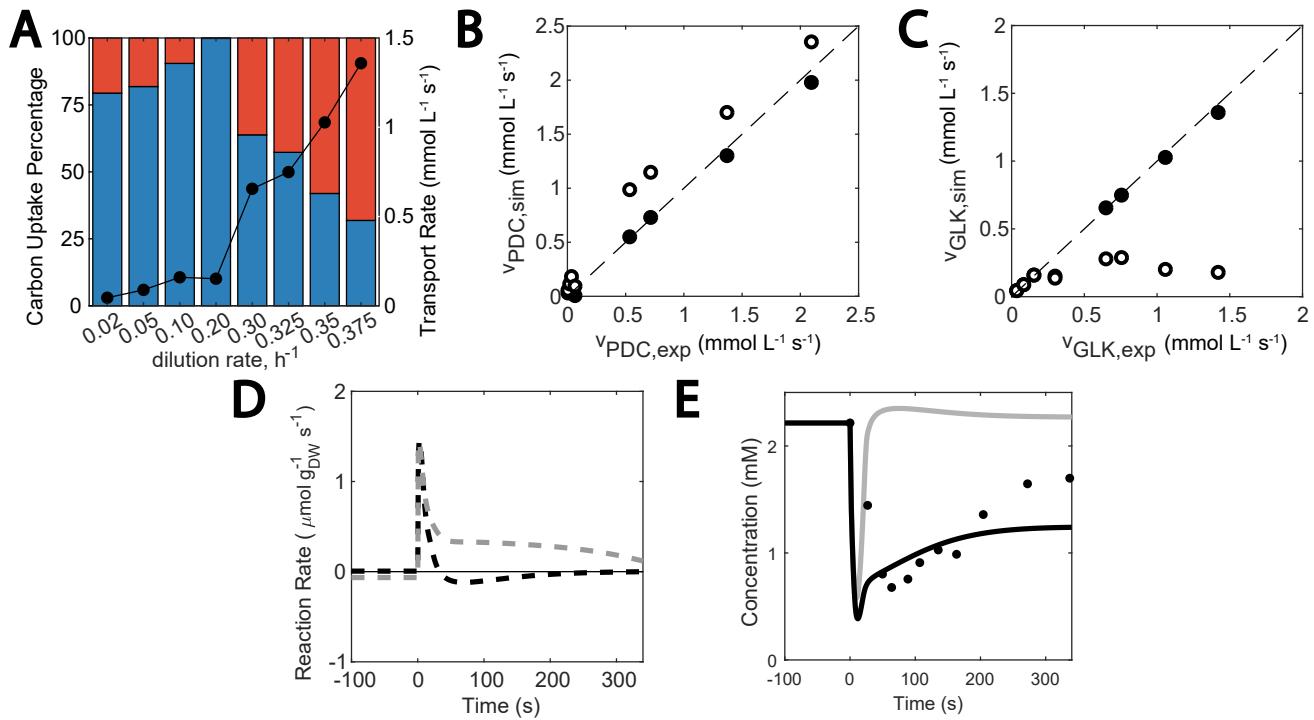


Figure 8. Contribution of different modules to glycolytic regulation. (A) Carbon flux distribution at different growth rates: flux towards sink reactions (blue), and towards ethanol and glycerol synthesis (red). The black dots show the simulated glucose uptake rate. (B) Effect of knocking out mitochondrial flux: Simulated PDC reaction rate plotted against experimental PDC reaction rate. Reference model (black) and effect of knockout carbon flux to respiration (white). (C) Effect of no oxygen: Simulated HXK reaction rate plotted against experimental HXK reaction rate. Reference model (black) and effect of knocking out NADH mitochondrial respiration (white). (D) Imbalance between upper and lower glycolysis (calculated as $v_{\text{GLK}} - 2*v_{\text{GAPDH}}$) occurring if the trehalose cycle is knocked out. The black and gray dashed lines show the *WT* and Trehalose cycle-knocked out strain, respectively. (E) Sensitivity of the ATP paradox to inosine salvage pathway. When the inosine salvage pathway is active (black), a decay in ATP concentrations can be observed, while knocking the cycle (gray) leads to a conservation of ATP steady state concentration.

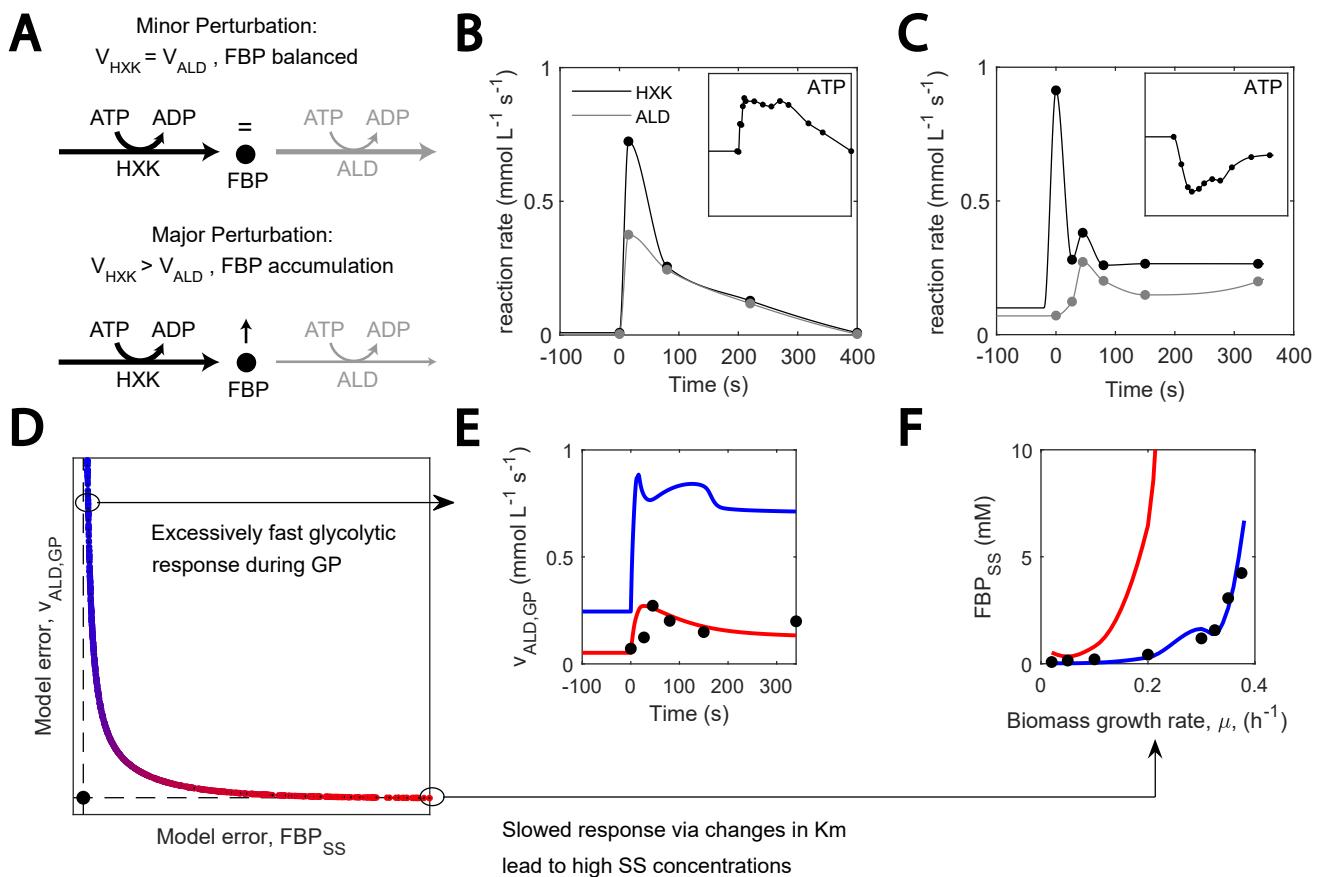
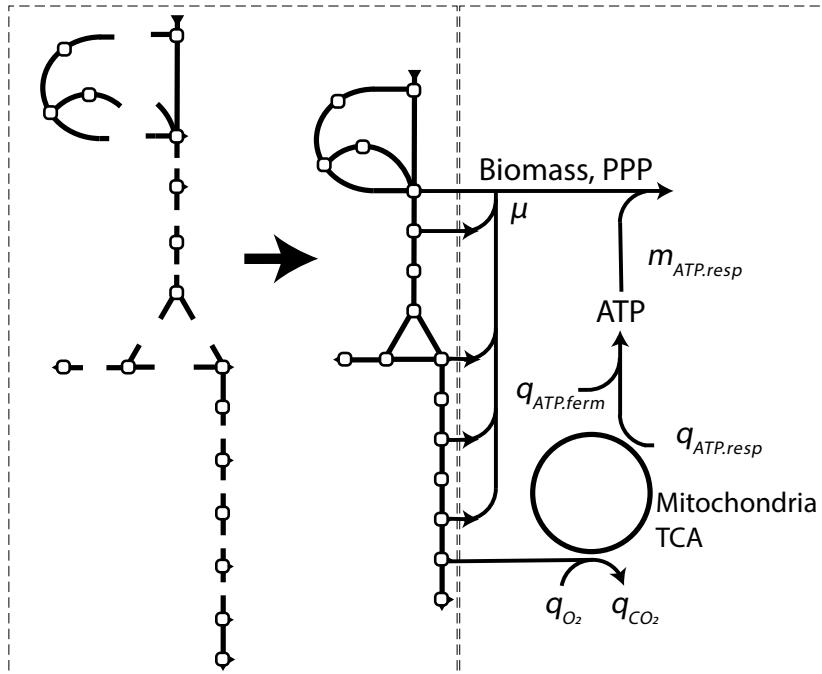


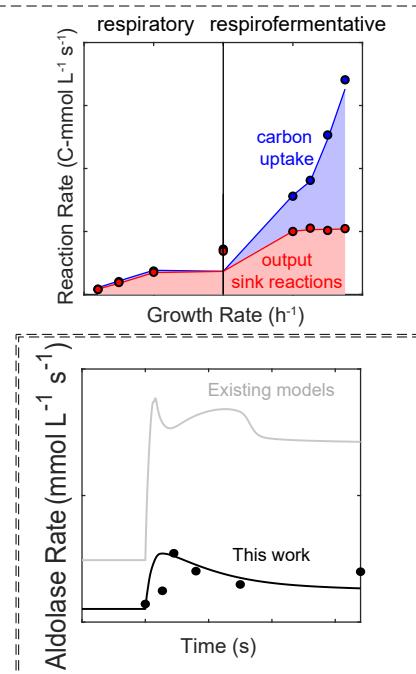
Figure 9. A transient imbalance between ALB and FBP can be explained by missing regulation.

(A) A large glucose perturbation triggers FBP accumulation, (B-C) HXK and ALD reaction rate and ATP concentration upon glucose perturbation under feast famine regime, 0.4 mM (Suárez-Mendez et al., 2014), and single, 110 mM glucose perturbation (van Heerden et al., 2014), (D) Pareto front observed when fitting ALD reaction rate during GP and FBP SS concentrations. Red and blue colors show the zone with lower model error for ALD and FBP, respectively. (E-F) Model fits for ALD reaction rate during GP and FBP SS concentrations, respectively. Colors correspond with (D).

Model decomposition



Physiological information



**Slow response
Upper glycolysis**

Graphical abstract