

Research article

Open Access

A method for estimation of elasticities in metabolic networks using steady state and dynamic metabolomics data and linlog kinetics

I Emrah Nikerel*, Wouter A van Winden, Walter M van Gulik and Joseph J Heijnen

Address: Department of Biotechnology, TU Delft, Julianalaan 67, 2628 BC Delft, The Netherlands

Email: I Emrah Nikerel* - i.e.nikerel@tnw.tudelft.nl; Wouter A van Winden - w.a.vanwinden@tnw.tudelft.nl; Walter M van Gulik - w.m.vangulik@tnw.tudelft.nl; Joseph J Heijnen - j.j.heijnen@tnw.tudelft.nl

* Corresponding author

Published: 21 December 2006

Received: 20 July 2006

BMC Bioinformatics 2006, 7:540 doi:10.1186/1471-2105-7-540

Accepted: 21 December 2006

This article is available from: <http://www.biomedcentral.com/1471-2105/7/540>

© 2006 Nikerel et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Dynamic modeling of metabolic reaction networks under *in vivo* conditions is a crucial step in order to obtain a better understanding of the (dis)functioning of living cells. So far dynamic metabolic models generally have been based on mechanistic rate equations which often contain so many parameters that their identifiability from experimental data forms a serious problem. Recently, approximative rate equations, based on the linear logarithmic (linlog) format have been proposed as a suitable alternative with fewer parameters.

Results: In this paper we present a method for estimation of the kinetic model parameters, which are equal to the elasticities defined in Metabolic Control Analysis, from metabolite data obtained from dynamic as well as steady state perturbations, using the linlog kinetic format. Additionally, we address the question of parameter identifiability from dynamic perturbation data in the presence of noise. The method is illustrated using metabolite data generated with a dynamic model of the glycolytic pathway of *Saccharomyces cerevisiae* based on mechanistic rate equations. Elasticities are estimated from the generated data, which define the complete linlog kinetic model of the glycolysis. The effect of data noise on the accuracy of the estimated elasticities is presented. Finally, identifiable subset of parameters is determined using information on the standard deviations of the estimated elasticities through Monte Carlo (MC) simulations.

Conclusion: The parameter estimation within the linlog kinetic framework as presented here allows the determination of the elasticities directly from experimental data from typical dynamic and/or steady state experiments. These elasticities allow the reconstruction of the full kinetic model of *Saccharomyces cerevisiae*, and the determination of the control coefficients. MC simulations revealed that certain elasticities are potentially unidentifiable from dynamic data only. Addition of steady state perturbation of enzyme activities solved this problem.

Background

Metabolic activities of living cells are accomplished by a well regulated, highly coupled network of numerous

enzyme catalyzed reactions and selective membrane transport systems. To engineer such systems, enzymatic, transport and regulatory functions of the cells are manip-

ulated via the use of recombinant DNA technology [1]. Within the purpose of metabolic engineering, i.e. rational redesign of the metabolic systems, the highly relevant question of which (combination of) perturbation should be applied in order to increase the productivity of the microorganism is addressed. The answer to this question requires information on both the regulatory level and the metabolic level. Ter Kuile and Westerhoff showed by their Hierarchical Control Analysis that the pathway flux is rarely controlled solely by gene expression, but that metabolite levels are also relevant [2].

In this context, we focus on the metabolic reaction network level. Modeling metabolic reaction systems is usually based on stoichiometric sometimes followed by kinetic modeling.

One of the initial steps in the modeling of metabolic reaction networks is to determine the structure and steady state characteristics of a given network using stoichiometric information alone. Steady state models describe time invariant fluxes, gathered from steady state experiments; hence they reflect the structural characteristics of the system. Metabolic Flux Analysis (MFA) and Metabolic Network Analysis (MNA) were developed as powerful tools to analyze such flux data. At steady state, the mass balances over the metabolites in the metabolic network yield a set of linear relations between the metabolic fluxes which can be expressed as:

$$\mathbf{S} \cdot \mathbf{v} = \mathbf{0} \quad (1)$$

Where \mathbf{S} is the ($m \times r$) stoichiometric matrix and \mathbf{v} is the ($r \times 1$) vector of metabolic fluxes, where m is the number of balanced metabolites and r is the number of fluxes. Here, the system is in most of the cases underdetermined so that there are an infinite number of possible solutions. The realized solution depends on the kinetic properties of the involved reactions; this information is seldom known. To bypass the need of information on kinetics of individual reactions, alternative mathematical approaches have been proposed in the past to obtain a unique solution. An example is the constraint based optimization approach which is based on assumed optimality criteria, e.g. maximum growth, given biochemical, thermodynamic and irreversibility constraints and maximal reaction rates [3-5]. Later, Segre *et al.* proposed the optimality constraint that requires maximization of biomass formation while minimization of metabolic adjustment (MOMA) in order to obtain a unique flux distribution of a mutant strain. In their approach, they defend their optimality criterion that a knock-out mutant strain would optimize its biomass production rate by changing minimally its metabolic fluxes from the wild type strain [6]. An alternative metabolic modeling framework which uses a fitness function is

the cybernetic approach. This approach assumes that an organism is an optimal strategist in utilizing all available sources with maximum efficiency. The expression and activity of the enzymes that catalyze network functionality are regulated by cybernetic control variables obtained from the solution of a constrained optimization problem [7-9].

Despite a number of successful applications especially in mixed substrate and prediction of knock-out lethality, all stoichiometric modeling approaches have their limitations, e.g. they can not predict time courses of the cellular processes, are based on "assumed goals" of the cell, and do not give insight in molecular events occurring in the cells, since information about the kinetic properties of the individual enzymes are not required. In the light of the above arguments, it is apparent that, to advance our understanding of the (dis)functioning of living cells a systems biology approach is needed, whereby the use of dynamic mathematical models of metabolic reaction networks to describe the complex kinetic behavior and interactions (allosterical, feedback and feed forward effects, cofactor coupling, compartmentation, intracellular transport, etc.) is becoming increasingly relevant.

The kinetic behavior of many important enzymes occurring in metabolic networks have been studied extensively, however these studies have generally been performed under non-physiological conditions in test tubes (*in vitro*), and therefore the applicability of these results to the *in vivo* metabolism is doubtful [10-13]. Teusink *et al.* showed that discrepancies exist between the *in vivo* measured changes of the concentrations of the glycolytic metabolites and their estimates using models based on mechanistic rate equations and *in vitro* determined parameters [12]. This basic problem invalidates detailed models of metabolism containing kinetic parameters which have been determined *in vitro*. Therefore, it is preferred to base the kinetic analysis of metabolic networks on *in vivo* studies of intact cellular networks. These *in vivo* studies are based on steady state and/or dynamic perturbations of a metabolic network starting from a reference steady state that is defined by its fluxes, enzyme activities, metabolite levels, extracellular concentrations, and kinetic parameters.

It is also important to notice that the number of parameters that is typically involved in the traditional mechanistic equations is very large which causes identifiability problems due to parameter insensitivity. Despite the information richness of data obtained from dynamic perturbations in *in vivo* experiments, there is a limit for identification of the parameters. For these reasons, it seems justified not to limit ourselves to the available complex mechanistic enzyme kinetics which must also be consid-

ered as approximations of the true *in vivo* behavior [14,15]. Approximative enzyme kinetic formats, which contain much fewer kinetic parameters are therefore of general interest. An overview of different approximative kinetic formats (linear, power law, loglin and linlog kinetics) used in metabolic network modeling is given by Heijnen [16].

One of the proposed formats is, linlog kinetics, which has been introduced for modeling of *in vivo* kinetics and for metabolic redesign, and shown to have a good approximation quality, standardized format and relatively few parameters [16,17]. In linlog kinetics, all the rate equations have the same mathematical structure in which the relation between rates and enzyme levels is proportional, while for metabolite levels, a linear sum of logarithmic concentration terms is proposed. All variables are considered relative to a reference steady state (Eq. (2)). The linlog approximation is valid in the neighborhood of the reference state (defined by J^0 , x^0 and c^0 in equation (2)), but quite large changes of metabolite concentrations, enzyme levels and fluxes are allowed [17]. The parameters (ε_x^v and ε_c^v) in the kinetic equations are the same scaled

elasticities $\left(\varepsilon_x^v = \frac{x^0}{J^0} \frac{\partial v}{\partial x} \Big|_{x=x^0} \right)$ that are used in Metabolic

Control Analysis (MCA). It is important to note that the elasticity parameters appear in the model in a linear fashion.

$$\frac{v}{J^0} = \frac{e}{e^0} \left(1 + \varepsilon_x^v \ln \left(\frac{x}{x^0} \right) + \varepsilon_c^v \ln \left(\frac{c}{c^0} \right) \right) \quad (2)$$

When the elasticities are known, a full dynamic model of the whole metabolic network can be set up using linlog kinetics. Such a model allows in principle the calculation of control coefficients also under dynamic conditions. In linlog kinetics, the elasticities are the kinetic parameters represented in the elasticity matrix. From these, and a given network structure specified in the form of a stoichiometric matrix, the control coefficients for a reference condition can be calculated from the summation and connectivity relations developed in the framework of MCA. Also the change in control coefficients upon large changes in enzyme levels can be calculated [18]. Moreover, the linlog formulation enables the analytical solution of the mass balances for steady state metabolite and flux levels in the metabolic network, providing the solution of the metabolic redesign problem, i.e. determination of the optimal enzyme levels that maximize a certain flux while the total amount of enzyme and the metabolite levels are con-

strained. Visser *et al.*, reported a successful application of linlog kinetics in an *in silico* study that aims to solve this metabolic redesign problem [19].

In order to determine the kinetic parameters of a model of a given *in vivo* metabolic system, *in vivo* perturbations of the complete metabolic network have to be performed. There are two main types of perturbations that can be imposed on the system: steady state and dynamic perturbations. In steady state perturbations, usually the enzyme activity of one (sometimes more) of the reactions is changed by adding specific inhibitors or activators or by genetically changing the enzyme activity, resulting in a new steady state. In steady state perturbations an important problem has been addressed by Kacser and Burns, which is the determination of the set of reactions that has to be perturbed in steady state fashion, in order to be able to determine all elasticities for a given metabolic network which resulted in their "double modulation" [20]. They showed that for a simple linear chain of reaction, perturbation of the activities of the first and last enzymes allows determination of the elasticities of all enzymes of the chain under the condition that each enzyme is only responsive to its substrate and product, which rules out the fact that feedback loops are present. The theoretical basis is presented in later studies which showed that determination of the elasticities for any enzyme in such a simple chain requires two perturbations, one upstream and one downstream of the enzyme concerned [21]. Giersch and Cornish-Bowden extended the double perturbation approach to more complex pathways containing branch points, regulatory loops, and conserved moieties and they provided guidelines to list the possible reactions to be modulated in order to determine the elasticities for arbitrary metabolic networks [22]. From the obtained list a minimum set of steady state perturbations, to be imposed on a specific network, can be chosen. As an alternative to the analysis of Giersch and Cornish-Bowden, Hofmeyr and Cornish-Bowden offered co-response analysis, to identify the mono-functional units that respond together to any perturbation applied [23]. These mono-functional units have to be dissected in order to determine the elasticities belonging to these groups.

Linlog kinetics has been successfully applied to estimate the elasticity parameters for a linear pathway from sets of steady state metabolite concentrations and enzyme activities [18]. Additionally, Heijnen *et al.* proposed a method to obtain flux control coefficients around a branch point, from large enzyme perturbation leading to large steady state flux perturbations, using the linlog kinetic format [24]. This approach allows obtaining explicit solutions for steady state flux and metabolite levels as a function of large changes in enzyme activities.

An alternative to steady state perturbations are dynamic perturbations, in which the system, being initially at the reference state, is disturbed to create time dependent data of transient metabolite levels. The dynamic metabolite profiles are typically obtained as a series of snapshots in time. A recent example of these are *in vivo* measurements of a number of metabolites in rather dense time sequences of a few seconds or minutes, using 'rapid sampling' methods with subsequent high precision metabolite measurement techniques [25-28]. Transient data are rich in information and allow determination of the time hierarchy of the different elements of the metabolic network and the causal relationships between the network elements. It is possible to exploit these data to estimate the parameters of a traditional full kinetic model [13,29]. Such a model can subsequently be used to calculate the values of the elasticities at a given reference steady state. Alternatively, when the linlog kinetic format is used the elasticities can be directly estimated from the transient data. This was demonstrated by Kresnowati *et al.* who estimated the elasticities of a small example network from transient concentration data, assuming that the dynamic fluxes are unknown [30].

An important difference between steady state and dynamic perturbations is that in the former both the steady state enzyme levels (e_i) and the metabolite concentrations (x_i) have to be measured, whereas in rapid dynamic perturbation experiments only x_i is required because the enzyme activities can be considered constant within a sufficiently short time window following the perturbation. Note that flux data are required in both methods. However, fluxes are not independent variables as they follow from the measured intra- and extracellular metabolite concentrations and the proper mass balances.

In this work, we present a method to estimate kinetic parameters (elasticities), using linlog kinetics, using metabolome data obtained from steady state and dynamic perturbation experiments. As such, this can be considered as an extension of the work by Visser and Heijnen [17], Wu *et al.*, [18] and Kresnowati *et al.*, [30] who have provided the theoretical framework and a small practical application of the linlog kinetic format. We first generalize the notation to make it applicable to networks of arbitrary size and complexity. Additionally, we address further the question of parameter identification from experimental data. We monitor the propagation of error throughout the proposed parameter estimation procedure and we determine the subset of identifiable parameters. To illustrate the proof of principle, we apply the presented theory on *in silico* generated data, with realistic experimental settings, to be able to compare the obtained results with the "known truth". This data is generated using a pre-

viously published glycolytic pathway model [31] of *Saccharomyces cerevisiae*.

Results

In silico pulse experiments

In these experiments, the steady state was perturbed by increasing the extracellular glucose concentration. When the external glucose concentration is increased, the intracellular glucose concentration also increases as the sugar is transported through facilitated diffusion. From Figure 1a it can be seen that the G6P concentration increases around 6% in the first five seconds and then starts to decrease, as the ATP concentration drops due to the phosphorylation of the glucose in the hexokinase reaction. ATP starts to recover after about 15 seconds due to increased ATP production downstream of glycolysis. This is in agreement with the literature on the glucose uptake dynamics following a sudden increase in the external sugar concentration [32]. Following this ATP fluctuation, the PEP concentration also decreases and then recovers. In the first five minutes, the concentrations of the external metabolites increase because of their increased production.

After this initial period of 2 min, in which the external glucose concentration is nearly constant, the concentration of external glucose slowly decreases due to continuous consumption and wash-out from the reactor. The cells follow this glucose drop by dropping the levels of intracellular metabolites. For the extracellular products (polysaccharides, glycerol, ethanol) the washout from the reactor is larger than their production and hence their concentration also drops. After about 200 minutes, the original steady state is restored (Figure 1b). Such a pulse experiment therefore always delivers a highly dynamic dataset, followed by a pseudo steady state dataset.

Parameter estimation

The dynamic data obtained during the first five minutes of the glucose pulse experiment (Figure 1a) were used to estimate the elasticities via the linlog parameter estimation procedure outlined in the Methods section (see section Determining elasticities from dynamic perturbation data). In Figure 1a, the simulation of the same dynamic perturbation using the estimated elasticities is also given. In Figure 2, the estimated elasticities are compared with the theoretically calculated elasticities, derived from mechanistic rate equations at the reference state.

Although the linlog simulation results do not differ much from the noise-free experimental data generated with the mechanistic model, some estimated elasticities are far from the expected theoretical values. The difference is in some cases in magnitude, for other cases there is a sign contradiction. Specifically, 6 out of 16 elasticities (E3, E5,

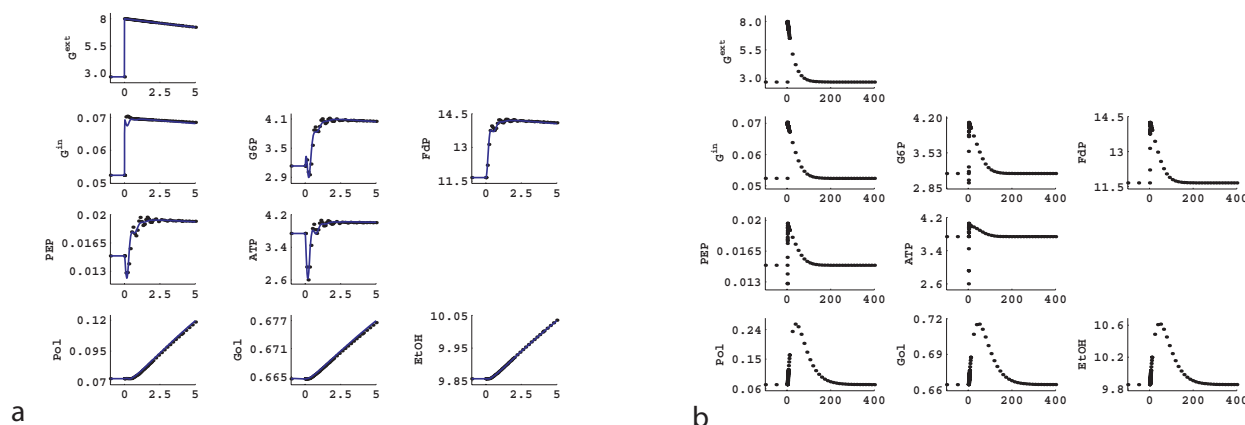


Figure 1

Response of the network to an increase in external glucose concentration. a) First five minutes, a "rapid sampling" experiment (black) and the simulation of the same perturbation with linlog kinetics (blue) using the elasticities estimated from the experimental data from dynamic perturbation. b) Long term response of the cells to the glucose pulse. The time is presented in minutes; intracellular metabolites are given in $\mu\text{mol gDW}^{-1}$; extracellular metabolites are given in mM.

E6, E7, E8, E16) are predicted in very good agreement with the theoretical ones (difference being around 10%), 6 estimated elasticities (E2, E9, E10, E12, E14, E15) agree in sign, but the magnitudes differ by more than 50% (in some cases up to 200%), and 4 elasticities (E1, E4, E11, E13) are estimated with a sign contradiction. Notice that in the case of E12, the 200% deviation is already expected. This elasticity explains the effect of G6P on the polysaccharide formation rate and in the mechanistic model this effect is explained with a Hill type kinetic equation with a Hill coefficient of 8.25, hence this rate is very sensitive to the changes in G6P. To mimic this behavior, E12 is estimated to be much higher than its theoretical value.

The poor identification of the elasticities has two possible reasons: either the experimental design is poor or there is a structural problem resulting that some interactions in the network can not be resolved from the available information. From an experimentalist point of view, it can be seen from Figure 1a that the changes in the intracellular metabolites during the first 300 seconds are between 25–30% which can be easily detected with the current measurement techniques [33]. The extracellular glucose and polysaccharides also change to a detectable extent, but the changes in the ethanol (and therefore glycerol) are only 2–3% which is hard to detect. To check if the poor experimental design causes problems in the parameter identification, we have altered the experimental design to create larger changes in these concentrations. First the biomass concentration was increased to 15 gDW L^{-1} . Also at the same moment that the fermentor was pulsed with the glucose, the inflow and outflow of the fermentor was stopped so the operation was effectively switched from though-

flow mode to batch mode, and the extracellular metabolites were not washed out anymore. This allowed more rapid accumulation of the secreted products resulting in much larger changes. Also, the change in the extracellular glucose concentration was more pronounced because there was no further addition of glucose after the pulse. These changes did not result in considerable changes in the intracellular metabolite profiles; they follow the slight change in the external glucose profile. The results of the new experimental design are depicted in Figure 3 which represents the new experimental data where only the significantly changed external metabolite profiles are depicted, together with the linlog simulation using the estimated elasticities. However, no significant improvement was achieved in the identification of the elasticities compared to the initial experimental design (data not shown). Hence it can be concluded that the problem is not due to poor experimental design.

A closer look at the estimated elasticities reveals that the elasticities belonging to the part of the metabolic network, between V_{HK} and V_{GAPD} are correctly estimated, whereas the estimation of the elasticities belonging to the lower part of glycolysis and the uptake reaction is poor. This is mainly due to the fact that the information content of the metabolite data is insufficient to resolve the complex interactions of metabolites and enzymes. Specifically, in the V_{PK} reaction rate, the feed forward activation of FdP and the mass action effect of PEP are assumed. The perturbation of the system via an increase in the external glucose concentration results in an increase in the V_{PK} flux. However, this does not allow a separate determination as to what extent the change in V_{PK} is due to mass action effect

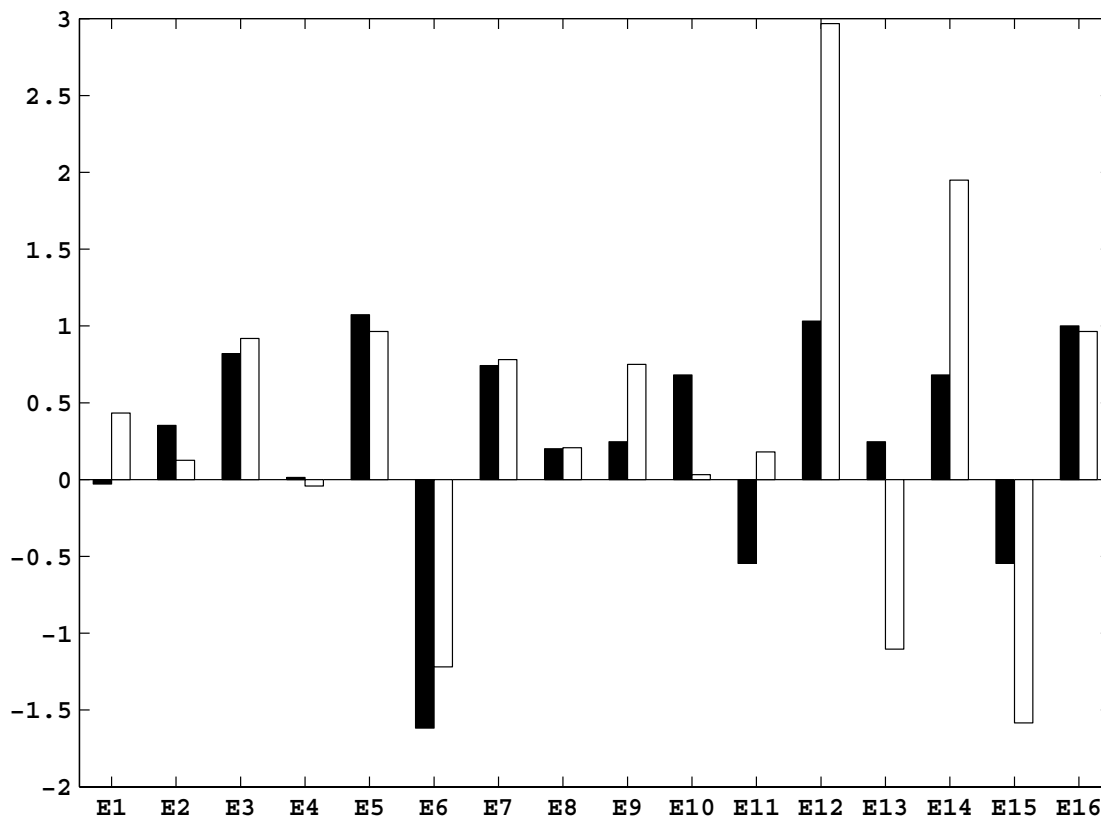


Figure 2

Results of the estimation of the elasticities using dynamic data only, comparison of the theoretical elasticities (black), with the linlog elasticities (white).

of PEP or due to feed forward effect of FdP. This results in identifiability problems for the elasticity parameters that describe the ethanol production (E9, E10, and E11). The same argument is valid for the elasticities belonging to the glycerol producing reaction, since in the *in silico* network the glycerol production is assumed to be proportional with the ethanol production; hence the same combined effect is seen in the glycerol production, resulting in an identifiability problem for the latter triplet of elasticities (E13, E14, E15)

In order to analyze the information content of the data, the Fisher Information Matrix (FIM) is calculated as described in [30]. FIM is an indicator of the information content with respect to the parameters and due to the parameter linearity in linlog kinetics FIM is independent of the values of the kinetic parameters. It can be calculated as $Y^T Y$, Y being the design matrix appearing in equation

(9). The singular values of the FIM hold information on the number of linear dependencies between the columns of the data matrix. In our system, we checked the singular values of FIM and concluded that not all the singular values are equally significant, which shows that there are some colinearities within the data matrix. This is also evident from the condition number, which equals the ratio of the highest to the lowest eigenvalues of the matrix and is ideally 1 in a completely uncorrelated case. In this case the condition number was calculated to be 5.0×10^6 . This fact supports the previous argument that using only one single set of dynamic data is not sufficient to resolve all the elasticity values.

In order to obtain better estimates of the elasticities, steady state perturbation data are introduced. The steady state enzyme perturbation data are generated as described in the Methods section (see section Steady state perturba-

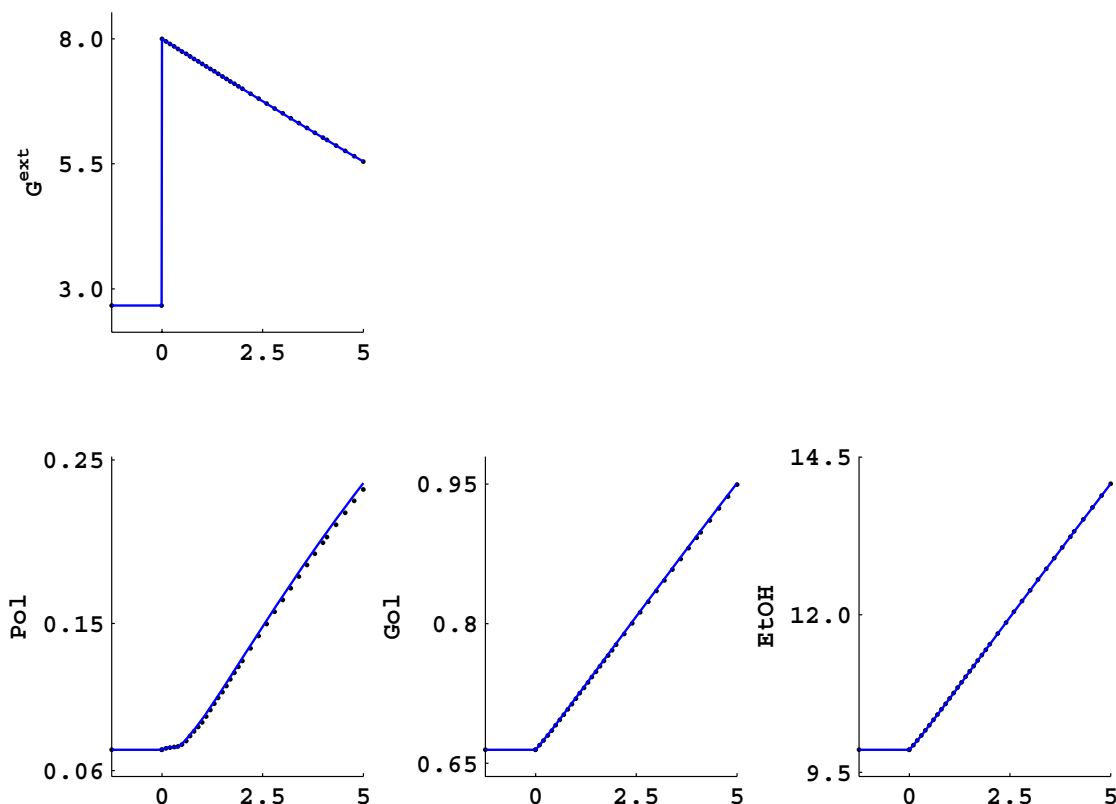


Figure 3

The results of the new experimental design (The inlet and outlet of the reactor is blocked just after the glucose is increased. The biomass concentration is 15 gDW L^{-1} , see text for further details in the experimental design) (black) and simulation of the same perturbation with linlog kinetics (blue).

tions): in order to resolve the feed forward effect of FdP on the V_{PK} reaction, the enzyme amount of the V_{GAPD} reaction is perturbed in addition to the V_{IN} and V_{ATPase} . Using the steady state perturbation data in Table 1, and the calculation procedure described in the Methods section (see section Determining elasticities from steady state perturbation data), we have estimated the elasticities for the *in silico* network under study. After the estimation of the elasticities using steady state data only, we have also simulated the dynamic perturbation of Figure 1a using linlog kinetics, using the estimated elasticities. The results are presented in Figure 4. Figure 4a shows the experimental data and the linlog simulation of the same pulse, and Figure 4b compares the elasticities estimated using the steady state data with the theoretical elasticities. As can be seen from Figure 4b, using the steady state data, the elasticities are estimated in good agreement with the theoret-

ical elasticities. However, as can be observed from Figure 4a, the linlog model deviates significantly from the experimental data, so the estimated elasticities have to be further refined.

Having a good initial estimate for the linlog elasticities from the steady state perturbation data only, we used in the second step both the dynamic and steady state data for the parameter estimation procedure. The nonlinear regression procedure allows combining both available steady state and dynamic experimental data, so that all parameters can be accurately estimated. After the non-linear fit to the data, we have obtained the final estimation of the elasticities. The simulation of the pulse experiment with these final estimates, and the comparison of the elasticities are given in Figure 5a and Figure 5b respectively,

Table 1: Relative effects of steady state perturbations on metabolite concentrations and rates. The metabolites and the rates are given relative to their reference states

	Perturbation Applied		
	$e_{IN}/e_{IN}^0 = 1.2$	$e_{ATPase}/e_{ATPase}^0 = 0.9$	$e_{GAPD}/e_{GAPD}^0 = 0.8$
Normalized metabolites			
G ^{ext'}	0.793	1	1
G ^{in'}	1.128	0.999	1
G6P [']	1.15	1.103	1
FdP [']	1.093	0.959	1.369
PEP [']	1.131	1.055	0.908
ATP [']	1.045	1.071	1
Pol [']	3.053	2.208	1
Gol [']	1.077	0.983	1
EtOH [']	1.077	0.983	1
Normalized rates			
V _{IN} [']	1.104	1	1
V _{HK} [']	1.104	1	1
V _{PFK} [']	1.077	0.983	1
V _{GAPD} [']	1.077	0.983	1
V _{PK} [']	1.077	0.983	1
V _{Pol} [']	3.053	2.208	1
V _{Gol} [']	1.077	0.983	1
V _{ATPase} [']	1.045	0.964	1

which shows that most of the elasticities are correctly estimated and that the experimental data are well described.

Analysis of parameter identifiability and model reduction

In the previous sections, we have seen that we could not determine all the parameters from dynamic perturbation data only and we have introduced the steady state perturbation data. After determination of all of the elasticities, using both types of perturbation data, we analyzed back the identifiability of each parameter under noise. In order to carry the identifiability and error propagation analysis, MC simulation was used. As described in the Methods section (see section Error propagation analysis), 10% relative error was added to the noise free data represented in Figure 1a (rapid sampling experiment), and the non-linear estimation procedure was implemented using the elasticities estimated from the steady state data, as the initial guesses. After repeating this scheme 50 times, we obtained a distribution for each of the elasticities from which we calculated the relative error for each elasticity as the standard deviation per mean of the corresponding elasticity (Table 2).

The adjustable parameter λ in equation (6) determines the relative importance of the different sources of data. In our case, these are dynamic and steady state data. Notice

that during the MC simulations, noise is added only to dynamic data. For steady state perturbations, the number of data points that can be obtained is theoretically infinite. This implies that, considering the central limit theorem, the steady state data can, in principle, be considered as noise free. However, the aim of the MC simulations was to elucidate the potentially unidentifiable elasticities using dynamic perturbation data, and therefore we chose λ to be equal to zero. The other extreme ($\lambda = 1$) would suppress the effect of noise and wouldn't lead to detection of poorly identifiable elasticities. The values in between, are up to the choice of the modeler, depending on how many data points have been obtained from a dynamic pulse experiment, the standard deviations of the measurements, etc. We have implemented different values for the value of λ , but the outcomes of the MC simulation, i.e. which elasticities can hardly be identified, did not change qualitatively.

From the distribution of each elasticity, the p-value, indicating the probability that the actual elasticity is zero and that the estimated value was caused by the random error only, is calculated via t-test and the results are given in Table 2 (second column). The elasticities E1, E4, E9, and E15 have a relative standard deviation higher than 200%, so these elasticities are not identifiable under the 10%

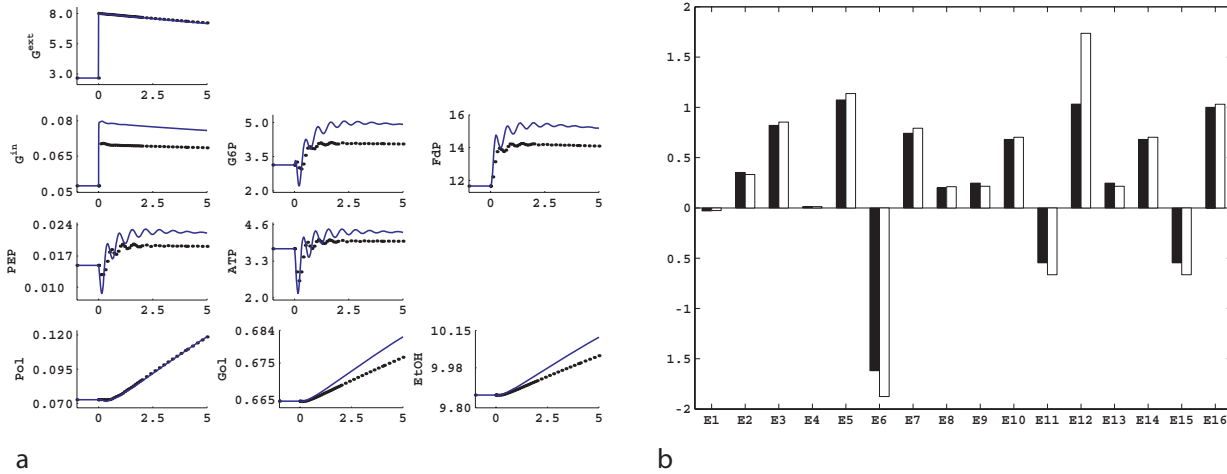


Figure 4 Linlog model simulation of the dynamic perturbation of Figure 1a using elasticities obtained from steady state data only. a) Experimental data (black) and the linlog simulation (blue). Units are the same as Figure 1. b) Comparison of the theoretical elasticities (black), with the linlog elasticities (white) estimated from steady state perturbation data.

noise present. Therefore we conclude that these elasticities are clearly candidates for model reduction by assigning them to be zero ($E1, E4, E9, E15 = 0$).

A closer look reveals that E9 represents the feed-forward effect of FdP on V_{PK} . E4 and E15 represent the effect of ATP on V_{HK} and V_{GoI} respectively and E1 is the feedback inhibition of G6P on V_{IN} . Notice that the elasticities of the reactions of V_{IN} , V_{PK} and V_{GoI} were already badly identi-

fied, when we estimated the elasticities using dynamic data only (Figure 2).

Using the arguments stated above, we set the values of these four ($E1, E4, E9, E15$) elasticities to zero, implemented the non-linear regression step, with the new sparser elasticity matrix, and estimated the elasticities using steady state and dynamic data. The results are presented in Figure 6a, where the experimental data and the linlog simulation using the new set of elasticities are given

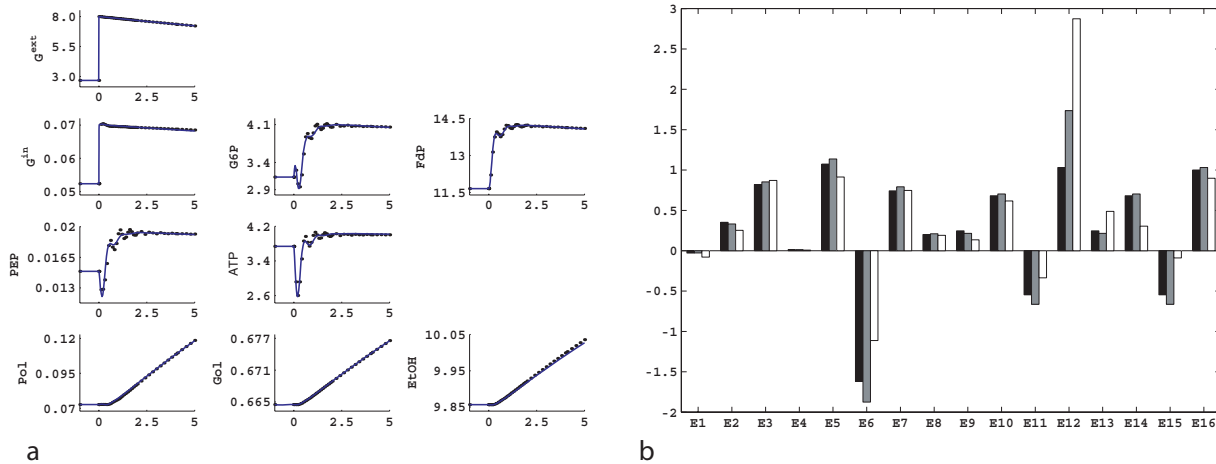


Figure 5 Results of the final estimation. a) Experimental data (black) and the linlog simulation using final elasticities (blue). Units are the same as Figure 1. b) Comparison of the theoretical elasticities (black), initial estimates from steady state perturbation data (grey) and the final estimates (white).

Table 2: The relative deviations of the elasticities, expressed as the standard deviation per mean*100, and the p-value at which the null hypothesis, that the actual elasticity is zero and that the estimated value was caused by the random error only, can not be rejected

Elasticity Parameter	Relative standard deviation $\frac{\sqrt{\sigma^2}}{\mu} \cdot 100$ [%]	p-value
E1	235.3	2.1E-03
E2	9.1	0
E3	10.8	0
E4	295.4	1.0E-02
E5	15.7	0
E6	26.7	0
E7	12.1	0
E8	150.9	1.1E-05
E9	342.4	2.2E-02
E10	60.5	3.3E-16
E11	144.6	5.5E-06
E12	12.0	0
E13	143.0	4.5E-06
E14	159.9	2.7E-05
E15	360.2	2.8E-02
E16	84.5	2.3E-11

and in Figure 6b where the comparison of the estimated elasticities with the theoretical ones is presented. It is clear that with the new set of 12 elasticities (replacing the original 16) the dynamic pulse can also be simulated in good agreement with the experimental data. The value of the objective function increased only by 0.04%, when the degrees of freedom increased by 21% (degrees of freedom increased from 19 (= 35 data points - 16 parameters), to 23 (= 35 data points - 12 parameters)) showing that elimination of these four elasticities in fact improved the quality of the fit.

Cross validation of the obtained model

In order to validate the estimated elasticities, we have performed a cross validation study, for which we have generated an independent dataset, consisting of steady state perturbation data, and we have predicted the effect of that perturbation with our estimated parameters, and compared the results. Since we already perturbed V_{IN} , V_{GAPD} and V_{ATPase} and included the resulting data in the identification dataset (Table 1), we had to select another enzyme in the pathway as the target for independent steady state perturbation. Among the remaining enzymes, we chose to inhibit the activity of V_{PFK} by 50%, because this enzyme assumes many interactions and is known to be a complex enzyme, which would be a challenge to cross validate for our network. The comparison between the model prediction and the experimental results is given in Figure 7. The level of extracellular and intracellular glucose did not change, whereas the levels of G6P and polysaccharides increased considerably and the levels of PEP, FdP, ethanol

and glycerol decreased. The linlog model predicted this steady state perturbation with maximal deviation of only 8.5% in G6P. We conclude that the linlog model with the estimated parameters performed satisfactorily for the test case considered.

Calculation of the systemic properties

After the estimation of the elasticities, we have calculated the flux control coefficients (C^J) using equation (11). The comparison of these with the theoretical control coefficients of the main flux (J_{PK}) is presented in Figure 8. The flux is controlled mainly by three enzymes: the hexose transporter (e_{IN}), and to a lesser extent by the phosphofructokinase (e_{PFK}), and the ATPase (e_{ATPase}). These findings are in qualitative agreement with the literature [31]. It is noticeable that the elimination of the four elasticities did not change these results significantly; the same three reactions still have the main control of the glycolytic flux.

The concentration control coefficients (C^{x0}) are also calculated using equation (10). Figure 9 gives the comparison of the C^{x0} s of the two branch point metabolites, G6P and FdP. The levels of both metabolites can be increased by increasing the activity of the hexose transporter. Increasing the activity of e_{ATPase} will decrease the level of G6P, whereas it increases the level of FdP. Increasing the activity of e_{PFK} will decrease and increase the levels of G6P and FdP respectively. Lastly, an increase in the activity of e_{GAPD} will result in a slight increase in G6P level and a decrease in FdP level. These results are also in agreement with the previous studies. As expected, the model reduction did not

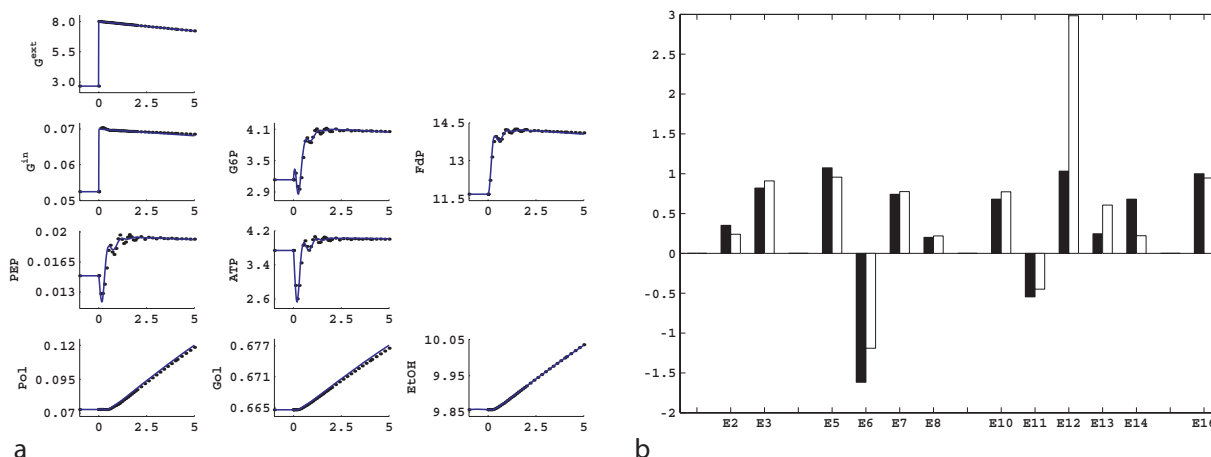


Figure 6
 Results of the estimation of the elasticities, after model reduction. a) Experimental data (black) and the linlog simulation with the reduced elasticity matrix (blue). Units are the same as Figure 1. b) Comparison of the theoretical elasticities (black), with the linlog elasticities (white).

have a significant effect on the estimated control coefficients.

Discussion

Large mathematical models are needed for finding the targets for engineering realistic metabolic systems. Currently, the available large models of several organisms e.g. *Saccharomyces cerevisiae* are mostly stoichiometric models. Although these models are highly relevant for knockout studies and determining the capabilities of the network considered, they have some limitations, e.g., they cannot predict the time courses, do not give insights in molecular events and are based on the "assumed goals" of the cells. Hence, for driving engineering interventions, kinetic models are needed. Here, we have presented a method to estimate the elasticities using the linlog kinetic format, from a combination of steady state and dynamic pulse experiments. In the past, there have been efforts to extract MCA parameters from transient metabolite data. Liao and Delgado presented a method to obtain the control coefficients from such data [34]. Later, it was shown by Elde and Zacchi through MC simulations that this method is highly sensitive to noise [35]. Moreover a full dynamic kinetic model is not obtained. Here, we have described an estimation procedure which directly yields the elasticities as parameters instead of control coefficients, hence allowing the construction of a full kinetic model. Subsequently, the control coefficients can be obtained from the summation and connectivity theorems. Such a full kinetic model can also be used to simulate the effect of different perturbations on metabolism of a microorganism in a fermentor.

The method presented here assumes the availability of dynamic perturbation data. Although the list of metabolites is growing, not all of the metabolites in the cell can be measured. In a recent study, Wang and coworkers described an extension of the MCA, under uncertainty [36-38] in which the authors described a framework to calculate the control coefficients when either there are no measurements on metabolites, or the available measurements are subject to high uncertainty. Their proposed framework can be considered as complementary to the method presented in this paper. In a case where accurate measurements are available, the present paper provides a mathematical approach to obtain the elasticities as kinetic parameters, from which one can calculate the control coefficients.

When compared to previous studies where linlog kinetics have been applied in combination with highly idealized linear pathways or small networks with branch points, the model in this work represents a more realistic case. The model used in this study represents an intermediate size system with branch points and conserved moieties. The proposed method can easily be extended to larger networks that are needed not only to understand the functioning of living cells, but also to infer engineering applications of e.g. microorganisms.

In the example considered in this paper, we have assumed that the zero entries of the elasticity matrix are known. This is a reasonable assumption, since numerous enzymes in the primary metabolism of many organisms are extensively studied and there is a dedicated public compen-

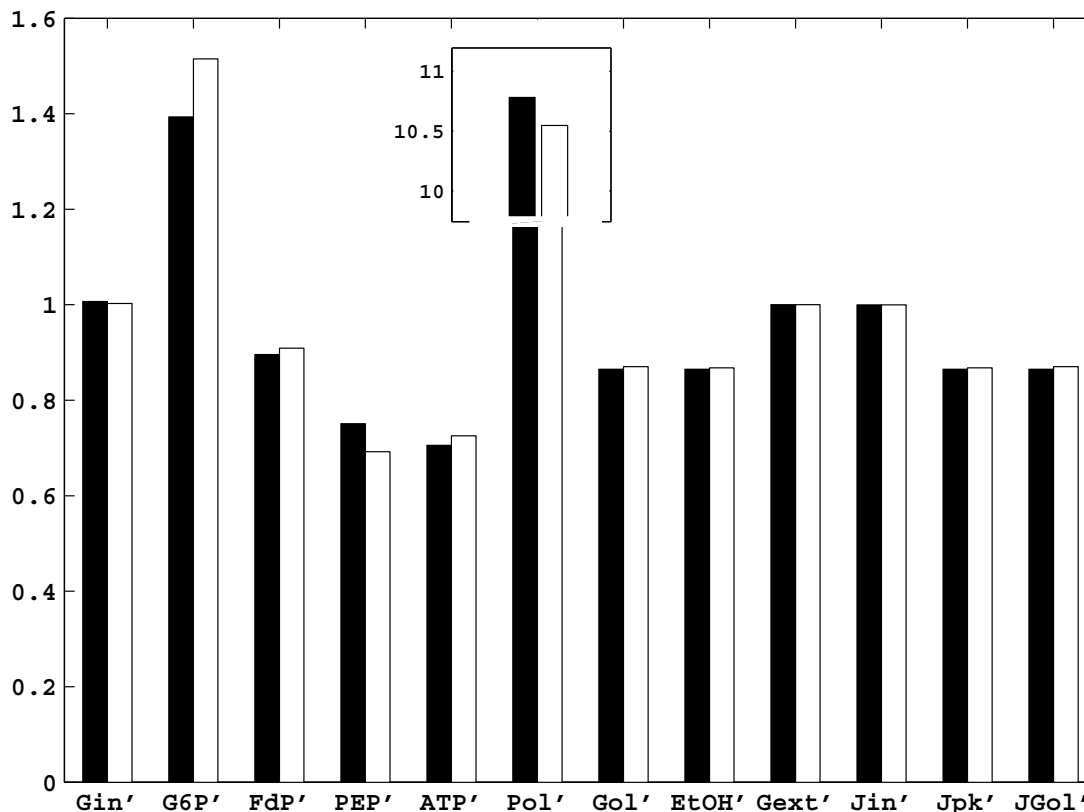


Figure 7

Results of the cross validation study. An independent steady state perturbation is introduced, in which the activity of V_{PFK} is decreased by half (see text). A comparison of the normalized *in silico* experimental results (black) with normalized model prediction (white) are given. The concentrations are normalized with respect to the reference conditions, given in Table 3.

dium for such information [39], allowing the use of *a priori* knowledge on elasticities which are zero.

In the absence of any *a priori* knowledge on the zero elasticities, the elasticities have to be estimated from the full elasticity matrix, for which vastly more experimental data are needed. In the MCA literature, there are several attempts to determine the different perturbations needed in order to determine all elasticities [22,23]. However, one needs to consider that an enzyme will never be affected by all metabolites; there is a physical limit for e.g. the maximum number of binding sites of each enzyme in a biological network.

From the results presented here, it has been shown that, despite the rich information content of the data obtained from dynamic experiments, not every elasticity of the network could be correctly estimated. This problem of

parameter identifiability would be much more pronounced in a case where all (possible) interactions are taken into account, caused by the combinatorial explosion of number of parameters to be estimated. In the current work, in order to resolve some of the interactions which could not be resolved from the dynamic data, steady state enzyme perturbations have been introduced. In order to get full kinetic models of microorganisms, accurate measurements of metabolite concentrations resulting from independent perturbations are needed, such as presented in [28] and [40]. At this point it is important to state that the property of the linlog kinetic format, that the rate equations are linear in the kinetic parameters, allows simultaneous use of alternative datasets by concatenating them in one parameter estimation scheme, i.e. it is straightforward to extend the data matrix Y in equation (9) and the vector χ_{sim} in equation (6), with the data from alternative dynamic and steady state pertur-

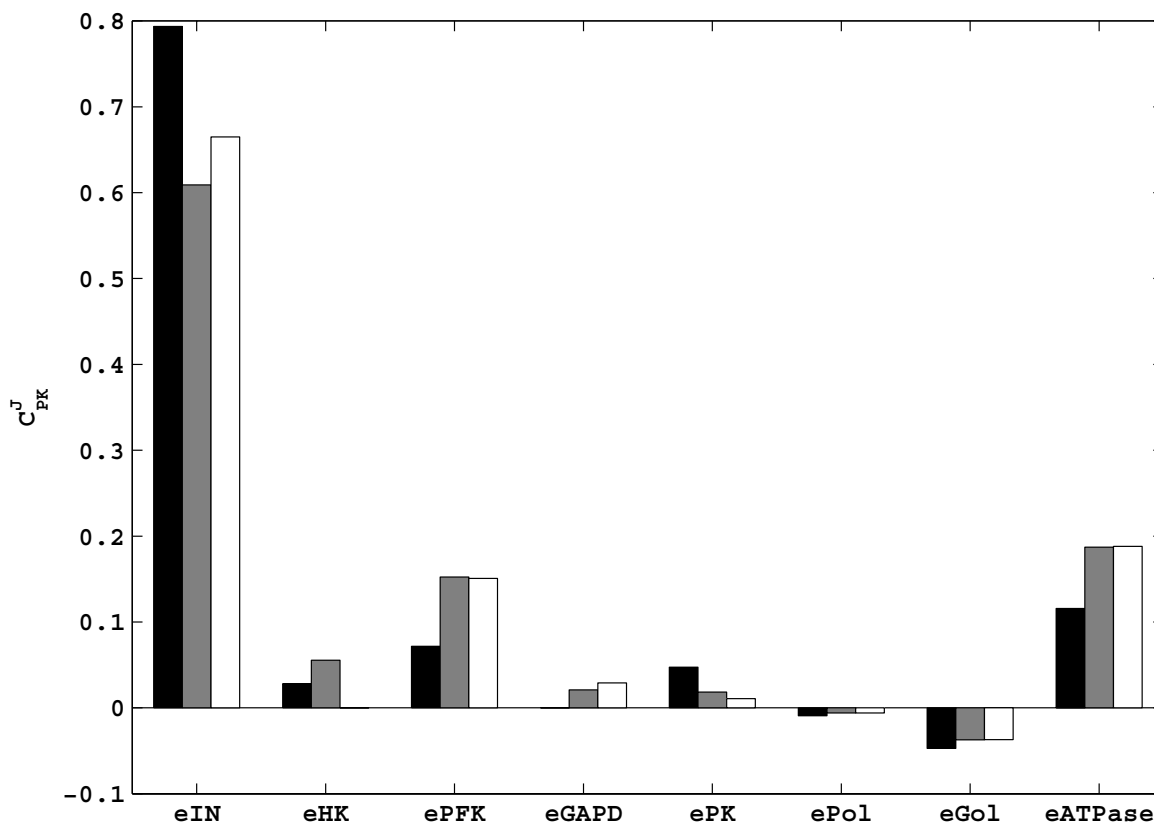


Figure 8

Comparison of the systemic properties, flux control coefficients (C_J^0) of the ethanol flux. First column (black): theoretical C_J^0 s calculated using theoretical elasticities. Second column (grey): C_J^0 s calculated using estimated elasticities in Figure 5b. Third column (white): C_J^0 s calculated using the estimated elasticities of the reduced model (Figure 6b).

bation experiments. In addition, the decoupling of the parameters is also immediate i.e. we can isolate poorly identifiable elasticities and estimate the remaining ones accurately and thereby reduce the problem.

It is worth mentioning that the obtained parameters (elasticities) are true kinetic parameters that reflect the properties of the enzymes with respect to the corresponding metabolites. They can be assumed to remain invariant as long as the enzyme keeps its properties in response to the changes in the environmental conditions. Additionally, since the linlog kinetics provides an approximation to the actual rate of the corresponding reaction over certain interval, the elasticity parameter that performs best may be different than the theoretical value of the corresponding elasticity. This should be kept in mind when comparing some of the less well determined elasticity values to the theoretical ones.

Conclusion

Constructing dynamic models of metabolic reaction networks under *in vivo* conditions using data obtained from perturbation experiments remains still a challenging problem in the area of systems biology. In this contribution, we presented a method which allows the determination of the elasticities directly from experimental data from typical dynamic and/or steady state perturbation experiments. These elasticities allow the reconstruction of the full kinetic model of the glycolysis of *Saccharomyces cerevisiae*, and the determination of the control coefficients. We further show by *a posteriori* parameter identifiability analysis that a subset of elasticities could not be identified using dynamic perturbation data only. Introduction of additional experimental information, i.e. steady state experiments, solved this parameter identification problem.

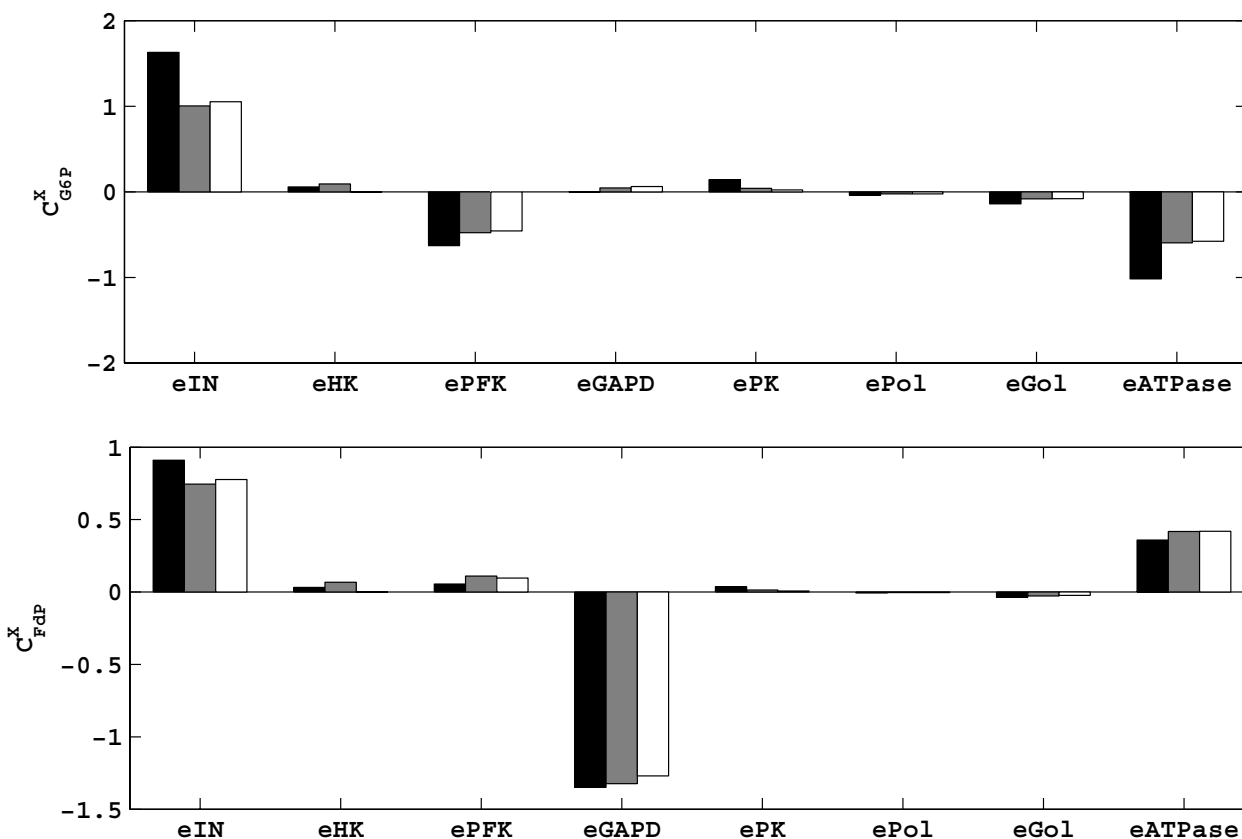


Figure 9

Comparison of the systemic properties, concentration control coefficients (C^{x0}) of the two branch point metabolites: G6P (upper panel) and FdP (lower panel). First column (black): theoretical C^{x0} s calculated using theoretical elasticities. Second column (grey): C^{x0} s calculated using estimated elasticities in Figure 5b. Third column (white): C^{x0} s calculated using the estimated elasticities of the reduced model (Figure 6b).

Considering the description of the dynamics, results of the cross validation studies and the final values of the elasticities, we conclude that linlog kinetics, although being an approximate kinetic format, performs very satisfactorily for estimating elasticities from data obtained from dynamic simulation of a mechanistic model, that was realistic in terms of biochemical complexity (glycolysis), noise added and sampling frequency.

Methods

In linlog kinetics, all rate equations have the same mathematical structure: proportionality to the enzyme level and linearity in the parameters (elasticities, ϵ_x^v and ϵ_c^v) as represented by equation (2). This equation can be generalized in vector form to represent the rate vector of the metabolic network under consideration

$$v = J^0 \cdot e' \cdot (i + E^x \ln(x') + E^c \ln(c')) \quad (3)$$

where the vector v is the $(r \times 1)$ rate vector, r being the number of rates, J^0 is the square diagonal matrix containing the reference state fluxes ($J^0 = \text{diag}(J_i^0) \ i = 1, \dots, r$), e' is the square diagonal matrix containing relative enzyme levels ($e' = \text{diag}(e_i/e_i^0) \ i = 1, \dots, r$), i is the $(r \times 1)$ vector of ones, E^x and E^c are the $(r \times m_x)$ and $(r \times m_c)$ elasticity matrices, m_x and m_c being the number of intracellular and extracellular metabolites respectively, and x' and c' are the $(m_x \times 1)$ and $(m_c \times 1)$ vectors containing relative concentrations of the intracellular and extracellular metabolites respectively ($x' = x_j/x_j^0 \ j = 1, \dots, m_x$ and $c' = x_k/x_k^0 \ k =$

1,...,m_c). In each term, the superscript ⁰ indicates the reference state values.

Determining elasticities from steady state perturbation data

Given that for steady state perturbation experiments, we obtain information on fluxes (J_i's) using the mass balances and the measured metabolite concentrations (x_j), we can directly use the equation (3) for the estimation of the elasticities. When it is rearranged, the equation (3) can be presented in the following standard linear model:

$$\mathbf{a} = \mathbf{Y} \cdot \mathbf{b} \quad (4)$$

where **a** is the (r × 1) measurement vector that contains the measured normalized fluxes and normalized enzyme levels (**a** = (J⁰e')⁻¹ v - i), **b** the (p × 1) vector that contains the non-zero elasticities of the original elasticity matrices E^x and E^c, and **Y** is the (r × p) design matrix of which each ith row contains as nonzero elements ln(x_j/ x_j⁰) and ln(c_k/ c_k⁰) at positions corresponding with the non-zero elasticities ε^{v_i}_{x_j} and ε^{v_i}_{c_k} in vector **b**. The equation (4) can be solved to obtain the elasticities using linear regression according to:

$$\mathbf{b} = (\mathbf{Y}^T\mathbf{Y})^{-1} \mathbf{Y}^T \cdot \mathbf{a} \quad (5)$$

This shows that estimation of elasticities from steady state perturbations requires data on metabolite levels (ln(x_j/ x_j⁰) and ln(c_k/ c_k⁰) presented in matrix **Y**) enzyme activities (presented in matrix e') and steady state fluxes (presented in the vector v and the matrix J⁰).

Determining elasticities from dynamic perturbation data

Since in a dynamic perturbation experiment, the rate information can not be obtained directly, we will follow different procedure here and we will only make use of the time profiles of the measured metabolites (x_j's) in treating dynamic perturbation data. The elasticities are, in this case, estimated via non-linear parameter estimation procedure, in which the objective function to be minimized is the weighed squared error between the experimental response and simulation results. The general form of the objective function is given in equation(6):

$$\text{objfun} = (1 - \lambda) \sum_{i=1}^m \alpha_i \cdot \left(\sum_{j=1}^q \left(\frac{c_{\text{exp},i,j} - c_{\text{sim},i,j}}{x_{0,i}} \right)^2 \right) + \lambda \cdot \sum_{k=1}^r \left(\frac{J'_{k,\text{exp}} - J'_{k,\text{sim}}}{J_k^0} \right)^2 \quad 0 \leq \lambda \leq 1 \quad (6)$$

the limits of the summations, m, q and r are the number of metabolites, measured time points and fluxes respectively. In equation (6), χ_{exp} is the experimentally measured, χ_{sim} is the simulated (using linlog kinetics)

metabolite matrix containing intracellular and extracellular metabolite concentrations. Two additional parameters are introduced here:

γ to weigh the effect of uncertainty in different metabolites and λ to weigh the effect of two different sources of data, namely steady state and dynamic perturbation data. In the second term of the equation(6), J'_{sim} and J'_{exp} represent the experimental and simulated steady state fluxes. The equation (6) contains the rate information implicitly, since the χ_{sim} results from the integration of the set of ode's representing the dynamics of the system. Note also that although we already explained the treatment of the steady state perturbation data in the previous section, we explicitly included the second term in equation (6) again; to state clearly that this form of the objective function allows also the integration of data from different types of experiments.

Two main classes of optimization algorithms are available to minimize the objective function in equation (6): greedy algorithms and evolutionary algorithms. Moles *et al.* presented a comparison of global optimization methods used for parameter estimation in biochemical pathways [41]. In that review, they discussed various global optimization methods and concluded that the algorithm that uses evolutionary strategy using stochastic ranking performed best. On the other hand, they also pointed out that, generally, evolutionary algorithms require high computational effort. The alternative, greedy algorithms, are fast, but in turn require an initial estimate close to the optimal solution. A good initial estimate is necessary not only to evade local minima and improve the solution performance, (i.e. convergence time, finding a global optimum) but also to prevent highly stiff systems, which increase the computation time. In this work, we chose to use a greedy (simplex) algorithm, mainly because the linlog kinetic format has the advantage to provide a good initial estimate that can be obtained directly from the experimental data via linear regression (see below) so that the method presented in this paper does not require the robustness of the evolutionary algorithms towards the initial estimate. It is noteworthy that alternative approximative kinetic formats such as the two proposed formats of BST (i.e. GMA or S-system forms) lack this advantage of providing a good initial guess to the non-linear regression step. With these formats, zero is generally assumed as the initial guess for the non-linear parameter estimation problem [42].

To obtain initial estimates, we start with the general dynamic model of a metabolic system in a typical rapid pulse experiment in a chemostat which is given by the mass balances for the m_x intracellular (x) and m_c extracellular (c) metabolites:

$$\begin{aligned} \frac{dx}{dt} &= S \cdot v - \mu \cdot x \\ \frac{dc}{dt} &= D \cdot (c_{feed} - c) + S_c \cdot v - c_X \end{aligned} \quad (7)$$

Where x and c are the $(m_x \times 1)$ and $(m_c \times 1)$ vectors containing concentrations of intracellular and extracellular metabolites respectively, expressed in $\mu\text{mol gDW}^{-1}$ and $\mu\text{mol L}^{-1}$ respectively, S and S_c are $(m_x \times r)$ and $(m_c \times r)$ stoichiometric matrices for the intra and extracellular metabolites respectively; v is the $(r \times 1)$ biomass specific reaction rate vector in $\mu\text{mol gDW}^{-1} \text{hr}^{-1}$, c_{feed} is the $(m_c \times 1)$ vector containing concentration of the extracellular metabolites in the feed expressed in $\mu\text{mol L}^{-1}$, D is the dilution rate (hr^{-1}), μ is the biomass specific growth rate (hr^{-1}) and c_X is the biomass concentration in gDW L^{-1} .

Substitution of v by linlog kinetic rate equation (3) in the mass balance equation (7) and assuming that there is no change in the enzyme levels ($e' = I_{r \times r}$) yields:

$$\begin{aligned} \frac{dx}{dt} &= S \cdot J^0 \cdot (i + E^x \cdot \ln(x') + E^c \cdot \ln(c')) - \mu \cdot x \\ \frac{dc}{dt} &= D \cdot (c_{feed} - c) + S_c \cdot J^0 \cdot (i + E^x \cdot \ln(x') + E^c \cdot \ln(c')) \cdot c_X \end{aligned} \quad (8)$$

After rearrangement (and taking into account that due to steady state of the reference, $S \cdot J^0 \cdot i = 0$), the set of equations are integrated for each metabolite, from t_i to t_{i+1} ,

$$\begin{aligned} \Delta x + \mu \cdot \int_{t_i}^{t_{i+1}} x dt &= S \cdot J^0 \cdot [E^x E^c] \cdot \int_{t_i}^{t_{i+1}} \begin{bmatrix} \ln(x') \\ \ln(c') \end{bmatrix} dt \\ \Delta c - \left(D \cdot \int_{t_i}^{t_{i+1}} (c_{feed} - c) dt + S_c \cdot J^0 \cdot i \cdot c_X \cdot \Delta t \right) &= S_c \cdot J^0 \cdot [E^x E^c] \cdot \int_{t_i}^{t_{i+1}} \begin{bmatrix} \ln(x') \\ \ln(c') \end{bmatrix} dt \cdot c_X \end{aligned}$$

This can be presented in the following standard linear model:

$$\begin{bmatrix} a^x \\ a^c \end{bmatrix} = \begin{bmatrix} Y^x \\ Y^c \end{bmatrix} \cdot b \quad (9)$$

Where

$$\begin{aligned} a^x &= \Delta x + \mu \cdot \int_{t_i}^{t_{i+1}} x dt \\ a^c &= \Delta c - \left(D \cdot \int_{t_i}^{t_{i+1}} (c_{feed} - c) dt + S_c \cdot J^0 \cdot i \cdot c_X \cdot \Delta t \right) \end{aligned}$$

and Y is a $((q - 1) \cdot (m_x + m_c)) \times p$ matrix, combining Y^x and Y^c , containing the time integrals of the logarithm of the

normalized metabolite concentrations, and S, S_c, J^0 . Here q is the number of measured time points. b is a $(p \times 1)$ vector containing the unknown elasticity coefficients, which is estimated by linear regression similarly as above in eq. (5). To obtain the integrals of equation(9), a linear interpolation of concentration between the measurements at t_i and t_{i+1} is used.

Although both equations ((6) and (5)) are based on least square principle, the use of equation (6) (where b of equation (5) is used as an initial estimate) has advantages, namely, further improvement of the quality of the estimated parameters, because initial linear regression assumes that errors are only present in the dependent variables (a in Eq. (5)), whereas errors in the measured metabolites in fact also affect the independent variables (matrix Y in Eq.(4) and Eq.(9)). Furthermore, the non-linear optimization allows the incorporation of additional degrees of freedom for the correction of errors in the metabolite levels at the first data point (t_0) used for the model simulation (integration of Eq.(8)). Moreover, during the non-linear parameter estimation, linear interpolation between the logarithms of the measured metabolites is not needed anymore. An additional ease of the non-linear regression is in the introduction of the adjustable parameter γ to weigh the effect of one (or more) metabolite(s) on the fit. This is useful, when the measurement precision is poor for a certain metabolite or when the relative change in one metabolite is small compared to the others. The obvious choice for γ would be the inverse of the standard deviation of the corresponding metabolite. Although the weighing parameter γ can be introduced in the linear regression step as well, introducing the factor γ in the non-linear regression step is much more straightforward.

Calculation of the systemic properties

Having estimated the elasticities, we can calculate scaled flux and concentration control coefficients

$$\left(C_{v_k}^J = \frac{v_k^0}{J} \cdot \frac{\partial J}{\partial v_k} \Big|_0, C_{v_k}^{x_i} = \frac{v_k^0}{x_i^0} \cdot \frac{\partial x_i}{\partial v_k} \Big|_0 \right), \text{ using classical summation and connectivity theorems [34]:}$$

$$C^{x0} = -L^x \cdot (S \cdot J^0 \cdot E^{x0} \cdot L^x)^{-1} \cdot S \cdot J^0 \quad (10)$$

$$C^{J0} = I + E^{x0} \cdot C^{x0} \quad (11)$$

Here L^x is the metabolite link matrix, which links the dependent metabolites to the independent ones and J^0, S and E^{x0} are steady state flux matrix, stoichiometric matrix and the elasticity matrix as represented in the previous sections. The scaled response coefficient (R^{J0}) for the external metabolites can also be calculated [43]:

$$R^{j0} = C^{j0} \cdot E^{c0} \quad (12)$$

Metabolic model

The glycolytic pathway of *Saccharomyces cerevisiae* has been extensively studied in terms of enzyme kinetics and metabolite levels. Over the last 30 years, modeling of the glycolysis of yeast cells has been applied for several reasons, such as simulation of physiology to understand regulation under dynamic conditions, MCA to amplify and/or redirect metabolic flux [29,31,44-47], and to investigate how changing environmental conditions change metabolism.

In order to test the proposed method for estimation of elasticities from steady state and dynamic perturbations and the applicability of linlog kinetics, noise-free data were generated using the modified version of the mechanistic model of Galazzo and Bailey describing the glycolysis of yeast cells, which contains 8 reactions and 9 metabolites (Figure 10). The set of mechanistic kinetic equations are highly non-linear and contain 41 parameters. The parameters used were taken from Galazzo and Bailey [31] using the experimental settings for $\text{pH}^{\text{ext}} = 5.5$ with suspended cells in the original study. The glucose uptake rate of the original model was modified in order to be able to mimic realistic fermentation conditions. The details of the metabolic model, the list of mechanistic rate equations and the mass balances describing the kinetics of the metabolic network of Figure 10 are given in the Appendix.

Reactor model

The chosen glycolytic model was not designed to describe growth, so in order to prevent washout of the cells and to keep a constant amount of biomass within the fermentor, the organism was assumed to be cultivated in a chemostat with complete biomass retention. The glucose concentration in the feed was set at a low value resulting in low steady state ethanol and glycerol concentrations, thus allowing measurable changes in those metabolites when a perturbation was applied. Furthermore the biomass concentration was set at 5 gDW L⁻¹ and the dilution rate was set at 1.5 hr⁻¹ to provide a reasonable substrate load (q_s) and glycolytic flux. The resulting reference steady state conditions (fluxes, metabolite levels) are represented in Table 3.

Linlog kinetic model

The number of non-zero entries of the elasticity matrix ($[E^x E^c]$ in equation(3)) defines the number of elasticities to be estimated. Specifically, the elasticity matrix has 72 entries (8 reactions \times 9 metabolites) and it followed that 16 of these were nonzero for the Galazzo and Bailey model which was extended with the alterations on the uptake reaction, mentioned in the Appendix. For the

present network, the structure of the elasticity matrix is presented in Figure 11. The linlog model contains 16 elasticities and 8 reference rates as parameters whereas the mechanistic model has 41 parameters.

Dynamic perturbations

To obtain transient data, the reference steady state (Table 3) was perturbed by increasing the extracellular glucose concentration (C^{ext}) at $t = 0$ from 2666 $\mu\text{mol L}^{-1}$ to 8000 $\mu\text{mol L}^{-1}$, where after the relaxation of the system was simulated. The levels of the metabolites were recorded until the system returned to the reference state.

In choosing the number of data points, i.e. the sampling frequency, special attention had been paid to use a realistic number of data points that an experimenter can obtain from a typical rapid sampling experiment (for a short time period) or from a dynamic pulse experiment (for a longer time period). The data were obtained as follows: in the first 120 seconds, where the initial dynamics after the glucose uptake are important, 20 samples uniformly distributed over time were taken, in the next 120 seconds, 10 samples and in the last minute 5 samples were taken. For long term experiments, from the 5th minute until the 10th minute, 5 samples were taken, and from the 10th minute until the end (400 min) 30 uniformly distributed samples were taken. Metabolites for which concentrations were assumed to be available at these times are intracellular glucose, G6P, FdP, PEP, ATP and extracellular metabolites glucose, ethanol, glycerol and polysaccharides.

Steady state perturbations

For the steady state perturbations, the enzyme levels were modulated and the system was allowed to reach a new steady state. The new steady state conditions (fluxes, metabolite levels) were measured.

The minimum number of independent steady state perturbations needed in order to identify all the elasticities for a reaction equals the number of non-zero elasticities for that rate expression. From Figure 11, where the elasticity matrix is shown, it can be inferred that the minimum number of perturbations to be applied per reaction is 3, due to the V_{PK} and V_{Gol} reactions that have three non-zero elasticities (E9–E10–E11 and E13–E14–E15 respectively).

As the reaction to be perturbed, V_{IN} was chosen because it is at the beginning of the glycolysis and therefore affects the entire pathway. As the second reaction to be perturbed, V_{ATPase} was chosen because it directly affects the level of ATP which is the effector of 6 out of 8 reactions. Lastly, V_{GAPD} was chosen for two reasons: 1) to resolve the branch point relations around G6P and 2) to resolve the feed forward effect of FdP on V_{PK} . Note that since the production rate of glycerol (V_{Gol}) was coupled (proportional)

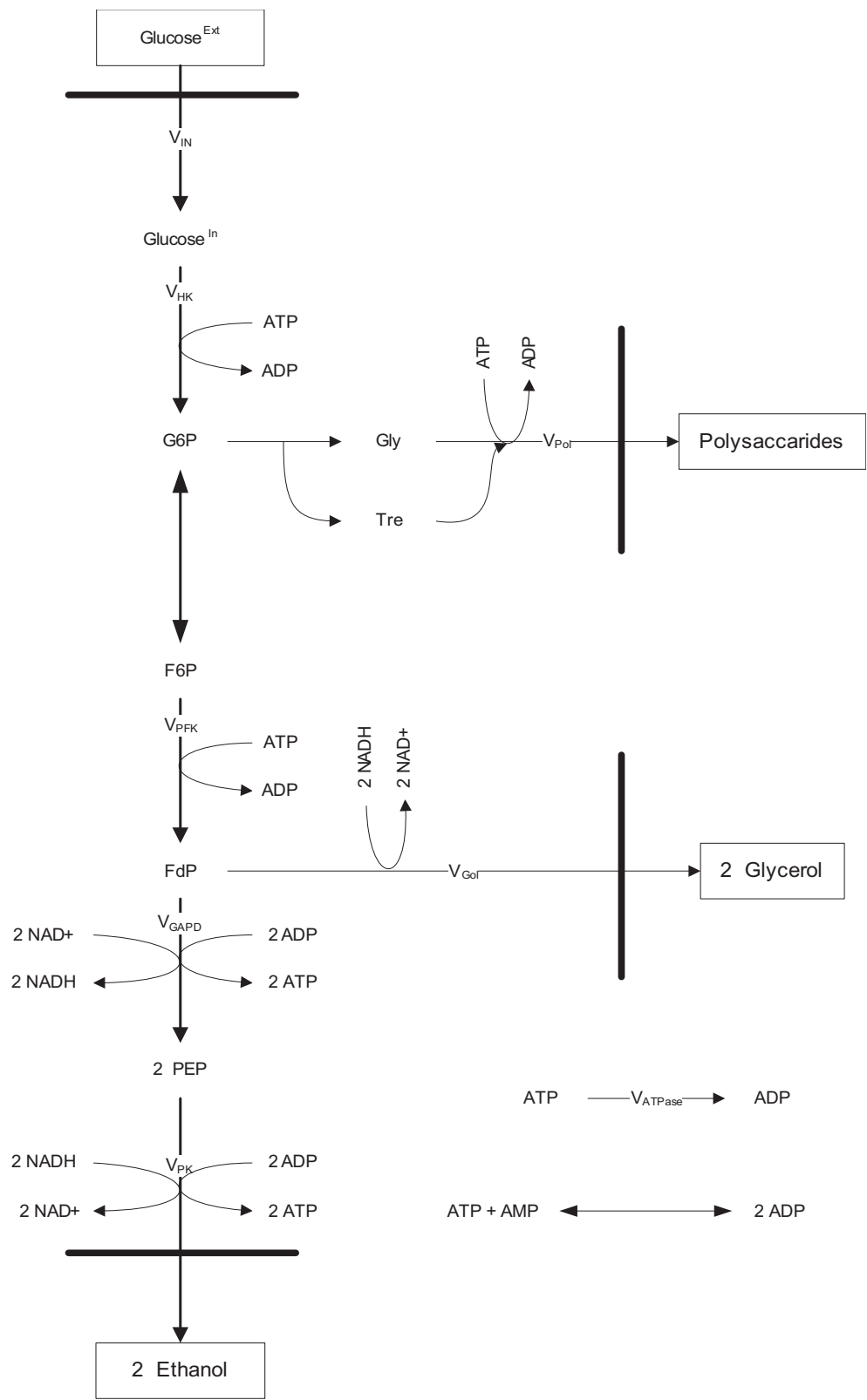


Figure 10
Glycolytic Pathway of *Saccharomyces cerevisiae* (Adapted from Galazzo and Bailey, 1990).

Table 3: Reference conditions of fermentation and microorganism

Fermentation parameters	
$D = 1.5 \text{ hr}^{-1}$, $c_{\text{Glucose}}^{\text{feed}} = 8000 \mu\text{mol L}^{-1}$, $c_X = 5 \text{ gDW L}^{-1}$	
Biological parameters	
Extracellular metabolite concentrations [$\mu\text{mol L}^{-1}$]	
G^{ext}	2666.8
Pol	72.7
Gol	664.5
EtOH	9856.5
Macroscopic Fluxes [$\mu\text{mol gDW}^{-1} \text{ hr}^{-1}$]	
q_s	1600
q_{Pol}	21.8
q_{Gol}	199.4
q_{EtOH}	2957
Intracellular metabolite concentrations [$\mu\text{mol gDW}^{-1}$]	
G^{in}	0.0524
G6P	3.133
FdP	11.657
PEP	0.0149
ATP	3.738
Intracellular rates [$\mu\text{mol gDW}^{-1} \text{ hr}^{-1}$]	
V_{IN}	1600
V_{HK}	1600
V_{PFK}	1578.2
V_{GAPD}	1478.5
V_{PK}	2957
V_{Pol}	21.8
V_{Gol}	199.4
V_{ATPase}	2714

to the production of ethanol (V_{PK}), the steady state branch point relations around FdP can be reformulated as follows:

At steady state, $J_{\text{PFK}} = J_{\text{GAPD}} + J_{\text{Gol}}$ and $J_{\text{GAPD}} = J_{\text{PK}}/2$. Additionally, as described above, $J_{\text{Gol}} = k \cdot J_{\text{PK}}$, hence $J_{\text{PFK}} = (1 + 2 \cdot k) \cdot J_{\text{GAPD}}$ where the value of k is given in the Appendix. Therefore, there is no need to introduce additional steady state perturbations to resolve the branch point relations around FdP. Table 1 summarizes the steady state perturbations used as data in this study. For example, in the first column, the activity of the glucose uptake rate was increased by 20% and the system is allowed to reach the new steady state, the fluxes and the metabolite levels relative to the reference state are presented in the corresponding column. Similarly, the second column represents the inhibition of V_{ATPase} by 10% and the third represents the data after the inhibition of V_{GAPD} by 20%.

Error propagation analysis

In order to inspect whether the estimated elasticities are robust against experimental noise in metabolite concentrations, we performed Monte Carlo (MC) simulations. To this end, random errors of 10% are repeatedly added to form 50 sets of "noisy data", which were used in the non-linear parameter estimation procedure. This yields a distribution of estimated elasticities which allows calculating the relative standard deviation for each elasticity as the ratio of standard deviation to the mean of that elasticity distribution.

Parameter identifiability and model reduction

The elasticities that have a high relative standard deviation are suspected to be poorly identifiable in the network. Potentially unidentifiable elasticities were recognized by calculating a p-value for testing the null hypothesis that the true mean of a given elasticity distribution is zero

	G^{in}	G6P	FdP	PEP	ATP	G^{ext}
v_{In}	$E1$	$E2$
v_{HK}	$E3$.	.	.	$E4$.
v_{PFK}	.	$E5$.	.	$E6$.
v_{GAPD}	.	.	$E7$.	$E8$.
v_{PK}	.	.	$E9$	$E10$	$E11$.
v_{Pol}	.	$8 \cdot E12$
v_{Gol}	.	.	$E13$	$E14$	$E15$.
v_{ATPase}	$E16$.

Figure 11

Nonzero entries of the elasticity matrix. Adapted from Galazzo and Bailey, 1990, with modifications in the v_{In} reaction. For the modifications in the uptake reaction, see text.

through the statistic $\mu/(\sigma/\sqrt{N})$ that has a t-distribution with N degrees of freedom, where N is the number of MC simulations. The elasticities that were found to be potentially unidentifiable were set to zero, thus reducing the number of non-zero entries in the elasticity matrix. Subsequently, for the reduced model, the remaining elasticities were re-estimated, using the same steady state and dynamic perturbation data and the parameter estimation procedure explained in the previous sections.

Software used

MATLAB software (version 6.5, Stat-Ease Inc., Minneapolis, USA) was used for generating the *in silico* data and analyses of the data and the parameters obtained. The non-linear minimization problem of parameter estimation was solved using the Nelder-Mead simplex search method.

Authors' contributions

IEN set up the methodology, performed the calculations and drafted the manuscript. WvW supervised the calculations, the organization of the manuscript and revised the manuscript. WvG contributed in organizing and revising the manuscript. JJH supervised the whole project. All of the authors read and approved the final manuscript.

Appendix**Details of the *in silico* metabolic model**

In this paper, the model of Galazzo and Bailey describing the glycolysis of yeast cells is used. This section describes the details of the metabolic model. The model consists of 5 intracellular (glucose^{intracellular}, G6P, FdP, PEP and ATP), 4 extracellular (glucose^{extracellular}, polysaccharides, glycerol and ethanol) metabolites and 8 reactions. In their work, a mathematical model for the description of ethanol production by non-growing *Saccharomyces cerevisiae* has been constructed using non-linear mechanistic rate equations. The glucose uptake is modeled by symmetrical carrier model (see below for the changes in the uptake reaction),

the hexokinase reaction is described by a two substrate single displacement mechanism, the phosphofructokinase and pyruvate kinase reactions are described by allosteric kinetics according to Monod-Wyman-Changeux, polysaccharide formation is explained by Hill type kinetics and the glyceraldehyde 3-phosphate dehydrogenase reaction is described by a kinetic expression which takes into account the crossed product inhibition by G3P and NAD⁺ and competitive inhibition by AMP, ADP and ATP. The ATP consumption is assumed to have first order kinetics.

One property of the original model is that the glycerol production is assumed to be stoichiometrically coupled to ethanol formation, i.e. there is no separate dynamics for the glycerol formation. Additionally, the NAD⁺ and NADH levels are assumed to be constant, as given below.

As it was originally published, the glucose influx was a linear function of the glucose-6-phosphate concentration only; hence, it was insensitive to the changes in extracellular glucose concentration. To mimic realistic fermentation conditions, where one can perturb the organism by adding a specific compound such as glucose to a chemostat, we have modified the glucose uptake reaction and modeled the glucose uptake by a symmetrical carrier model, suggesting that the transport of glucose across the cell membrane occurs via facilitated diffusion, as proposed by Teusink et al., [4]. The list of mechanistic rate equations used is given in the following section.

Mechanistic rate equations used to construct the in silico network. (Adapted from Galazzo and Bailey, 1990)

The rate equations of the reactions in Figure 10 are given by:

$$V_{IN} = 17.223 \frac{(G^{ext} - G^{in})}{1.1918} \frac{1}{1 + \frac{G^{ext}}{1.1918} + \frac{G^{in}}{1.1918} + 0.91 \frac{G^{ext} \cdot G^{in}}{1.1918 \cdot 1.1918}}$$

$$V_{HK} = 68.5 \cdot \left(\frac{1}{\frac{0.0062 \cdot 0.1}{G^{in} \cdot ATP} + \frac{0.11}{G^{in}} + \frac{0.1}{ATP} + 1} \right)$$

$$V_{PFK} = 31.7 \frac{166.67 \cdot F6P \cdot ATP \cdot R}{R^2 + LT^2}$$

with

$$R = 1 + F6P + 16.67 ATP + 16.67F6P \cdot ATP$$

$$T = 1 + 5 \cdot 10^{-4} F6P + 16.67 ATP + 0.00833F6P \cdot ATP$$

$$L = L_0 \left(\frac{1 + 0.76AMP}{1 + 40AMP} \right)^2, L_0 \text{ is pH dependent}$$

$$V_{GAPD} = 49.9 \left[1 + \frac{0.0025}{G3P} + \frac{0.18}{NAD^+} \left(1 + \sum_{i=1}^3 \frac{\zeta_i}{K_i} \right) + \frac{0.0025 \cdot 0.18}{G3P \cdot NAD^+} \left(1 + \frac{NADH}{0.0003} \right) \left(1 + \sum_{i=1}^3 \frac{\zeta_i}{K_i} \right) \right]^{-1}$$

with

$$\zeta_{i,i=1,2,3} = AMP, ADP, ATP \quad K_{i,i=1,2,3} = 1.1, 1.5, 2.5$$

$$V_{PK} = \left(\frac{3440}{1 + \frac{K_{pH}}{H^+}} \right) \frac{3.1976 \cdot PEP \cdot ADP \cdot R + 0.0004PEP \cdot ADP \cdot T \cdot L}{R^2 + LT^2}$$

with

$$R = 1 + 159.88PEP + 0.2ADP + 3.1976PEP \cdot ADP$$

$$T = 1 + 0.02PEP + 0.2ADP + 0.004PEP \cdot ADP$$

$$L = L_0 \cdot \left(\frac{1 + 0.05 \cdot FdP}{1 + 5 \cdot FdP} \right)^2, L_0 \text{ pH dependent}$$

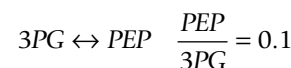
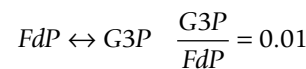
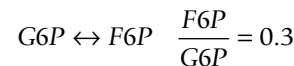
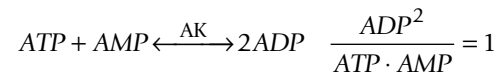
$$V_{Pol} = 1.1 \cdot V_{Gly} V_{Gly}^{app} = \frac{14.31 \cdot G6P^{8.25}}{2^{8.25} + G6P^{8.25}} \cdot \left(\frac{1}{\frac{1}{UDPG} \left(1 + \frac{1.1}{G6P} \right) + 1} \right)$$

$$V_{Gol} = 0.06742 \cdot V_{PK}$$

$$V_{ATPase} = 12.1 \cdot ATP$$

Equilibrium relations

The following metabolites were assumed in equilibrium throughout the simulated time window:



Mass balances

The concentrations of the extracellular metabolites are described by the following mass balances:

$$\frac{dc_{G^{ext}}}{dt} = D \cdot (c_{G^{ext}}^{feed} - c_{G^{ext}}) - V_{IN} \cdot c_X$$

$$\frac{dc_{Pol}}{dt} = D \cdot (c_{Pol}^{feed} - c_{Pol}) + V_{Pol} \cdot c_X$$

$$\frac{dc_{Gol}}{dt} = D \cdot (c_{Gol}^{feed} - c_{Gol}) + V_{Gol} \cdot c_X$$

$$\frac{dc_{EtOH}}{dt} = D \cdot (c_{EtOH}^{feed} - c_{EtOH}) + V_{PK} \cdot c_X$$

Acknowledgements

The authors would like to thank Peter J.T. Verheijen for critically reading the manuscript. This project was financed by the Kluyver Centre for Genomics of Industrial Fermentation, a Center of Excellence sponsored by the Netherlands Genomics Initiative.

References

- Bailey JE: **Toward a science of metabolic engineering.** *Science* 1991, **252**:1668-1675.
- ter Kuile BH, Westerhoff HV: **Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway.** *FEBS Lett* 2001, **500**:169-171.
- Varma A, Palsson BO: **Metabolic Capabilities of Escherichia coli: I. Synthesis of Biosynthetic Precursors and Cofactors.** *Journal of Theoretical Biology* 1993, **165**:477-502.
- Varma A, Palsson BO: **Metabolic Capabilities of Escherichia coli II. Optimal Growth Patterns.** *Journal of Theoretical Biology* 1993, **165**:503-522.
- van Gulik WM, Heijnen JJ: **A metabolic network stoichiometry analysis of microbial growth and product formation.** *Biotechnology and Bioengineering* 1995, **48**:681-698.
- Segre D, Vitkup D, Church GM: **Analysis of optimality in natural and perturbed metabolic networks.** *PNAS* 2002, **99**:15112-15117.
- Varner J, Ramkrishna D: **Metabolic engineering from a cybernetic perspective. 2. Qualitative investigation of nodal architectures and their response to genetic perturbation.** *Biotechnol Prog* 1999, **15**:426-438.
- Varner J, Ramkrishna D: **Metabolic engineering from a cybernetic perspective. 1. Theoretical preliminaries.** *Biotechnol Prog* 1999, **15**:407-425.
- Ramkrishna D, Kompala DS, Tsao GT: **Are microbes optimal strategists?** *Biotechnol Prog* 1987, **3**:121-126.
- Richey B, Cayley DS, Mossing MC, Kolka C, Anderson CF, Farrar TC, Record MT: **Variability of the Intracellular Ionic Environment of Escherichia-Coli - Differences Between Invitro and In vivo Effects of Ion Concentrations on Protein-Dna Interactions and Gene-Expression.** *Journal of Biological Chemistry* 1987, **262**:7157-7164.
- Schaefer U, Boos W, Takors R, Weuster-Botz D: **Automated sampling device for monitoring intracellular metabolite dynamics.** *Analytical Biochemistry* 1999, **270**:88-96.
- Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV, Snoep JL: **Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry.** *European Journal of Biochemistry* 2000, **267**:5313-5329.
- Vaseghi S, Baumeister A, Rizzi M, Reuss M: **In vivo dynamics of the pentose phosphate pathway in Saccharomyces cerevisiae.** *Metab Eng* 1999, **1**:128-140.
- Savageau MA: *Biochemical Systems Analysis, A study of function and Design in Molecular Biology* Addison-Wesley; 1976.
- Voit EO: *Computational Analysis of Biochemical Systems. A practical guide for biochemists and Molecular Biologists.* Cambridge University Press; 2000.
- Heijnen JJ: **Approximative kinetic formats used in metabolic network modeling.** *Biotechnology and Bioengineering* 2005, **91**:534-545.
- Visser D, Heijnen JJ: **Dynamic simulation and metabolic redesign of a branched pathway using linlog kinetics.** *Metabolic Engineering* 2003, **5**:164-176.
- Wu L, Wang WM, van Winden WA, van Gulik WM, Heijnen JJ: **A new framework for the estimation of control parameters in metabolic pathways using lin-log kinetics.** *European Journal of Biochemistry* 2004, **271**:3348-3359.
- Visser D, Schmid JW, Mauch K, Reuss M, Heijnen JJ: **Optimal redesign of primary metabolism in Escherichia coli using linlog kinetics.** *Metabolic Engineering* 2004, **6**:378-390.
- Kacser H, Burns JA: **MOlecular democracy: who shares the controls?** *Biochem Soc Trans* 1979, **7**:1149-1160.
- Giersch C: **Determining Elasticities from Multiple Measurements of Steady-State Flux Rates and Metabolite Concentrations - Theory.** *Journal of Theoretical Biology* 1994, **169**:89-99.
- Giersch C, Cornish-Bowden A: **Extending double modulation: Combinatorial rules for identifying the modulations necessary for determining elasticities in metabolic pathways.** *Journal of Theoretical Biology* 1996, **182**:361-369.
- Hofmeyr JH, Cornish-Bowden A: **Co-response Analysis: A New Experimental Strategy for Metabolic Control Analysis.** *Journal of Theoretical Biology* 1996, **182**:371-380.
- Heijnen JJ, van Gulik WM, Shimizu H, Stephanopoulos G: **Metabolic flux control analysis of branch points: an improved approach to obtain flux control coefficients from large perturbation data.** *Metabolic Engineering* 2004, **6**:391-400.
- Ostergaard S, Olsson L, Nielsen J: **In vivo dynamics of galactose metabolism in Saccharomyces cerevisiae: Metabolic fluxes and metabolite levels.** *Biotechnology and Bioengineering* 2001, **73**:412-425.
- Theobald U, Mailinger W, Baltés M, Rizzi M, Reuss M: **In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae .1. Experimental observations.** *Biotechnology and Bioengineering* 1997, **55**:305-316.
- Visser D, van Zuylen GA, van Dam JC, Oudshoorn A, Eman MR, Ras C, van Gulik WM, Frank J, van Dedem GWK, Heijnen JJ: **Rapid sampling for analysis of in vivo kinetics using the BioScope: A system for continuous-pulse experiments.** *Biotechnology and Bioengineering* 2002, **79**:674-681.
- Visser D, van Zuylen GA, van Dam JC, Eman MR, Proll A, Ras C, Wu L, van Gulik WM, Heijnen JJ: **Analysis of in vivo kinetics of glycolysis in aerobic Saccharomyces cerevisiae by application of glucose and ethanol pulses.** *Biotechnology and Bioengineering* 2004, **88**:157-167.
- Rizzi M, Baltés M, Theobald U, Reuss M: **In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae .2. Mathematical model.** *Biotechnology and Bioengineering* 1997, **55**:592-608.
- Kresnowati MTAP, van Winden WA, Heijnen JJ: **Determination of elasticities, concentration and flux control coefficients from transient metabolite data using linlog kinetics.** *Metabolic Engineering* 2005, **7**:142-153.
- Galazzo JL, Bailey JE: **Fermentation Pathway Kinetics and Metabolic Flux Control in Suspended and Immobilized Saccharomyces-Cerevisiae.** *Enzyme and Microbial Technology* 1990, **12**:162-172.
- Theobald U, Mailinger W, Reuss M, Rizzi M: **In-Vivo Analysis of Glucose-Induced Fast Changes in Yeast Adenine-Nucleotide Pool Applying A Rapid Sampling Technique.** *Analytical Biochemistry* 1993, **214**:31-37.
- Wu L, Mashego MR, van Dam JC, Proell AM, Vinke JL, Ras C, van Winden WA, van Gulik WM, Heijnen JJ: **Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly C-13-labeled cell extracts as internal standards.** *Analytical Biochemistry* 2005, **336**:164-171.
- Liao JC, Delgado J: **Dynamic Metabolic Control-Theory - A Methodology for Investigating Metabolic-Regulation Using Transient Metabolic Data.** *Annals of the New York Academy of Sciences* 1992, **665**:27-38.

35. Ehlde M, Zacchi G: **Influence of experimental errors on the determination of flux control coefficients from transient metabolic concentrations.** *Biochemical Journal* 1996, **313**:721-727.
36. Wang L, Hatzimanikatis V: **Metabolic engineering under uncertainty. I: framework development.** *Metab Eng* 2006, **8**:133-141.
37. Wang L, Hatzimanikatis V: **Metabolic engineering under uncertainty--II: analysis of yeast metabolism.** *Metab Eng* 2006, **8**:142-159.
38. Wang L, Birol I, Hatzimanikatis V: **Metabolic control analysis under uncertainty: framework development and case studies.** *Biophys J* 2004, **87**:3750-3763.
39. Schomburg I, Chang A, Schomburg D: **BRENDA, enzyme data and metabolic information.** *Nucl Acids Res* 2002, **30**:47-49.
40. Mashego MR, van Gulik WM, Heijnen JJ: **Metabolome dynamic responses of *Saccharomyces cerevisiae* on simultaneous rapid perturbations in external electron acceptor and electron donor.** *FEMS Yeast Res* 2006, **accepted**.
41. Moles CG, Mendes P, Banga JR: **Parameter Estimation in Biochemical Pathways: A Comparison of Global Optimization Methods.** *Genome Res* 2003, **13**:2467-2474.
42. Voit EO, Almeida J: **Decoupling dynamical systems for pathway identification from metabolic profiles.** *Bioinformatics* 2004, **20**:1670-1681.
43. Visser D, Heijnen JJ: **The Mathematics of Metabolic Control Analysis revisited.** *Metabolic Engineering* 2002, **4**:114-123.
44. Cortassa S, Aon MA: **Metabolic Control Analysis of Glycolysis and Branching to Ethanol-Production in Chemostat Cultures of *Saccharomyces-Cerevisiae* Under Carbon, Nitrogen, Or Phosphate Limitations.** *Enzyme and Microbial Technology* 1994, **16**:761-770.
45. Cortassa S, Aon JC, Aon MA: **Fluxes of Carbon, Phosphorylation, and Redox Intermediates During Growth of *Saccharomyces-Cerevisiae* on Different Carbon-Sources.** *Biotechnology and Bioengineering* 1995, **47**:193-208.
46. Delgado J, Meruane J, Liao JC: **Experimental-Determination of Flux Control Distribution in Biochemical Systems - Invitro Model to Analyze Transient Metabolite Concentrations.** *Biotechnology and Bioengineering* 1993, **41**:1121-1128.
47. Schlosser PM, Riedy TG, Bailey JE: **Ethanol-Production in Bakers-Yeast - Application of Experimental Perturbation Techniques for Model Development and Resultant Changes in Flux Control Analysis.** *Biotechnology Progress* 1994, **10**:141-154.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

