

# Simplified modelling of metabolic pathways for flux prediction and optimization: lessons from an *in vitro* reconstruction of the upper part of glycolysis

Julie B. FIÉVET\*†, Christine DILLMANN\*, Gilles CURIEN† and Dominique DE VIENNE\*<sup>1</sup>

\*UMR de Génétique Végétale INRA (Institut National de la Recherche Agronomique)/UPS (Université Paris-Sud)/CNRS (Centre National de la Recherche Scientifique)/INAPG, Ferme du Moulon, 91190 Gif-sur-Yvette, France, and †Département Réponse et Dynamique Cellulaires, Laboratoire de Physiologie Cellulaire Végétale, CEA (Commissariat à l'Energie Atomique)/CNRS/Université Joseph Fourier/INRA, CEA-Grenoble, 38054 Grenoble, France

Explicit modelling of metabolic networks relies on well-known mathematical tools and specialized computer programs. However, identifying and estimating the values of the very numerous enzyme parameters inherent to the models remain a tedious and difficult task, and the rate equations of the reactions are usually not known in sufficient detail. A way to circumvent this problem is to use 'non-mechanistic' models, which may account for the behaviour of the systems with a limited number of parameters. Working on the first part of glycolysis reconstituted *in vitro*, we showed how to derive, from titration experiments, values of effective enzyme activity parameters that do not include explicitly any of the classical kinetic constants. With a maximum of only

two parameters per enzyme, this approach produced very good estimates for the flux values, and enabled us to determine the optimization conditions of the system, i.e. to calculate the set of enzyme concentrations that maximizes the flux. This fast and easy method should be valuable in the context of integrative biology or for metabolic engineering, where the challenge is to deal with the dramatic increase in the number of parameters when the systems become complex.

**Key words:** biochemical pathway, flux computation, glycolysis, hyperbolic fitting, optimal enzyme concentration, systemic enzyme parameter.

## INTRODUCTION

The 'omics' approaches are challenging biochemical practice. The complete genome sequence data and microarray technologies reveal frequent occurrences of gene families with complex patterns of gene regulations, quantitative proteomics give expression data for hundreds of enzymes and isoenzymes, and metabolomics techniques give access to thousands of major metabolites. The way all these systemic variables are affected by development, environmental conditions and/or genetic background gives invaluable information to investigate the nature, organization and control of complex metabolic networks. Modelling metabolic systems relies on classical theoretical methods, such as differential equations, and on powerful computer programs. With these tools, extensive efforts have been made to achieve a comprehensive dynamic description of some essential pathways, such as glycolysis, especially in yeast [1,2], *Trypanosoma brucei* [3], erythrocytes [4–6] and muscle [7]. However, such approaches, which require a long-term effort of numerous laboratories for simple pathways, have severe limitations for more complex systems. The main obstacle for biochemical modelling is usually not the availability of mathematical and informatics tools, but the determination of the parameter values. Single rate equations may include more than ten parameters {as arises in the equation for PFK (6-phosphofructokinase; EC 2.7.1.11) [2]}, and the number of parameters increases dramatically with the size and complexity of the systems. Except for the concentration of the most abundant enzymes, which can be estimated with quantitative proteomics (e.g. [8]), there is no simple or direct way to measure essen-

tial parameters such as catalytic constants, Michaelis–Menten constants, or inhibition constants. For example, the robot-based platform recently developed to measure activities of approx. 20 enzymes in *Arabidopsis* is promising, but does not provide values for such essential parameters [9]. *In vivo*, it is questionable whether all the parameters can be estimated from perturbation experiments [10], in particular because the range of variation of metabolite levels is limited by homeostasis constraints [11]. For some well-described pathways, computer simulations can be performed to estimate the parameter values by fitting the model with *in vivo* observations [12]. In any case, estimating the parameter values remains a tedious and time-consuming task, with multiple factors of inaccuracy, and no assurance that all the key parameters are identified.

Consequently, modelling efforts based on conceptual shortcuts are essential to simulate complex cellular behaviours from a limited amount of biological data. Flux balance analysis (see [13] for a review) represents a radical approach in this direction, since no kinetic information is included in the models. Flux balance analysis does not attempt to predict the exact behaviour of metabolic networks, but uses known constraints on the integrated function of multiple enzymes to separate the state that a system can reach from those it cannot. Constraints considered are the steady-state mass balance of metabolites, the effective irreversibility of reactions due to thermodynamic constraints, or the maximum flux capacity of enzymes or transport proteins. In the MCA (metabolic control analysis) [14,15] or in the BST (biochemical systems theory) [16], rate laws are not explicitly written, but kinetic parameters are not ignored. Both theoretical frameworks use systemic coefficients (control coefficients or

Abbreviations used: BST, biochemical systems theory; CK, creatine kinase; cv, coefficients of variation; FBA, aldolase; G3PDH, glycerol-3-phosphate dehydrogenase; HK, hexokinase; MCA, metabolic control analysis; PFK, 6-phosphofructokinase; PGI, phosphoglucosomerase; TPI, triose-phosphate isomerase.

<sup>1</sup> To whom correspondence should be addressed (email devienne@moulon.inra.fr).

logarithmic gains) and local coefficients (elasticity or kinetic order), which can be estimated experimentally, to quantify the behaviour of the systems. Derived from MCA or BST, various so-called 'non-mechanistic' models have been proposed that aim at using approximate equations to describe the behaviour of the systems with a reduced number of parameters. The power law representations, such as the S-systems, are obtained by specifying a rate constant (the multiplicative term) and writing one power term for each variable that directly influences the rate law (metabolites, enzymes and effectors) [17]. The (log)linear kinetic models rely on knowledge about the strength of interaction among the rates of enzyme-catalysed reactions and the concentrations of the various metabolites and regulators of the network [18]. The linlog kinetics proposes general expressions giving steady-state fluxes and metabolite levels as a function of enzyme concentration and response coefficients [11]. Finally, the 'tendency modelling' approach results in rate equations containing a minimal number of parameters, because it assumes that the behaviour of the metabolites is determined more by the stoichiometric structure of the network than by the exact enzyme mechanisms [19].

The present study is in the line of the modelling approaches using a limited number of parameters. We show how to use a simplified formalism based on MCA to derive global parameters that account for the kinetic behaviour of enzymes in their metabolic context. According to the structure of the pathway and to the position of the enzyme in the pathway, one or two parameters per enzyme are sufficient. We evaluated the validity of the method on a pathway reconstituted *in vitro*, an approach that just proved to be very efficient to study the systemic properties of the pathways, such as elasticities, flux control coefficients and transition time control coefficients [20–22], dynamic properties [23,24] and allosteric controls [25]. The upstream part of glycolysis, from HK (hexokinase; EC 2.7.1.1) to G3PDH (glycerol-3-phosphate dehydrogenase; EC 1.1.1.8), allowed us to measure the flux for a large number of different distributions of enzyme concentrations, and to compare the values with the theoretical predictions. We showed that the predicted flux values were tightly correlated to the measured flux values. In addition we derived equations to calculate the optimal distribution of enzyme concentrations for a given total enzyme content allocated to the network. Again the predicted values were fully consistent with observations. As the method does not require the estimation of any classical kinetic parameter such as  $K_m$ ,  $V_{max}$  or  $K_i$ , the experimental effort for modelling and determining the conditions for optimization is thus significantly reduced. Application of the approach to metabolic engineering and to *in vivo* measurements is discussed.

## THEORETICAL DEVELOPMENTS

### Linear pathway of Michaelis–Menten enzymes

The flux  $J$  through a linear pathway of  $n$  unimolecular reversible reactions catalysed by Michaelis–Menten enzymes far from saturation is [14,15]:

$$J = \frac{S_1 - \frac{S_{n+1}}{K_{1,n+1}}}{\sum_{j=1}^n \frac{K_{mj}}{V_j K_{1,j}}} \quad (1)$$

where  $S_1$  and  $S_{n+1}$  are respectively the concentrations of the initial and final substrates of the pathway,  $K_{1,n+1}$  is the product of the equilibrium constants of all the reactions,  $K_{1,j}$  is the product of the equilibrium constants from reaction 1 to reaction  $S_{j-1} \rightarrow S_j$ ,

$K_{mj}$  is the Michaelis–Menten constant of enzyme  $j$  and  $V_j$  is the maximum velocity of enzyme  $j$ . Since the maximum velocity is given as  $V_j = E_j k_{catj}$ , with  $E_j$  the concentration of enzyme  $j$  and  $k_{catj}$  its catalytic constant, we have:

$$J = \frac{S}{\sum_{j=1}^n \frac{K_{mj}}{k_{cat} K_{1,j} E_j}} \quad (2)$$

where  $S = [S_1] - [S_{n+1}]/K_{1,n+1}$ .

For a given total enzyme concentration,

$$\sum_{j=1}^n E_j = E_{tot}$$

there is an optimal set of enzyme concentrations  $\{E_1^*, E_2^*, \dots, E_n^*\}$ , that is, a distribution that maximizes the flux. At maximum flux, we have for all enzyme  $i$  [26,27]:

$$\frac{E_i^*}{E_{tot}} = \frac{1/\sqrt{A_i}}{\sum_{j=1}^n 1/\sqrt{A_j}} \quad (3)$$

where  $A_j = k_{catj} K_{1,j}/K_{mj}$ .

### Estimating systemic kinetic parameters

Equations (2) and (3) do not apply to pathways such as glycolysis that contain non-Michaelian–Menten and regulated enzymes. Nevertheless, we wanted to evaluate the extent to which a similar formalism could help in computing estimates of flux values and determining the conditions for optimizing metabolic systems of any complexity.

Consider a system converting a metabolite  $S_1$  into  $S_{m+1}$  through a series of reactions catalysed by  $n$  reversible enzymes far from saturation. As branching is possible,  $n \geq m$ . With a single input ( $S_1$ ) and output ( $S_{m+1}$ ), the flux through the whole pathway at steady state is the rate of consumption of  $S_1$ , which is equal to the rate of formation of  $S_{m+1}$ . On increasing the concentration  $E_i$  of enzyme  $i$ , the concentrations of other enzymes being fixed, the response of the flux is expected to display a quasi-hyperbolic ascending curve (saturation curve). This intuitive prediction is corroborated by an abundant theoretical and experimental literature (see the Discussion section). However, depending on the topology, the regulations and the constraints in the network, various mechanisms could result – at least theoretically – in a decrease in the flux when the concentration of certain enzymes increases (some examples are presented in the Discussion section). Moreover, non-linear behaviour, such as oscillations, can prevent a stable steady state being observed.

In the following developments, we considered only stable steady states, but for more generality we considered that the asymptotic response of the flux upon enzyme increase may be ascending or descending.

For fitting to experimental curves, we chose an expression of the flux  $J$  derived from eqn (2):

$$J = \frac{S}{\sum_{j=1}^n \frac{1}{A_j E_j + p_j}} = \frac{1}{\sum_{j=1}^n \frac{1}{SA_j E_j + Sp_j}} \quad (4)$$

where  $S$  and  $E_j$  have the same definition as above. The parameter  $A_j$  is a complex systemic kinetic parameter that depends on the type of enzyme and regulation, and takes the above value

( $A_j = k_{catj}K_{1,j}/K_{mj}$ ) only for Michaelis–Menten non-regulated enzymes in a linear pathway. Note that  $A_j$  may be negative, for example if the enzyme belongs to the backward branch of a substrate cycle. The dimension of  $A_j$  is  $\mu\text{M}^{-1} \cdot \text{s}^{-1}$ . The parameter  $p_j$  accounts for the ‘dispensability’ of enzyme  $j$ : if  $p_j = 0$ , then  $J = 0$  when  $E_j = 0$ , and if  $|p_j| \neq 0$ , then  $J \neq 0$  when  $E_j = 0$ . For example, in the system considered (see Scheme 1 in the Experimental section), removing the TPI (triose-phosphate isomerase; EC 5.3.1.1) does not bring the flux to zero because of the branching of the pathway at the reaction catalysed by the FBA (aldolase; EC 4.1.2.13). So the  $p_{\text{TPI}}$  value is not null, unlike those of the other enzymes. The dimension of  $p_j$  is  $\text{s}^{-1}$ .

All those parameters can be estimated with titration experiments, by varying in turn the concentration of each enzyme, the other concentrations being kept constant. For enzyme  $i$ , we have the following expression of the flux:

$$J = \frac{1}{\frac{1}{SA_i E_i + Sp_i} + \sum_{j \neq i}^n \frac{1}{SA_j E_j + Sp_j}} \quad (5)$$

When  $E_i$  tends to infinity, the flux tends to the value:

$$J_{\infty i} = \frac{1}{\sum_{j \neq i}^n \frac{1}{SA_j E_j + Sp_j}} \quad (6)$$

Thus the flux may be given as:

$$J = \frac{1}{\frac{1}{SA_i E_i + Sp_i} + \frac{1}{J_{\infty i}}} \quad (7)$$

The three parameters  $SA_i$ ,  $Sp_i$  and  $J_{\infty i}$  of this function of  $E_i$  can be estimated by hyperbolic fitting of the titration curves [we used the *nls* package of the version 2.0.1 of the R software (<http://www.r-project.org/>)]. Note that the ordinate of the intercept for  $E_i = 0$ , when  $Sp_i \neq 0$ , is:

$$J_{0i} = \frac{1}{\frac{1}{Sp_i} + \frac{1}{J_{\infty i}}} \quad (8)$$

Hence:

$$Sp_i = \frac{J_{0i} \cdot J_{\infty i}}{J_{\infty i} - J_{0i}} \quad (9)$$

For indispensable enzymes, we have  $Sp_i = 0$ , and the flux value tends to 0 when  $E_i$  tends to 0. With all the  $SA_i$  and  $Sp_i$  estimates, the flux can be computed from eqn (4) for any set of  $E_i$  values.

### Determination of the optimal distribution of enzyme concentrations

For a given total enzyme concentration, the optimal distribution of the concentrations  $\{E_1^*, E_2^*, \dots, E_n^*\}$  for a flux expressed by eqn (4) can be calculated using the Lagrange multiplier method (see the Appendix). Writing  $e_i^* = E_i^*/E_{\text{tot}}$ , we get:

$$e_i^* = \frac{1/\sqrt{A_i}}{\sum_j 1/\sqrt{A_j}} \left( 1 + \sum_j \frac{p_j}{A_j E_{\text{tot}}} \right) - \frac{p_i}{A_i E_{\text{tot}}} \quad (10)$$

This equation is valid only if  $E_{\text{tot}}$  corresponds to a fixed total molarity. If  $E_{\text{tot}}$  is a fixed total weight of enzymes (e.g. in g/l), as considered in the present study, the molecular masses of enzymes

have to be included in the optimization equation, as shown in the Appendix. We get:

$$g e_i^* = \frac{1/\sqrt{A_i}}{\sum_j 1/\sqrt{A_j}} \left( 1 + \sum_j \frac{p_j}{A_{j,g} E_{\text{tot}}} \right) - \frac{p_i}{A_{i,g} E_{\text{tot}}} \quad (11)$$

where  $A'_j = A_j/M_j$ , with  $M_j$  the molecular mass of enzyme  $j$ , and where the left subscript ‘g’ means that the enzyme concentrations are in g/l.

## EXPERIMENTAL

### Chemicals

D-Glucose, Pipes, ATP, NADH, phosphocreatine and all enzymes were obtained from Sigma. HK, PGI (phosphoglucosomerase; EC 5.3.1.9) and TPI were purified from baker’s yeast. FBA, G3PDH and CK (creatine kinase; EC 2.7.3.2) were purified from rabbit muscle, and PFK was from *Bacillus stearothermophilus*. All the purchased commercial enzymes were crystalline suspensions in ammonium sulphate. These proteins were desalted by using PD-10 desalting columns according to the manufacturer’s recommendations (Amersham Biosciences). The protein concentrations were determined at  $\lambda = 205 \text{ nm}$  [28]. The molecular masses of the enzymes are: PGI, 61 300 Da; PFK, 34 000 Da; FBA, 39 212 Da; and TPI, 26 795 Da.

### Measurements of kinetic parameters of isolated enzymes

All enzyme assays were carried out in a thermoregulated quartz cuvette (25°C), in a buffer containing 50 mM Pipes (pH 7.5), 100 mM KCl and 5 mM magnesium acetate. HK activity was measured with 1  $\mu\text{M}$  lactate dehydrogenase, 1  $\mu\text{M}$  pyruvate kinase, 1 mM phosphoenolpyruvate and various concentrations of glucose. PGI activity was measured in the forward direction with 0.5  $\mu\text{M}$  PFK, 5  $\mu\text{M}$  FBA, 3  $\mu\text{M}$  G3PDH and glucose 6-phosphate as substrate. PFK activity was measured with 0.25  $\mu\text{M}$  FBA, 0.41  $\mu\text{M}$  TPI, 0.9  $\mu\text{M}$  G3PDH and various concentrations of fructose 6-phosphate. FBA activity was measured with 0.5 mM TPI and 1 mM G3PDH and fructose 1,6-bisphosphate as substrate. TPI activity was assayed with 1  $\mu\text{M}$  G3PDH and glyceraldehyde 3-phosphate as substrate.

For estimating the kinetic parameters of isolated enzymes, we used the classical Eadie–Hofstee representation [29] where the maximal velocity  $V$  and the Michaelis–Menten constant  $K_m$  were estimated by fitting the linear regression model

$$Y_i = \alpha - \beta X_i + \varepsilon_i$$

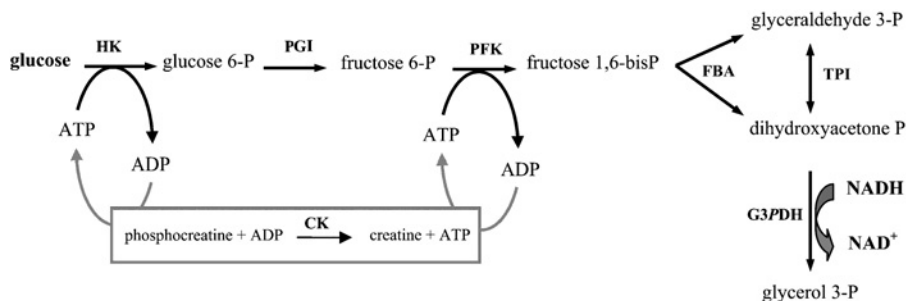
where  $Y_i = v_i/S_i$  are experimental data points for the reaction rate  $v_i$  at substrate concentration  $S_i$ ;  $X_i = v_i$ ; and  $\varepsilon_i$  is a random error term. In this representation,  $K_m = 1/\beta$  and  $V = \alpha/\beta$ . Unbiased estimators  $\hat{K}_m$  and  $\hat{V}$ , and their corresponding sampling variance, can be found by a Taylor series expansion of the ratios around the true values  $\alpha$  and  $\beta$  (see for example [30]):

$$\hat{K}_m = \frac{1}{\hat{\beta}} - \frac{s_{\hat{\beta}}^2}{\hat{\beta}^3} \quad \text{and} \quad \text{Var}(\hat{K}_m) = \frac{s_{\hat{\beta}}^2}{\hat{\beta}^4}$$

$$\hat{V} = \frac{\hat{\alpha}}{\hat{\beta}} - \frac{s_{\hat{\beta}}^2}{\hat{\beta}^3} + \frac{\text{Cov}(\hat{\alpha}, \hat{\beta})}{\hat{\beta}^2}$$

$$\text{and} \quad \text{Var}(\hat{V}) = \frac{s_{\hat{\alpha}}^2}{\hat{\beta}^2} - \frac{s_{\hat{\beta}}^2}{\hat{\beta}^4} - 2 \frac{\text{Cov}(\hat{\alpha}, \hat{\beta})}{\hat{\beta}^3}$$

where  $\hat{\alpha}$ ,  $\hat{\beta}$ ,  $s_{\hat{\alpha}}^2$ ,  $s_{\hat{\beta}}^2$  and  $\text{Cov}(\hat{\alpha}, \hat{\beta})$  are estimated from the linear regression of  $v_i/S_i$  on  $v_i$ .



**Scheme 1** Upstream part of glycolysis reconstructed *in vitro*

**Table 1** Kinetic parameters of the enzymes

Maximal rates ( $V$ ) are expressed as units (i.e.  $\mu\text{mol}/\text{min}$ ) per mg of total protein. Because  $V$  and  $K_m$  are estimated from ratios in the Eadie–Hofstee representation, unbiased estimators and S.D. values of these parameters were calculated by Taylor series expansions (see the Materials and methods section).

Enzyme	Experimental determination			Manufacturer values		
	$V \pm \text{S.D.}$ (units/mg of protein)	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_m \pm \text{S.D.}$ (mM)	$V$ (units/mg of protein)	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_m$ (mM)
PGI	$1370 \pm 60$	1410	$1.65 \pm 0.15$	504	514	1.4
PFK	$70 \pm 1$	41.7	$0.033 \pm 0.001$	164	150	0.029
FBA	$42 \pm 2$	27.7	$0.012 \pm 0.002$	14	9.3	0.05
TPI	$14690 \pm 1900$	6680	$2.38 \pm 0.45$	9500	4220	1.23

### *In vitro* pathway and flux measurement

The system constructed *in vitro* was operated following Scheme 1.

The ATP pool is kept constant via the regeneration reaction catalysed by CK. HK is saturated due to the high glucose concentration (100 mM). The HK concentration was fixed in all experiments at 0.1  $\mu\text{M}$  so that the glucose 6-phosphate input is the same in all the experiments. G3PDH, catalysing the oxidation of NADH, was kept constant, at high concentration (1  $\mu\text{M}$ ), in order to use the slope of the linear NADH decay as a measurement of the steady-state flux through the whole pathway. NADH consumption was monitored every 2 s with an Uvikon 850 spectrophotometer at 390 nm from 60 to 120 s. The assays were carried out at 25 °C in a buffer containing 50 mM Pipes (pH 7.5), 100 mM KCl, 20 mM phosphocreatine, 3 mM NADH and 5 mM magnesium acetate. Each reaction was started by addition of 1 mM ATP.

### Titration curves

We varied in turn the concentrations of each enzyme, the concentration of the others being maintained at reference values computed on the basis of their physiological concentration estimated in the yeast strain S288C [8]. The reference values are 0.15  $\mu\text{M}$  for PGI, 0.29  $\mu\text{M}$  for PFK, 1.54  $\mu\text{M}$  for FBA and 0.84  $\mu\text{M}$  for TPI (respectively 9.1, 10.4, 60.1 and 22.3 mg/l). Hyperbolic fitting of the curves gave the relevant enzyme parameters that can be used for computing the predicted flux values (see the Theoretical developments subsection).

### Selecting distributions of enzyme concentrations for flux measurement

To assess the reliability of the predicted flux values, and the validity of the model giving the optimal enzyme concentrations, we measured the flux values for a series of 121 distributions of enzyme concentrations, the total amount being kept constant. The sum of PGI, PFK, FBA and TPI concentrations was fixed at  $E_{\text{tot}} = 101.9 \text{ mg/l}$ , chosen as the sum of the reference concentrations. In order to cover a large range of variation, we varied

the FBA concentration (the most abundant enzyme) from 0 to  $E_{\text{tot}}$ , taking ten values evenly distributed, and the proportions of the remaining enzymes (PGI, PFK and TPI) were drawn in  $\beta$  distributions with shape parameter  $\alpha = 1$  and scale parameter

$$\beta = \frac{1 - e_i^{\circ}}{e_i^{\circ}}$$

with  $e_i^{\circ}$  the reference proportion of enzyme  $i$ . In the 121 distributions obtained, enzyme concentrations ranged from 0 to 70 mg/l for PGI and PFK and from 1.66 to 66.9 mg/l for TPI.

Each flux measurement was repeated three times in three independent tubes. As the fluxes should be expressed in  $\mu\text{M}/\text{s}$ , the enzyme concentrations (in mg/l) were converted into  $\mu\text{M}$  using the molecular masses given by the manufacturer.

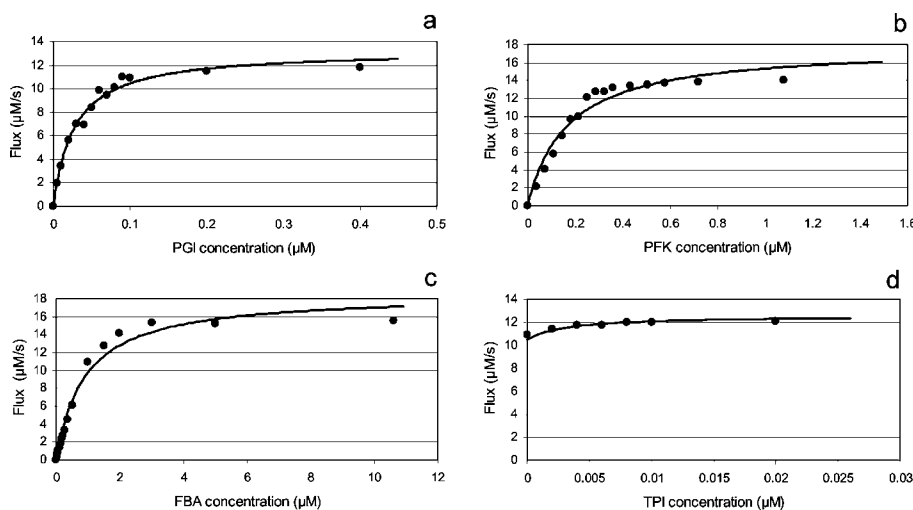
## RESULTS

### Estimation of the kinetic parameters of isolated enzymes

The kinetics of the commercial enzymes in the forward direction was studied under the same conditions as those used for the system constructed *in vitro*. The values of the enzyme kinetic parameters estimated using the Eadie–Hofstee representation were of the same order of magnitude as those indicated by the manufacturer (Table 1).

### Titration experiments in the pathway constructed *in vitro*

Relationships between the flux and the concentration of each enzyme of the upstream part of glycolysis constructed *in vitro*, and the curves of the hyperbolic fittings, are shown in Figure 1. For each enzyme, the parameters  $J_{\infty i}$ ,  $SA_i$  and  $Sp_i$  were estimated as indicated in the Theoretical developments subsection (Table 2). As expected,  $SA_i$  values were poorly correlated with the ratio  $k_{\text{cat}i}K_{1,i}/K_{mi}$  that can be estimated from isolated enzymes (results not shown). TPI alone is ‘dispensable’, since the flux with no TPI is estimated to be  $J_{0\text{TPI}} = 10.47 \mu\text{M}/\text{s}$ , whereas  $J_{0i}$  for other



**Figure 1** Titration curves

(a) PGI, (b) PFK, (c) FBA and (d) TPI. The flux ( $y$ -axis) is in  $\mu\text{M/s}$  and the concentration of the enzymes ( $x$ -axis) is in  $\mu\text{M}$ .

**Table 2** Systemic parameter values of the four enzymes estimated from titration curves

$J_\infty$ ,  $J_0$ ,  $SA$  and  $Sp$ , see the text.

Enzyme	$J_\infty \pm \text{S.D.}$ ( $\mu\text{M/s}$ )	$J_0 \pm \text{S.D.}$ ( $\mu\text{M/s}$ )	$SA \pm \text{S.D.}$ ( $\text{s}^{-1}$ )	$Sp \pm \text{S.D.}$ ( $\mu\text{M/s}$ )
PGI	$13.22 \pm 0.5$	0	$499.4 \pm 49.6$	0
PFK	$17.61 \pm 1.0$	0	$115.5 \pm 15.1$	0
FBA	$18.24 \pm 0.7$	0	$22.5 \pm 2.4$	0
TPI	$12.60 \pm 0.1$	$10.5 \pm 0.1$	$22\,940 \pm 430$	$61.8 \pm 4.5$

enzymes is zero. So we have  $Sp_{\text{TPI}} = 61.77 \mu\text{M/s}$ , and  $Sp_i = 0$  for PGI, PFK and FBA.

### Comparing predicted and measured flux values

With the estimates of the  $SA_i$  and  $Sp_i$ , we calculated a predicted flux value,  $J_{\text{pred}}$ , for any set of enzyme concentrations from eqn (4):

$$J_{\text{pred}} = \frac{1}{\frac{1}{499.4 \times E_{\text{PGI}}} + \frac{1}{115.5 \times E_{\text{PFK}}} + \frac{1}{22.5 \times E_{\text{FBA}}} + \frac{1}{22\,940 \times E_{\text{TPI}}} + 61.8}$$

This predictor was used to compute the flux values for 121 distributions of enzyme concentrations, and these values were compared with the flux values measured *in vitro*,  $J_{\text{obs}}$ . Table 3 gives the  $J_{\text{pred}}$  and the  $J_{\text{obs}}$  values and the S.D. of the  $J_{\text{obs}}$ . The cv (coefficients of variation) never exceeded 0.14, and 84% of the cv were below 0.05. The linear correlation coefficient between the predicted and measured flux value was high ( $r = 0.93$ ) and the intercept of the regression line was not statistically different from zero (Figure 2). The slope of the regression line was statistically higher than 1 ( $a = 1.38 \pm 0.18$ ), which means that the predictor overestimates the observed flux by a constant factor. Thus  $J_{\text{pred}}/a$  appears to be a reliable predictor of the flux in the *in vitro* system.

### Optimizing the metabolic system

As the total enzyme concentration was fixed, the relationship between the concentration of a particular enzyme and the flux can no longer be hyperbolic, since there is a global negative correlation between the concentrations. High concentration of one enzyme

reduces that of the others, which may decrease the flux. Consistent with this prediction, plotting the observed flux as a function of the concentration of any one enzyme showed that the experimental points were distributed below a hump-shaped curve (Figure 3). For a given concentration of one of the enzymes, the flux may vary between 0 and a local maximum, depending on the distribution of the concentrations among the other enzymes. The distributions were randomly drawn, which explains the high extent of scatter of the points (some enzyme concentrations may be very low by chance).

The optimal proportion of each enzyme in the system was computed using eqn (11), which gave the proportions 0.155 for PGI, 0.240 for PFK, 0.590 for FBA and 0.015 for TPI. As the total enzyme concentration,  $E_{\text{tot}}$ , was 101.9 mg/l, these values corresponded respectively to 15.8, 24.5, 60.1 and 1.5 mg/l. As expected, the flux value computed with these concentrations ( $J_{\text{pred max}}/a = 14.62 \mu\text{M/s}$ ) is higher than all the predicted or measured flux values, and the position of the optimal concentration for each enzyme corresponds to the top of the hump-shaped curve that 'envelops' all the points (Figure 3). For an enzyme concentration lower than the optimum, the local maximum flux increases with the enzyme concentration. For an enzyme concentration that exceeds the optimum, the local maximum flux decreases with the enzyme concentration.

### DISCUSSION

We used the reconstituted upstream part of glycolysis to evaluate a method for estimating *effective* enzyme parameters that account for the behaviour of the enzyme within its metabolic pathway. The flux values predicted from these parameters proved to be quite close to the measured flux values. The systemic parameter  $A_i$  that we defined includes all the kinetic and regulation properties of the enzyme, but its estimation does not require any measurement of parameters such as  $K_m$ ,  $k_{\text{cat}}$  or  $K_i$ . Thus the kinetic type of the enzyme does not interfere with the modelling. For instance, PFK, which is a regulated enzyme, was treated like PGI, a Michaelian enzyme. It is worth noting that  $A_i$  also contains information on the position of the catalysed reaction in the pathway, via the equilibrium constants. In linear pathways of Michaelis–Menten enzymes far from saturation,

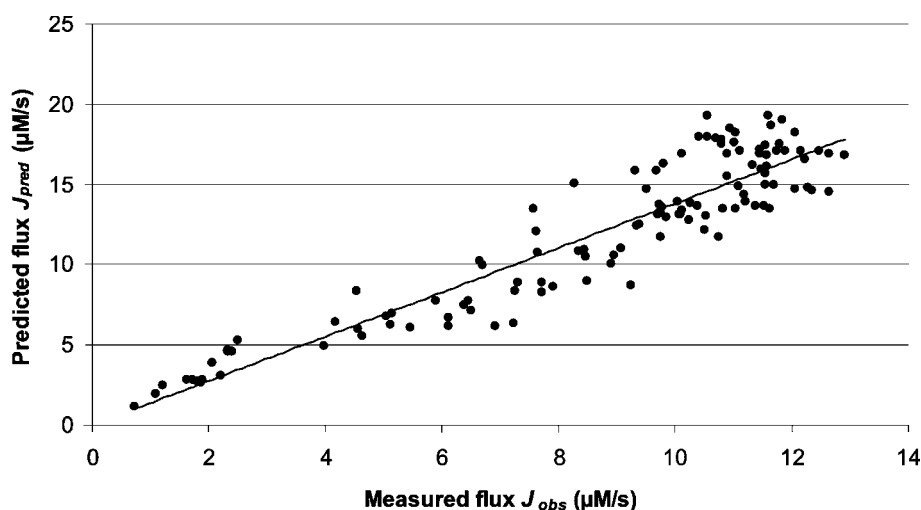
**Table 3 Predicted ( $J_{\text{pred}}$ ) and measured flux ( $J_{\text{obs}}$ ) for the 121 genotypes**

The enzyme concentrations are indicated in their respective columns. The last column (S.D.) indicates the standard deviation of the mean  $J_{\text{obs}}$  values computed with three replicates.

	PGI	PFK	FBA	TPI	$J_{\text{pred}}$	$J_{\text{obs}}$	S.D.		PGI	PFK	FBA	TPI	$J_{\text{pred}}$	$J_{\text{obs}}$	S.D.
1	39.32	12.38	25.47	24.71	10.49	8.46	0.11	62	20.00	42.50	28.50	10.90	13.45	7.57	0.65
2	51.97	14.22	25.47	10.23	10.92	8.45	0.23	63	35.00	32.50	8.50	25.90	4.59	2.39	0.21
3	12.20	27.64	35.66	26.39	14.36	11.19	0.22	64	20.00	45.00	32.50	4.40	15.02	8.28	0.33
4	32.45	28.74	35.66	5.03	15.84	9.68	0.14	65	55.00	7.50	22.50	16.90	8.40	4.53	0.65
5	15.66	23.52	45.85	16.86	17.11	12.15	0.22	66	11.50	25.00	47.50	17.90	16.89	10.11	0.34
6	15.55	30.53	45.85	9.96	17.96	10.40	0.22	67	22.50	30.00	8.50	40.90	4.54	2.34	0.10
7	6.79	20.43	56.04	18.63	15.71	11.55	0.16	68	25.00	27.50	10.00	39.40	5.26	2.49	0.07
8	3.33	5.75	66.23	26.58	8.74	9.24	0.04	69	45.00	37.50	8.50	10.90	4.63	2.32	0.06
9	9.01	8.80	66.23	17.85	13.61	11.39	0.24	70	3.98	2.28	81.52	14.12	5.52	4.62	0.06
10	4.24	25.75	66.23	5.67	14.94	11.70	0.30	71	3.81	2.71	81.52	13.85	6.17	6.12	0.12
11	8.44	9.53	66.23	17.69	13.93	11.22	0.10	72	6.40	2.69	86.61	6.19	6.72	6.11	0.15
12	5.92	5.25	76.42	14.30	10.03	8.90	0.06	73	4.75	2.63	81.52	13.00	6.27	5.13	0.19
13	5.79	3.30	76.42	16.38	7.51	6.38	0.29	74	5.38	3.01	86.61	6.90	7.09	6.49	0.09
14	4.23	2.62	76.42	18.62	6.09	5.46	0.10	75	6.37	4.17	76.42	14.94	8.86	7.71	0.34
15	9.40	2.58	86.61	3.31	6.77	5.05	0.13	76	4.66	4.03	81.52	11.69	8.27	7.71	0.11
16	3.72	1.95	86.61	9.61	4.90	3.99	0.13	77	24.36	7.48	56.04	14.01	13.23	9.73	0.11
17	7.36	3.21	86.61	4.72	7.77	6.47	0.08	78	28.10	8.42	50.95	14.43	13.58	9.77	0.13
18	37.90	12.60	35.66	15.73	13.24	9.74	0.05	79	8.76	5.41	66.23	21.49	10.55	8.96	0.08
19	29.71	6.62	56.04	9.53	12.52	9.39	0.22	80	23.65	7.59	61.14	9.52	13.77	9.72	0.18
20	13.29	8.21	56.04	24.36	13.11	10.08	0.05	81	6.82	14.17	76.42	4.49	16.17	11.33	0.38
21	20.60	7.18	66.23	7.89	13.61	11.52	0.08	82	5.26	11.19	71.33	14.12	13.48	11.03	0.16
22	9.88	3.73	86.61	1.67	8.92	7.31	0.04	83	9.64	16.24	66.23	9.79	17.45	11.56	0.23
23	3.40	2.81	86.61	9.08	6.20	6.92	0.24	84	4.56	4.53	71.33	21.48	8.59	7.92	0.11
24	70.00	5.00	5.00	21.90	2.44	1.22	0.08	85	8.62	3.47	86.61	3.19	8.36	7.25	0.11
25	20.00	10.00	5.00	66.90	2.60	1.87	0.04	86	5.35	4.48	76.42	15.65	8.97	8.50	0.09
26	40.00	10.00	40.00	11.90	13.13	9.70	0.55	87	22.56	7.84	50.95	20.55	12.94	9.86	0.05
27	55.00	15.00	12.00	19.90	5.98	4.56	0.06	88	23.34	7.80	56.04	14.72	13.48	11.63	0.14
28	15.00	15.00	55.00	16.90	16.79	12.90	0.53	89	27.65	9.99	45.85	18.40	13.89	10.05	0.09
29	10.00	20.00	15.00	56.90	6.98	5.15	0.26	90	17.75	4.96	66.23	12.95	10.79	8.36	0.15
30	40.00	35.00	5.00	21.90	2.78	1.72	0.02	91	18.53	4.91	71.33	7.12	10.97	9.08	0.35
31	45.00	40.00	12.00	4.90	6.43	4.18	0.22	92	8.31	6.98	61.14	25.47	11.69	10.74	0.23
32	15.00	40.00	45.00	1.90	18.22	12.07	0.81	93	15.24	5.45	76.42	4.78	11.75	9.76	0.03
33	40.00	45.00	7.00	9.90	3.86	2.07	0.12	94	11.04	15.68	61.14	14.03	17.10	12.47	0.17
34	3.00	45.00	50.00	3.90	12.10	10.50	0.35	95	7.06	14.74	76.42	3.67	16.53	12.23	0.13
35	15.00	50.00	5.00	31.90	2.76	1.79	0.00	96	19.12	16.02	56.04	10.71	17.85	10.71	0.19
36	25.00	50.00	12.00	14.90	6.40	4.18	0.15	97	11.51	13.37	66.23	10.79	16.91	12.63	0.15
37	25.00	50.00	20.00	6.90	10.19	6.64	0.10	98	22.62	8.82	50.95	19.51	13.68	10.37	0.08
38	35.00	60.00	5.00	1.90	2.80	1.89	0.01	99	45.65	13.30	25.47	17.47	10.72	7.65	0.32
39	15.00	65.00	20.00	1.90	9.95	6.69	0.11	100	30.21	11.88	45.85	13.96	14.93	11.54	0.07
40	25.00	70.00	2.00	4.90	1.14	0.74	0.08	101	9.06	16.44	56.04	20.35	15.97	11.49	0.10
41	33.00	1.00	66.23	1.66	3.08	2.20	0.06	102	5.08	15.50	71.33	9.99	14.77	12.28	0.13
42	1.00	33.00	66.23	1.66	6.30	7.23	0.01	103	19.19	17.00	56.04	9.67	18.20	11.05	0.29
43	4.00	25.00	66.23	6.66	14.50	12.65	0.21	104	33.82	18.87	35.66	13.55	14.66	9.50	0.18
44	7.00	12.00	66.23	16.66	14.61	12.35	0.21	105	21.50	7.41	56.04	16.94	13.05	10.52	0.07
45	10.00	10.00	66.23	15.66	14.68	12.05	0.07	106	33.81	9.61	45.85	12.63	13.82	10.26	0.03
46	13.00	13.00	66.23	9.66	17.09	11.90	0.14	107	11.15	8.50	61.14	21.11	13.48	10.82	0.19
47	16.00	7.00	66.23	12.66	13.13	10.10	0.29	108	23.46	10.70	50.95	16.79	14.92	11.08	0.25
48	19.00	7.00	66.23	9.66	13.34	10.11	0.27	109	16.40	17.41	50.95	17.14	17.04	11.75	0.10
49	22.00	12.00	66.23	1.66	17.51	10.80	n.d.	110	10.61	18.22	50.95	22.12	16.13	11.57	0.29
50	25.00	9.00	66.23	1.66	15.48	10.90	0.14	111	22.35	16.51	45.85	17.18	16.29	9.80	0.27
51	28.00	6.00	66.23	1.66	12.43	9.35	0.46	112	18.13	15.35	56.04	12.38	17.50	11.79	0.08
52	31.00	3.00	66.23	1.66	7.75	5.90	0.03	113	18.07	18.85	56.04	8.93	18.64	11.64	0.05
53	6.00	18.00	66.23	11.66	15.81	9.31	0.10	114	12.34	16.16	56.04	17.36	16.85	11.58	0.06
54	8.00	8.00	66.23	19.66	12.73	10.25	0.07	115	20.45	19.14	50.95	11.36	17.95	10.56	0.42
55	10.00	15.00	66.23	10.66	17.14	11.45	0.49	116	12.00	20.03	56.04	13.82	17.82	10.79	0.24
56	12.00	13.00	66.23	10.66	16.86	11.45	0.21	117	18.35	27.25	50.95	5.35	19.26	10.55	0.29
57	14.00	18.00	66.23	3.66	19.32	11.60	0.00	118	9.90	28.14	56.04	7.82	18.48	10.95	0.26
58	16.00	16.00	66.23	3.66	18.97	11.85	0.21	119	24.06	26.13	40.76	10.95	16.88	10.88	0.30
59	47.50	37.50	3.50	13.40	1.97	1.10	0.03	120	11.23	21.97	50.95	17.75	17.06	11.11	0.07
60	30.00	25.00	26.00	20.90	12.05	7.62	0.15	121	11.17	25.48	50.95	14.30	17.59	11.01	0.16
61	37.50	47.50	5.00	11.90	2.79	1.62	0.05								

$A_i$  includes merely the product of the equilibrium constants from the initial reaction to the reaction catalysed by enzyme  $i$  [14], but in the general case of complex networks there is no simple relationship between the equilibrium constants. The other parameter that we defined is the 'dispensability' ( $p_i$ ), which is

different from zero whenever removing the enzyme from the system does not drive the flux to 0, a situation that occurs in the case of reticulation in the system. The parameters  $A_i$  and  $p_i$  could not be estimated independently of  $S$ , a parameter that depends also on the equilibrium constants of the reactions and on the initial and



**Figure 2** Relationship between the measured flux  $J_{obs}$  and the predicted flux  $J_{pred}$

The regression line  $J_{pred}$  on  $J_{obs}$  is indicated.

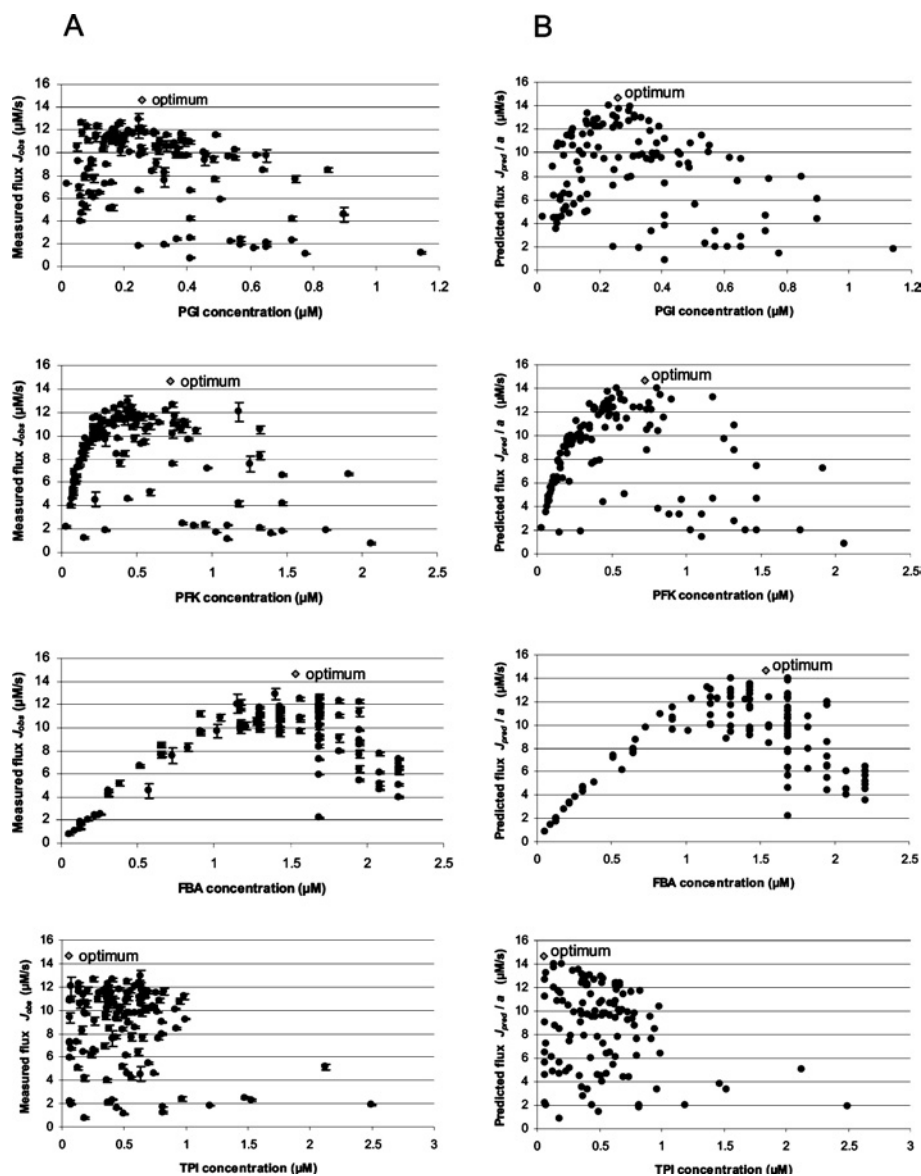
final substrates of the pathway, but this had no incidence on the calculus of the predicted flux.

The visual inspection of the titration curves used to estimate these parameters revealed that for every enzyme, the experimental points follow a slightly more convex curve than the fitted hyperbola, so that the estimated  $J_{oci}$  are higher than the apparent maximum experimental flux values. Thus the slope  $a$  of the regression line of  $J_{pred}$  on  $J_{obs}$  is higher than 1, so the right predictor of flux is  $J_{pred}/a$ . Other, more complex, functions could lead to better fittings. However, the quality of the prediction that we obtained with only two parameters does not justify making the model more complex. In addition, we think that the prediction could be even improved in a system in which the concentration of the first substrate could be maintained constant. Under our conditions, the first substrate of the system was glucose 6-phosphate; its input was constant because HK was saturated by glucose and did not vary, but its steady-state concentration depended on the concentration of the other enzymes. As a consequence, unlike what was assumed in the model, the parameter  $S$  could slightly vary, which inevitably increased the dispersion of the values.

The method for estimating global parameters can apply whether flux response towards enzyme variation is positive or negative, i.e. whether enzymes have positive or negative flux control coefficients. Here, the flux displayed a saturation curve for all enzymes, which means that the enzymes have a positive control coefficient for any concentration. This situation has been described for diverse pathways in many classical papers (e.g. [31–33]), and is theoretically expected in case of linear pathways [14] even if there are feedback or feedforward loops [34]. In this connection, it would be interesting to test the robustness of the method on the complete glycolysis, which would introduce strong feedback loops. In branched pathways, there are some particular situations where increasing an enzyme concentration could result in a decrease in the flux. (i) When the enzyme under study is located inside a feedback loop, negative control coefficients may be observed, depending on the parameter values of the network [34]. This is consistent with observations showing that increasing the trehalose-6-phosphate synthase may result in a decrease in trehalose synthesis, via the inhibition of HK by trehalose 6-phosphate [35]. (ii) When there is a substrate cycle (or 'futile' cycle), i.e. when a reaction, or a set of reactions, that converts metabolite  $S_i$  into  $S_j$ , is opposed by a second set that reconverts

$S_j$  into  $S_i$ , as observed for example with the cycle glucose–glucose 6-phosphate. Overexpression of glucose-6-phosphatase in rat hepatocytes decreases glucose 6-phosphate levels, with a negative control coefficient of this enzyme on glycogen synthesis [36]. In such cycles, increasing concentration of a backward enzyme is expected to decrease the whole flux to a minimal level, whereas a maximum flux value will be observed if the concentration of any one backward enzyme is null. (iii) When there is regulation by covalent modifications of enzymes, negative control coefficient may be observed. In the classical example of reversible phosphorylation of enzymes by protein kinases, the effects on the target enzymes can be activatory or inhibitory. Interestingly, when antagonistic enzymes in cycles are phosphorylated, usually one is activated and the other is inhibited [37]. Analogous, though more complex, mechanisms account for the high negative control coefficient of glycogen phosphorylase on glycogen synthesis [38]. All these situations necessarily result in negative values for  $SA_i$  and  $Sp_i$ , which is allowed in our modelling since we made no assumption on the sign of those parameters. (It is worth noting that when the enzyme under study does not belong to the branch the flux of which is measured, negative flux control coefficients are expected [39,40], but this situation is not relevant to our approach, because we considered the flux through the whole pathway, with input flux equals output flux.)

The equations that we derived also allowed us to find the optimal distribution of enzyme concentrations in case of constraint on total enzyme. Despite the fact that the amount of enzyme the cell allocates to a given metabolic pathway is necessarily limited, biochemical modelling implicitly considers in most cases that enzyme concentrations are independent of each other, i.e. there is no limitation of resources (matter and energy) in the cell. However, some results clearly indicate that there are constraints on the variation of enzyme concentrations, particularly for the most abundant ones, due to competition for space, energy and cellular resources. Thus global negative correlations may occur between enzyme concentrations, resulting in the so-called 'protein burden effect' [41]. For example, it has been shown that increasing the expression of various glycolytic enzymes in *Zymomonas mobilis* causes a significant decrease in the glycolytic flux and growth rate [42]. In terms of economy of resources, and hence from the evolutionary point of view, it is important to achieve a given flux value with the minimum amount of total enzyme. Formally, this



**Figure 3** Relationship between the concentration of each enzyme and the observed (A) and predicted (B) flux through the pathway

The grey lozenges correspond to the theoretical optimal concentration of each enzyme, i.e. the concentration that leads to maximum flux. The error bars in (A) correspond to  $\pm 1$  standard deviation.

question is equivalent to the one that we addressed, namely how to achieve the maximum flux value with a given amount of total enzyme (for a general treatment of this question, see [26,27,43]). The equations that we derived allow this optimal distribution to be found.

Another noticeable interest of predicting optimal distribution of enzyme concentrations is for *in vitro* metabolic engineering. For instance, a process for the production of 2-oxo-L-gulononic acid, a precursor to vitamin C, implies two enzymatic reactions exterior to the non-living permeabilized *Pantoea citrea* cells used in the bioreactor [44]. The optimal concentrations of those extra enzymes could be determined following the method described in the present paper. Similarly, the approach could be applied to functional protein chips based on mRNA–enzyme fusion molecules hybridized on a DNA microarray, where immobilized enzymes retain activity proportional to the amount of capture DNA, allowing modulation of the relative activity of enzymes;

this novel and elegant technique was used for the optimization of the trehalose synthesis pathway [35], but on an empirical base, while our theoretical treatment would make it rational.

The approach is not restricted to systems reconstituted *in vitro*, and can be applied to two experimental systems closer to *in vivo* conditions. With cell-free extracts, titration experiments using external enzymes are possible, as has been shown for characterizing control properties in glycolysis of mouse muscle [45] and rat liver [46]. Interestingly, it was shown that the kinetic parameters of commercial enzymes did not produce a reliable model of control properties of HK and PFK as compared with kinetic parameters determined in muscle extract [45]. In plants, isolated mesophyll cells from leaves of *Digitaria sanguinalis* keep functional plasmodesmata that allow the free exchange of low-molecular-mass compounds with the culture medium. This system has been used successfully to analyse the kinetic properties of the phosphoenolpyruvate carboxylase, with external



malate dehydrogenase added to the medium [47]. Varying external enzymes in such experimental systems would allow the parameters we defined to be estimated 'in cell'.

In conclusion, we developed a general method that allows us both to predict flux values from global enzyme parameters and to estimate optimal enzyme concentrations. This method is promising for integrative biology and could find applications in metabolic engineering.

## APPENDIX

### Calculation of the optimal concentrations of the enzymes

We used the Lagrange multiplier method to find the distribution of enzyme concentrations that maximizes the flux for a given amount of total enzyme. At maximum flux, we have:

$$\frac{\partial J}{\partial E_i} \left[ J - \lambda \left( \sum_{j=1}^n E_j - E_{\text{tot}} \right) \right] = 0. \quad (\text{A1})$$

Taking the expression of the flux of the main text (eqn 4), we have:

$$\frac{\partial J}{\partial E_i} = \frac{SA_i}{(A_i E_i^* + p_i)^2} \left( \sum_j \frac{1}{A_j E_j^* + p_j} \right)^2, \quad (\text{A2})$$

where  $E_i^*$  is the optimal concentration of enzyme  $i$ .

From eqn (A1), we see that all the derivatives are equal to  $\lambda$ , so we have:

$$\forall i, j \frac{A_i}{(A_i E_i^* + p_i)^2} = \frac{A_j}{(A_j E_j^* + p_j)^2}, \quad (\text{A3})$$

which is equivalent to:

$$\sqrt{\frac{A_i}{A_j}} = \frac{A_i E_i^* + p_i}{A_j E_j^* + p_j}. \quad (\text{A4})$$

Hence:

$$E_j^* = \frac{1}{\sqrt{A_j}} \left( \sqrt{A_i} E_i^* + \frac{p_i}{\sqrt{A_i}} \right) - \frac{p_j}{A_j}. \quad (\text{A5})$$

Introducing  $E_{\text{tot}}$ , it yields:

$$E_{\text{tot}} = \sum_j E_j^* = \sum_j \left( \frac{1}{\sqrt{A_j}} \left( \sqrt{A_i} E_i^* + \frac{p_i}{\sqrt{A_i}} \right) - \frac{p_j}{A_j} \right). \quad (\text{A6})$$

After rearrangements, we get:

$$E_i^* = \frac{1/\sqrt{A_i}}{\sum_j 1/\sqrt{A_j}} \left( E_{\text{tot}} + \sum_j \frac{p_j}{A_j} \right) - \frac{p_i}{A_i}. \quad (\text{A7})$$

This equation leads to eqn (10) of the main text, with

$$e_j^* = E_j^*/E_{\text{tot}}.$$

This general treatment, where implicitly the concentrations are expressed in mol/l, has to be adapted if the constraint on  $E_{\text{tot}}$  is a fixed total weight of enzymes, as we did in the present study.

We thank Professor Hans V. Westerhoff (Faculty of Earth and Life Sciences, Department of Molecular Cell Physiology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands) for fruitful discussions, Professor David Fell (School of Biological and Molecular Sciences, Oxford Brookes University, Oxford, U.K.) for useful suggestions and Professor Olivier C. Martin (Laboratoire de Physique Théorique et Modèles Statistiques, Université Paris-Sud, Orsay Cedex, France) and Dr Delphine Sicard (UMR de Génétique Végétale, INRA/UPS/CNRS/INAPG, Gif-sur-Yvette, France) for careful reading of this paper and relevant remarks. J. B. F. was supported by a Thesis grant from the French Ministère de la Jeunesse, de l'Éducation Nationale et de la Recherche.

In this case, the enzyme concentration,  ${}_g E_i$  (left subscript 'g' for g/l), can be written as:

$${}_g E_i = M_i E_i, \quad (\text{A8})$$

where  $M_i$  is the molecular mass of enzyme  $i$ . Thus the constraint is:

$$\sum_{j=1}^n M_j E_j = {}_g E_{\text{tot}}, \quad (\text{A9})$$

and eqn (A1) becomes:

$$\frac{\partial J}{\partial E_i} \left[ J - \lambda \left( \sum_{j=1}^n M_j E_j - {}_g E_{\text{tot}} \right) \right] = 0. \quad (\text{A10})$$

Here the derivatives of the flux are not equal, since we have:

$$\frac{\partial J}{\partial E_i} = \lambda M_i. \quad (\text{A11})$$

The flux equation can be rewritten to make the molecular masses apparent:

$$J = \frac{S}{\sum_{j=1}^n \frac{1}{A'_j M_j E_j + p_i}}, \quad (\text{A12})$$

with  $A'_j = A_j/M_j$ .

Hence eqn (A11) can be written as:

$$\frac{SA'_i}{(A'_i E_i^* + p_i)^2} = \lambda, \quad \left( \sum_j \frac{1}{A'_{j,g} E_j^* + p_i} \right)^2 = \lambda, \quad (\text{A13})$$

with  ${}_g E_i^*$  the optimal concentration of enzyme  $i$ .

Thus replacing  $A_i$  (respectively  $A_j$ ) by  $A'_i$  (respectively  $A'_j$ ), and  $E_i^*$  (respectively  $E_j^*$ ) by  ${}_g E_i^*$  (respectively  ${}_g E_j^*$ ) in eqns (3)–(7), we obtain directly:

$${}_g E_i^* = \frac{1/\sqrt{A'_i}}{\sum_j 1/\sqrt{A'_j}} \left( {}_g E_{\text{tot}} + \sum_j \frac{p_j}{A'_j} \right) - \frac{p_i}{A'_i}. \quad (\text{A14})$$

This equation leads to eqn (11) of the main text, with

$${}_g e_j^* = {}_g E_j^*/{}_g E_{\text{tot}}.$$

## REFERENCES

- 1 Theobald, U., Mailinger, W., Baltes, M., Rizzi, M. and Reuss, M. (1997) *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Experimental observations. *Biotechnol. Bioeng.* **55**, 305–316
- 2 Teusink, B., Passarge, J., Reijenga, C. A., Esgalhado, E., van der Weijden, C. C., Schepper, M., Walsh, M. C., Bakker, B. M., van Dam, K., Westerhoff, H. V. and Snoep, J. L. (2000) Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. *Eur. J. Biochem.* **267**, 5313–5329
- 3 Bakker, B. M., Michels, P. A., Opperdoes, F. R. and Westerhoff, H. V. (1997) Glycolysis in bloodstream form *Trypanosoma brucei* can be understood in terms of the kinetics of the glycolytic enzymes. *J. Biol. Chem.* **272**, 3207–3215
- 4 Bali, M. and Thomas, S. R. (2001) A modelling study of feedforward activation in human erythrocyte glycolysis. *C.R. Acad. Sci. III* **324**, 185–199
- 5 Mulquiney, P. J. and Kuchel, P. W. (1999) Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: equations and parameter refinement. *Biochem. J.* **342**, 581–596
- 6 Mulquiney, P. J. and Kuchel, P. W. (1999) Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: computer simulation and metabolic control analysis. *Biochem. J.* **342**, 597–604
- 7 Lambeth, M. J. and Kushmerick, M. J. (2002) A computational model for glycogenolysis in skeletal muscle. *Ann. Biomed. Eng.* **30**, 808–827
- 8 Fiévet, J., Dillmann, C., Lagniel, G., Davanture, M., Negroni, L., Labarre, J. and de Vienne, D. (2004) Assessing factors for reliable quantitative proteomics based on two-dimensional electrophoresis. *Proteomics* **4**, 1939–1949
- 9 Gibon, Y., Blaessing, O. E., Hannemann, J., Carillo, P., Hohne, M., Hendriks, J. H., Palacios, N., Cross, J., Selbig, J. and Stitt, M. (2004) A robot-based platform to measure multiple enzyme activities in *Arabidopsis* using a set of cycling assays: comparison of changes of enzyme activities and transcript levels during diurnal cycles and in prolonged darkness. *Plant Cell* **16**, 3304–3325
- 10 Rizzi, M., Baltes, M., Theobald, U. and Reuss, M. (1997) *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. *Biotechnol. Bioeng.* **55**, 592–608
- 11 Visser, D. and Heijnen, J. J. (2003) Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics. *Metab. Eng.* **5**, 164–176
- 12 Poolman, M. G., Fell, D. A. and Thomas, S. (2000) Modelling photosynthesis and its control. *J. Exp. Bot.* **51**, 319–328
- 13 Edwards, J. S., Covert, M. and Palsson, B. (2002) Metabolic modelling of microbes: the flux-balance approach. *Environ. Microbiol.* **4**, 133–140
- 14 Kacser, H. and Burns, J. A. (1973) The control of flux. *Symp. Soc. Exp. Biol.* **27**, 65–104
- 15 Heinrich, R. and Rapoport, T. A. (1974) A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. *Eur. J. Biochem.* **42**, 89–95
- 16 Savageau, M. A. (1971) Concepts relating the behavior of biochemical systems to their underlying molecular properties. *Arch. Biochem. Biophys.* **145**, 612–621
- 17 Torres, N. V., Voit, E. O., Glez-Alcon, C. and Rodriguez, F. (1997) An indirect optimisation method for biochemical systems: description of method and application to the maximization of the rate of ethanol, glycerol, and carbohydrate production in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **55**, 758–772
- 18 Hatzimanikatis, V. and Bailey, J. E. (1997) Effects of spatiotemporal variations on metabolic control: approximate analysis using (log)linear kinetic models. *Biotechnol. Bioeng.* **54**, 91–104
- 19 Visser, D., van der Heijden, R., Mauch, K., Reuss, M. and Heijnen, S. (2000) Tendency modeling: a new approach to obtain simplified kinetic models of metabolism applied to *Saccharomyces cerevisiae*. *Metab. Eng.* **2**, 252–275
- 20 Giersch, C. (1995) Determining elasticities from multiple measurements of flux rates and metabolite concentrations. Application of the multiple modulation method to a reconstituted pathway. *Eur. J. Biochem.* **227**, 194–201
- 21 Torres, N. V., Souto, R. and Melendez-Hevia, E. (1989) Study of the flux and transition time control coefficient profiles in a metabolic system *in vitro* and the effect of an external stimulator. *Biochem. J.* **260**, 763–769
- 22 Saavedra, E., Encalada, R., Pineda, E., Jasso-Chavez, R. and Moreno-Sanchez, R. (2005) Glycolysis in *Entamoeba histolytica*. Biochemical characterization of recombinant glycolytic enzymes and flux control analysis. *FEBS J.* **272**, 1767–1783
- 23 Schellenberger, W., Eschrich, K. and Hofmann, E. (1980) Selforganization of a glycolytic reconstituted enzyme system: alternate stable stationary states, hysteretic transitions and stabilization of the energy charge. *Adv. Enzyme Regul.* **19**, 257–284
- 24 Schellenberger, W. and Hervagault, J. F. (1991) Irreversible transitions in the 6-phosphofructokinase/fructose 1,6-bisphosphatase cycle. *Eur. J. Biochem.* **195**, 109–113
- 25 Curien, G., Ravel, S. and Dumas, R. (2003) A kinetic model of the branch-point between the methionine and threonine biosynthesis pathways in *Arabidopsis thaliana*. *Eur. J. Biochem.* **270**, 4615–4627
- 26 Kuchel, P. W. (1985) Kinetic analysis of multienzyme systems in homogeneous solution. In *Organized Multienzyme Systems* (Welch, G. R., ed.), pp. 303–380, Academic Press, New York
- 27 Heinrich, R., Schuster, S. and Holzhutter, H. G. (1991) Mathematical analysis of enzymic reaction systems using optimization principles. *Eur. J. Biochem.* **201**, 1–21
- 28 Scopes, R. K. (1974) Measurement of protein by spectrophotometry at 205 nm. *Anal. Biochem.* **59**, 277–282
- 29 Cornish-Bowden, A. (1995) *Fundamentals of Enzyme Kinetics*, Portland Press, London
- 30 van Kempen, G. M. and van Vliet, L. J. (2000) Mean and variance of ratio estimators used in fluorescence ratio imaging. *Cytometry* **39**, 300–305
- 31 Kacser, H. and Burns, J. A. (1981) The molecular basis of dominance. *Genetics* **97**, 639–666
- 32 Niederberger, P., Prasad, R., Miozzari, G. and Kacser, H. (1992) A strategy for increasing an *in vivo* flux by genetic manipulations. The tryptophan system of yeast. *Biochem. J.* **287**, 473–479
- 33 Rossignol, R., Faustin, B., Rocher, C., Malgat, M., Mazat, J.-P. and Letellier, T. (2003) Mitochondrial threshold effects. *Biochem. J.* **370**, 751–762
- 34 Sen, A. K. (1996) On the sign pattern of metabolic control coefficients. *J. Theor. Biol.* **182**, 269–275
- 35 Jung, G. Y. and Stephanopoulos, G. (2004) A functional protein chip for pathway optimization and *in vitro* metabolic engineering. *Science* **304**, 428–431
- 36 Aiston, S., Trinh, K. Y., Lange, A. J., Newgard, C. B. and Agius, L. (1999) Glucose-6-phosphatase overexpression lowers glucose 6-phosphate and inhibits glycogen synthesis and glycolysis in hepatocytes without affecting glucokinase translocation. Evidence against feedback inhibition of glucokinase. *J. Biol. Chem.* **274**, 24559–24566
- 37 Cohen, P. (1985) The coordinated control of metabolic pathways by broad-specificity protein kinases and phosphatases. *Curr. Top. Cell. Regul.* **27**, 23–37
- 38 Aiston, S., Hampson, L., Gomez-Foix, A. M., Guinovart, J. J. and Agius, L. (2001) Hepatic glycogen synthesis is highly sensitive to phosphorylase activity: evidence from metabolic control analysis. *J. Biol. Chem.* **276**, 23858–23866
- 39 LaPorte, D. C., Walsh, K. and Koshland, Jr, D. E. (1984) The branch point effect. Ultrasensitivity and subsensitivity to metabolic control. *J. Biol. Chem.* **259**, 14068–14075
- 40 Kruckeberg, A. L., Neuhaus, H. E., Feil, R., Gottlieb, L. D. and Stitt, M. (1989) Decreased-activity mutants of phosphoglucose isomerase in the cytosol and chloroplast of *Clarkia xantiana*. Impact on mass-action ratios and fluxes to sucrose and starch, and estimation of flux control coefficients and elasticity coefficients. *Biochem. J.* **261**, 457–467
- 41 Peter Smits, H., Hauf, J., Muller, S., Hobley, T. J., Zimmermann, F. K., Hahn-Hagerdal, B., Nielsen, J. and Olsson, L. (2000) Simultaneous overexpression of enzymes of the lower part of glycolysis can enhance the fermentative capacity of *Saccharomyces cerevisiae*. *Yeast* **16**, 1325–1334
- 42 Snoep, J. L., Yomano, L. P., Westerhoff, H. V. and Ingram, L. O. (1995) Protein burden in *Zymomonas mobilis*: negative flux and growth control due to overproduction of glycolytic enzymes. *Microbiology* **141**, 2329–2337
- 43 Lion, S., Gabriel, F., Bost, B., Fievet, J., Dillmann, C. and de Vienne, D. (2004) An extension to the metabolic control theory taking into account correlations between enzyme concentrations. *Eur. J. Biochem.* **271**, 4375–4391
- 44 Banta, S., Boston, M., Jarnagin, A. and Anderson, S. (2002) Mathematical modeling of *in vitro* enzymatic production of 2-keto-L-gulonate using NAD(H) or NADP(H) as cofactors. *Metab. Eng.* **4**, 273–284
- 45 Puigjaner, J., Rais, B., Burgos, M., Comin, B., Ovadi, J. and Cascante, M. (1997) Comparison of control analysis data using different approaches: modelling and experiments with muscle extract. *FEBS Lett.* **418**, 47–52
- 46 Torres, N. V., Mateo, F., Melendez-Hevia, E. and Kacser, H. (1986) Kinetics of metabolic pathways. A system *in vitro* to study the control of flux. *Biochem. J.* **234**, 169–174
- 47 Pierre, J. N., Prieto, J. L., Gadal, P. and Vidal, J. (2004) *In situ* C4 phosphoenolpyruvate carboxylase activity and kinetic properties in isolated *Digitaria sanguinalis* mesophyll cells. *Photosynth. Res.* **79**, 349–355