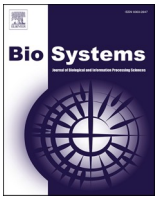
BioSystems 231 (2023) 104984

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems



Control analysis in the identification of key enzymes driving metabolic adaptations: Towards drug target discovery

Pedro de Atauri [a,b,*], Carles Foguet [c], Marta Cascante [a,b,*]

[a] Department of Biochemistry and Molecular Biomedicine & Institute of Biomedicine of Universitat de Barcelona, Faculty of Biology, Universitat de Barcelona, Barcelona, 08028, Spain
[b] Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD), Instituto de Salud Carlos III (ISCIII), Madrid, 28020, Spain
[c] British Heart Foundation Cardiovascular Epidemiology Unit and Victor Phillip Dahdaleh Heart and Lung Research Institute, University of Cambridge, Cambridge, CB2 0BD, United Kingdom

ARTICLE INFO

Keywords:
Metabolic control analysis
Control coefficients
Metabolic adaptations

ABSTRACT

Metabolic Control Analysis (MCA) marked a turning point in understanding the design principles of metabolic network control by establishing control coefficients as a means to quantify the degree of control that an enzyme exerts on flux or metabolite concentrations. MCA has demonstrated that control of metabolic pathways is distributed among many enzymes rather than depending on a single rate-limiting step. MCA also proved that this distribution depends not only on the stoichiometric structure of the network but also on other kinetic determinants, such as the degree of saturation of the enzyme active site, the distance to thermodynamic equilibrium, and metabolite feedback regulatory loops. Consequently, predicting the alterations that occur during metabolic adaptation in response to strong changes involving a redistribution in such control distribution can be challenging. Here, using the framework provided by MCA, we illustrate how control distribution in a metabolic pathway/network depends on enzyme kinetic determinants and to what extent the redistribution of control affects our predictions on candidate enzymes suitable as targets for small molecule inhibition in the drug discovery process. Our results uncover that kinetic determinants can lead to unexpected control distribution and outcomes that cannot be predicted solely from stoichiometric determinants. We also unveil that the inference of key enzyme-drivers of an observed metabolic adaptation can be dramatically improved using mean control coefficients and ruling out those enzyme activities that are associated with low control coefficients. As the use of constraint-based stoichiometric genome-scale metabolic models (GSMMs) becomes increasingly prevalent for identifying genes/enzymes that could be potential drug targets, we anticipate that incorporating kinetic determinants and ruling out enzymes with low control coefficients into GSMM workflows will facilitate more accurate predictions and reveal novel therapeutic targets.

1. Introduction

The year 2023 marks the 50th anniversary of Metabolic Control Analysis (MCA), which originated from two papers published by separate teams: Kacser and Burns (1973) and Heinrich and Rapoport (1974). It has been widely acknowledged that this methodology overturned the dogma that the control of a metabolic pathway is dependent on a single limiting step. Instead, using a formulation based on sensitivity coefficients, known as control coefficients, it introduced a more realistic perspective where control is distributed and, importantly, redistributed in the event of significant perturbations in metabolic pathways.

Control coefficients describe the effects of changes or perturbations in enzyme activities at the molecular level on metabolite concentrations or fluxes at the systemic level (Cornish-Bowden, 2012; Fell, 1997; Heinrich and Schuster, 1996; Miskovic and Hatzimanikatis, 2010; Sauro, 2019). However, this description is only accurate for small changes in enzyme activities that will not result in any redistribution of control. Nevertheless, significant redistributions of control can be expected for larger changes, reflecting the non-linearity of the kinetic mechanisms involved in the metabolic system. This redistribution of

Abbreviations: GSMM, Genome-Scale Metabolic Model; LP, Linear Programming; MCA, Metabolic Control Analysis.

* Corresponding authors. Department of Biochemistry and Molecular Biomedicine & Institute of Biomedicine of Universitat de Barcelona, Faculty of Biology, Universitat de Barcelona, Barcelona, 08028, Spain.
E-mail addresses: pde_atauri@ub.edu (P. de Atauri), martacascante@ub.edu (M. Cascante).

https://doi.org/10.1016/j.biosystems.2023.104984
Received 8 May 2023; Received in revised form 18 July 2023; Accepted 25 July 2023
Available online 26 July 2023
0303-2647/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

control is especially relevant when studying metabolic adaptations or reprogramming that occur in pathological situations in response to drug therapy or in biotechnological interventions. In such cases, in practice, only a partial picture of the system is available for predicting metabolic vulnerabilities to be exploited for biomedical/biotechnological purposes. We cannot determine the precise nature of the system's perturbation, which enzymatic activities have been affected, or to what extent they have been affected. Therefore, the exact mechanism of the perturbation remains unknown. The prediction of such vulnerabilities currently relies mostly on genome-scale metabolic models (GSMMs) (Ebrahim et al., 2013; Wang et al., 2018; Heirendt et al., 2019) based on the stoichiometric structure of the metabolic networks, the analysis of changes in fluxes of metabolite uptakes and secretions, and metabolite concentrations at the systemic level, while assuming that alterations in gene expression or protein levels are adequate to fully describe changes in enzyme activities at the molecular level (Katzir et al., 2019).

Even the success of constrain-based GSMMs, there is increasing evidence that these approaches have important limitations in accurately inferring the key enzyme changes that sustain metabolic adaptations triggered by system perturbations, such as those induced by a drug. The predictive capacity of GSMMs has been improved by incorporating additional constraints, such as ensuring that each metabolic flux does not exceed its maximum capacity, which is equal to the product of the enzyme's abundance and turnover number (Sánchez et al., 2017; Domenzain et al., 2022). In these enzyme-constrained models, sensitivity coefficients, such as control coefficients on the predicted growth rate and other predicted fluxes, are estimated based on these approximated maximum capacities (Nilsson and Nielsen 2016; Wilken et al., 2022). An improvement in the accuracy of control determination can be expected as these sensitivity coefficients gradually can incorporate all kinetic determinants. The recent publication of deep learning-predicted Michaelis constants (Km) on a genome-scale (Kroll et al., 2021) is expected to fuel the incorporation of kinetic determinants into GSMMs (Antolin and Cascante, 2021), generating an urgent need for tools that can thoroughly explore the role of kinetic determinants in control distribution, and understand how control redistribution can affect the identification of putative drug targets.

The specific objectives of the present study are to address how does control distribution depend on kinetic determinants other than the stoichiometric structure, and to what extent does the redistribution of control affect the predictions based on control coefficients. As a proof of concept, we first examine how control distribution depends on factors such as negative feedback regulation circuits or the presence of isoenzymes. Second, we investigate the impact of control redistribution, particularly focusing on the ability to identify altered enzyme activities that serve as key drivers of metabolic adaptations to mechanism-unknown perturbations. In this study, we use an inference strategy, based on linear programming (LP), that assumes control distribution remains constant after perturbation (de Atauri et al., 2021). Given a metabolic adaptation to a perturbation, theoretically, having complete knowledge of the control distribution, as well as the concentrations and reaction fluxes at the systemic level before and after the perturbation, should enable us to identify which key enzymatic activities were impacted at the molecular level. The identification of these key drivers is based on the predictive capacity of control coefficients. However, these predictions can deviate due to the reorganization of control distribution following the perturbation. An assessment of the distortion level can be conducted by comparing it with simulations based on detailed models that encompass complete knowledge of the rate laws involved in all molecular mechanisms. The same models are used to make predictions on key drivers and simulations.

To illustrate these points, we use two kinetic models. The first is a simple model that covers the upper part of glycolysis in mouse muscle (Puigjaner et al., 1997). This model is applied to address the two objectives. Additionally, we employ a second model that encompasses the entire central metabolism of *Escherichia coli* (Millard et al., 2017), which is used to address the second objective in a more complex scenario.

2. Methods

2.1. MCA formulation

In MCA, sensitivity coefficients are used to measure the dependencies of systemic variables, such as metabolic fluxes or metabolite concentrations, with respect to enzymes or parameters in metabolic networks (Cornish-Bowden, 2012; Fell, 1997; Heinrich and Schuster, 1996; Miskovic and Hatzimanikatis, 2010; Sauro, 2019). Table 1 summarizes the different types of sensitivity coefficients for a metabolic system with n internal metabolites ($i = 1, \ldots, n$) and m reaction steps ($j = k = 1, \ldots, m$). Here, x_i represents the systemic steady state metabolite concentration for metabolite i. J_j and v_k correspond to the same steady state measure. On the one hand, J_j is taken as the systemic steady state reaction flux that changes around the steady state as a system-dependent variable. On the other hand, v_k is taken as the molecular enzyme activity that changes around the steady state as a variable independent of the rest of the metabolic system. Both measures are applicable for the reaction step j=k. Molecular activities can be taken as proportional to enzyme concentration and can generally be approximated from protein or gene expression levels. However, they can also be modulated by phosphorylation or other covalent modifications (Martín-Bernabé et al., 2017). Systemic transport processes and molecular transporters, which can be relevant actors in metabolism, are treated as systemic fluxes and enzyme activities, respectively.

At the systemic level, for global changes in concentrations and fluxes that are system-dependent, sensitivity coefficients can be divided into control and response coefficients. Control coefficients describe variations in metabolite concentrations and reaction fluxes in response to perturbations in enzyme activities (v_k taken as parameters), while response coefficients describe the same variations in response to any other perturbation (parameter p). At the molecular level, for local changes in enzyme activities that behave independently of the rest of the system, elasticities are described as variations in enzyme activities (v_k taken as variables) in response to perturbations in metabolite concentrations, known as metabolite elasticities, or in response to any other perturbation (parameter p), known as parameter elasticities. The MCA formulation includes sets of theorems that establish dependencies between these sensitivities:

$$\sum_{k=1}^{m} C_{v_k}^{x_i} = 0 \text{ and } \sum_{k=1}^{m} C_{v_k}^{J_j} = 1 \text{ (summation theorems)} \quad (1)$$

Table 1
Sensitivity coefficients.

		a	b
concentration control coefficients	$C_{v_k}^{x_i} = \frac{v_{ko}}{x_{io}}\frac{dx_i}{dv_k} = \frac{d\log x_i}{d\log v_k}$	x_i	v_k
flux control coefficients	$C_{v_k}^{J_j} = \frac{v_{ko}}{J_{jo}}\frac{dJ_j}{dv_k} = \frac{d\log J_j}{d\log v_k}$	J_j	v_k
"metabolite" elasticities	$\varepsilon_{x_i}^{v_k} = \frac{x_{io}}{v_{ko}}\frac{\partial v_k}{\partial x_i} = \frac{\partial\log v_k}{\partial\log x_i}$	v_k	x_i
concentration response coefficients	$R_p^{x_i} = \frac{p}{x_{io}}\frac{dx_i}{dp} = \frac{d\log x_i}{d\log p}$	x_i	p
flux response coefficients	$R_p^{J_j} = \frac{p}{J_{jo}}\frac{dJ_j}{dp} = \frac{d\log J_j}{d\log p}$	J_j	p
"parameter" elasticities	$\varepsilon_p^{v_k} = \frac{p}{v_{ko}}\frac{\partial v_k}{\partial p} = \frac{\partial\log v_k}{\partial\log p}$	v_k	p

Each sensitivity coefficient is a dimensionless quantity that measures the fractional change in a variable a per fractional change in a parameter b around a steady-state ($x_{io}, J_{ko} = v_{ko}$). Notation as used in de Atauri et al. (2021). Alternative notations could be utilized. For example, control and response coefficients could be defined as partial derivatives, and what we have referred to as 'parameter elasticities' might also be known as 'π-elasticities' (Heinrich and Schuster, 1996).

and

$$\sum_{k=1}^{m} C_{v_k}^{x_a} \times \varepsilon_{x_b}^{v_k} = 0\ (a \neq b), \sum_{k=1}^{m} C_{v_k}^{x_a} \times \varepsilon_{x_a}^{v_k} = -1, \text{and} \sum_{k=1}^{m} C_{v_k}^{J_j} \times \varepsilon_{x_i}^{v_k} = 0\ (\text{connectivity theorems}) \quad (2)$$

For the estimation of control coefficients, various matrix formulations have been developed in the context of MCA (Cascante et al., 1989a, b; Fell and Sauro, 1985; Westerhoff and Kell, 1987; Reder, 1988; Small and Fell, 1989). These matrix methods imply all summation and connectivity dependencies in Table 1, along with stoichiometric dependencies of fluxes and concentrations of species involved in moiety conservations.

2.2. Model I. Upper part of glycolysis on mice muscle extracts

Model I is based on a published kinetic model that covers the upper part of glycolysis. This model was parametrized from experiments performed on mouse muscle extracts (Puigjaner et al., 1997). The kinetic laws, their parameters, and steady-state concentrations and fluxes were obtained from the original publication. The original irreversible Hill equation for phosphofructokinase (PFK) has been replaced with a reversible Hill equation (Hofmeyr and Cornish-Bowden, 1997). The equation was adjusted based on published Gibbs free energies of formation for the metabolites involved in the reaction, as well as the ADP/ATP ratio from human muscle (Fernandes et al., 2019). The fructose-1,6-biphosphate (FbP) concentration scaled by its half saturation constant (FbP/$K_{0.5}$) was adjusted to provide negative elasticities for

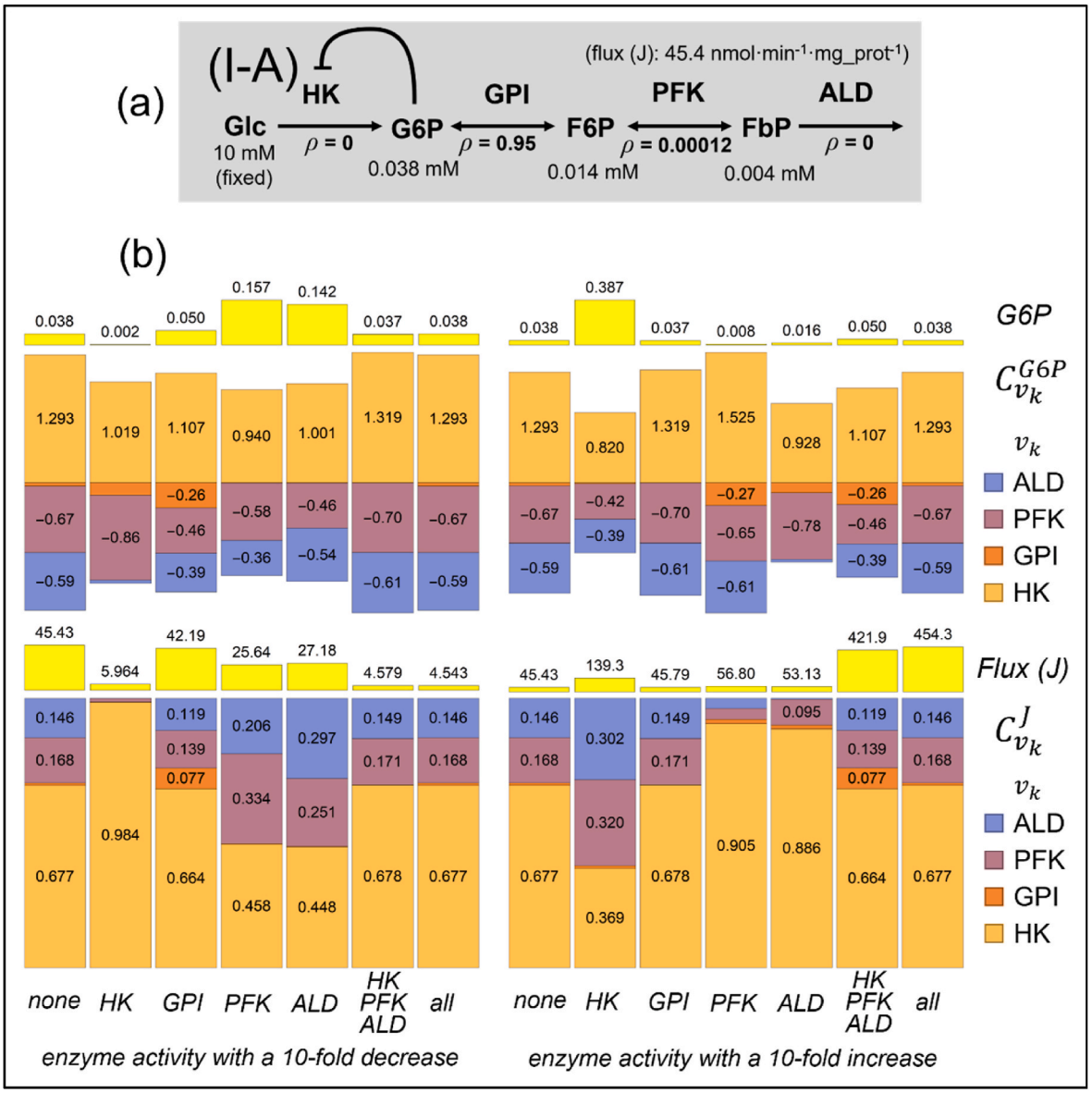


Fig. 1. Magnitudes of control coefficients and control redistribution. (a) Scheme for Model I-A. (b) Each column shows the distribution of control coefficients for flux (J) and G6P concentration, as well as the corresponding flux and G6P concentration, as estimated in Model I (Variant I-A). Each column corresponds to a ten-fold decrease or increase of enzyme activities: 1st) none; 2nd) HK; 3rd) GPI; 4th) PFK; 5th) ALD; 6th) simultaneously HK, PFK, and ALD; and 7th) simultaneously all enzymes. Changes in G6P concentrations/G6P-concentration control coefficients and the flux/flux control coefficients are in the upper and lower panels, respectively. Tenfold decreases and tenfold increases are in the left and right panels, respectively. Flux units are *nmol* min^{-1} *mg* $prot^{-1}$, and concentration units are millimolar (*mM*). ρ represents the disequilibrium ratio ($\rho = \Gamma/K$). Model I-A is described in the Appendix.

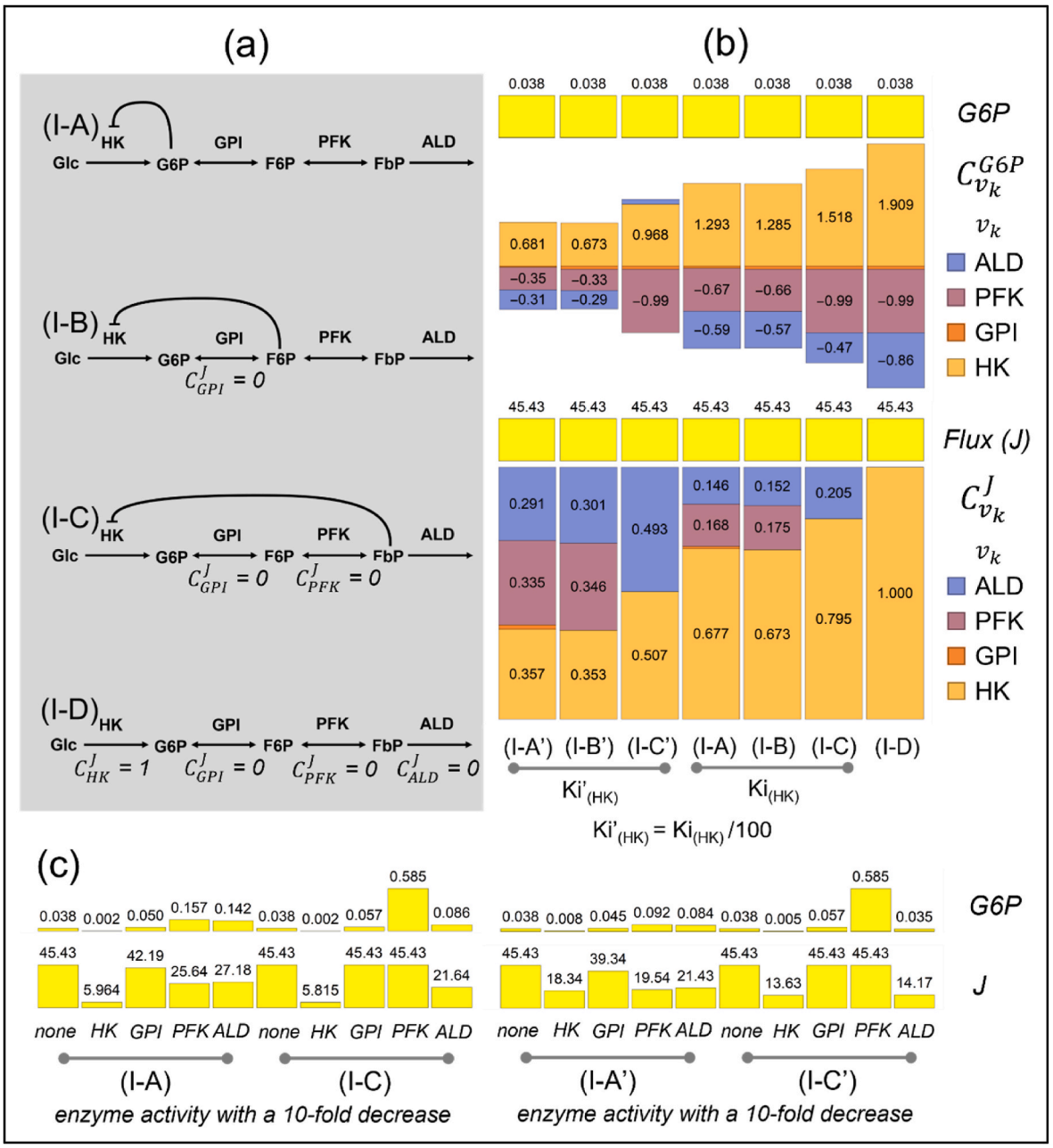


Fig. 2. Regulatory circuits and control distribution. (a) Schemes of four variations of Model I that vary in the metabolite responsible for product inhibition on HK (I-A, I–B, I–C), and without inhibition (I–D). (b) Distribution of control coefficients for flux (J) and G6P concentration, as well as the corresponding flux and G6P concentration, estimated in the four variations of Model I – I-A, I–B, I–C, and I-D – along with three additional variants of Models I-A, I–B, and I–C (designated as I-A', I–B', and I–C', respectively), in which the inhibition constant for HK was modified. Each column corresponds to a variant of Model I. All variants were adjusted to the same steady-state flux and concentrations. (c) Each column corresponds to a tenfold decrease in enzyme activities for Model I-A, I–C, I-A', and I–C': 1st) none; 2nd) HK; 3rd) GPI; 4th) PFK; 5th) ALD. Model I variants are described in the Appendix.

FbP in all the analysed steady states.

As described in Figs. 1–3, a total of five model variants (A, B, C, D, and E) were utilized for comparisons, which allowed for the assessment of the impact of different control determinants. Model I-A features a feedback inhibition on the first metabolic intermediary targeting the first enzyme. The feedback inhibition is shifted along the metabolic pathway to the second and third metabolic intermediaries in Models I–B and I–C, respectively. In contrast, Model I-D lacks any feedback inhibition. Model I-E, another variant of Model I-A, incorporates two isoenzymes catalysing the first reaction step. These five model variants were adjusted to provide the same steady-state flux and concentrations as the original model. Model I-A is the original model, but with an irreversible Hill equation. In models I–B to I-D, corrections were applied to maintain inhibitor concentration/inhibition constant (*Ki*) ratios, as well as apparent *Km* and *Vmax* for Hexokinase activities (HK), as in model I-A. In model I-E, the first step catalysed by HK was modified to be catalysed by two isoenzymes, which correspond to HK type I and type II in muscle (Ritov and Kelley, 2001; Wilson, 2003).

The model includes three system-dependent metabolites and four (five in model I-E) enzyme-catalysed reactions. A detailed description of the model equations is provided in the Appendix. **Model abbreviations**: G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; FbP, fructose-1,6-diphosphate; HK, hexokinase; GPI, glucose-phosphate isomerase; PFK, phosphofructokinase; ALD, aldolase.

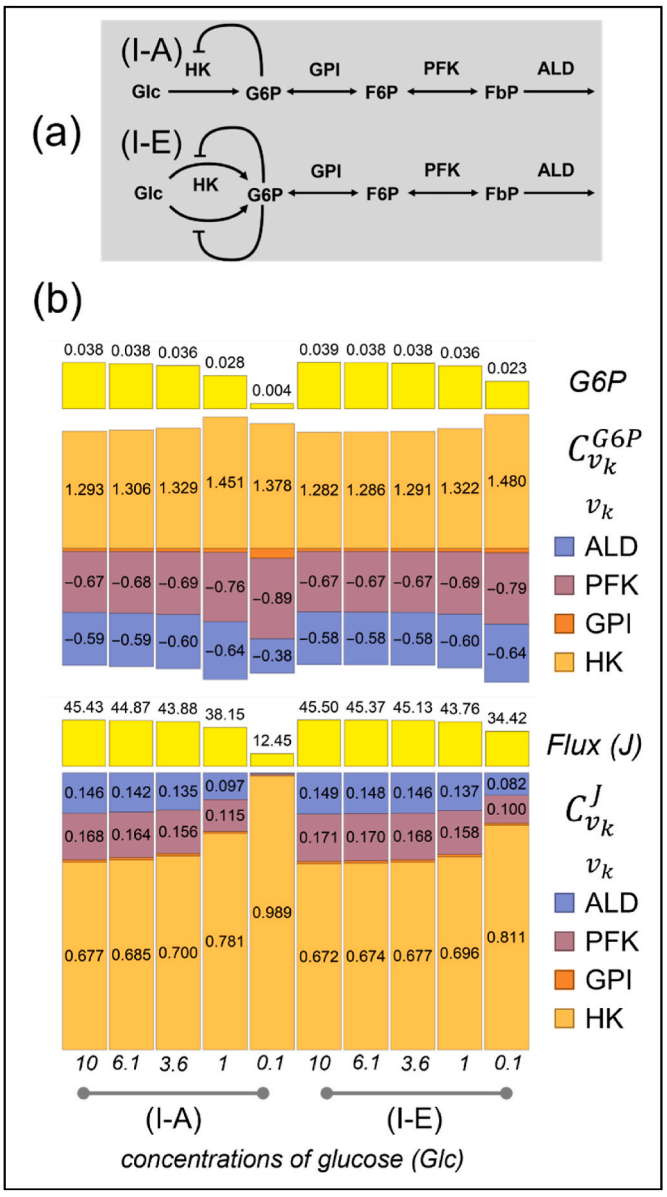


Fig. 3. Isoenzymes and control distribution. (a) Schemes for Models I-A (without isoenzymes) and I-E (with isoenzymes for HK activity). (b) Distribution of control coefficients for the flux (J) and G6P concentration, as well as the corresponding flux and G6P concentration, estimated in Models I-A and I-E. Each column corresponds to a concentration of glucose substrate. Model I variants are described in the Appendix.

2.3. Model II. Glucose uptake, catabolism, ATP production and growth in *Escherichia coli*

Model II is based on a published and detailed kinetic model that describes the central metabolism of *Escherichia coli* cultivated on glucose under aerobic conditions (Millard et al., 2017) and includes 62 metabolites and 67 reactions. The concentration of external glucose was set as constant.

2.4. Inference by interval reduction

Given a single known flux control coefficient and that variations in enzyme activity are small enough for this control coefficient to remain constant, by linear approximation predictions can be made about the change in the flux value (Fell, 1997). Predictions for larger changes are possible, although these linear predictions will be less accurate as the changes become larger. A response coefficient in response to a perturbation p can be rewritten as a sum of terms, each of which quantifies the variations in a metabolite concentration or reaction flux in response to the perturbation of one of the enzymes (Kholodenko, 1988; Cascante et al., 1989a; 1989b; Hofmeyr and Cornish-Bowden, 1991):

$$R^{x_i}_p = \sum_{k=1}^{m} C^{x_i}_{v_k} \times \varepsilon^{v_k}_p = C^{x_i}_{v_1} \times \varepsilon^{v_1}_p + \cdots + C^{x_i}_{v_m} \times \varepsilon^{v_m}_p \quad (3)$$

$$R^{J_j}_p = \sum_{k=1}^{m} C^{J_j}_{v_k} \times \varepsilon^{v_k}_p = C^{J_j}_{v_1} \times \varepsilon^{v_1}_p + \cdots + C^{J_j}_{v_m} \times \varepsilon^{v_m}_p \quad (4)$$

Under this definition, while each control coefficient describes variations in a metabolite concentration or reaction flux in response to the perturbation of a single enzyme activity, each response coefficient is presented as a general sensitivity coefficient. This describes the response to any perturbation, which could proportionally or not affect a single enzyme activity, or multiple enzyme activities simultaneously. According to this definition, a response coefficient will correspond to a control coefficient when a single activity is affected proportionally ($\varepsilon^{v_k}_p = 1$). These expressions integrate changes at both systemic and molecular levels, accounting for the effect of an infinitesimal or small perturbation p. Analogously to the single control coefficient, given known control coefficients, by linear approximation predictions can be made about the changes in fluxes and concentrations, which will become less accurate as the changes become larger. Assuming this loss of accuracy, using a constraint and optimization-based method, we recently applied an inference strategy that assumes constant control coefficients and incorporates constraints from Eqs. (3) and (4) to identify the key drivers of metabolic adaptations (de Atauri et al., 2021). Our aim was to identify the key unknown changes in enzyme activities that drive metabolic adaptations resulting from mechanism-unknown perturbations. Firstly, analogous to Eqs. (3) and (4), the following expression was derived,

$$\Delta\log x_i = C^{x_i}_{v_1} \times \Delta\log v_1 + \cdots + C^{x_i}_{v_m} \times \Delta\log v_m = \sum_{k=1}^{m} C^{x_i}_{v_k} \times \Delta\log v_k \quad (5)$$

$$\Delta\log J_j = C^{J_j}_{v_1} \times \Delta\log v_1 + \cdots + C^{J_j}_{v_m} \times \Delta\log v_m = \sum_{k=1}^{m} C^{J_j}_{v_k} \times \Delta\log v_k \quad (6)$$

where $\Delta\log x_i = \log(x_{ifinal}/x_{iinitial})$, $\Delta\log J_j = \log(J_{jfinal}/J_{jinitial})$, and $\Delta\log v_k = \log(v_{kfinal}/v_{kinitial})$ are log fold changes for systemic concentrations, systemic reaction fluxes, and molecular enzyme activities, respectively, describing the change from an initial state (i.e, before the perturbation) to a final state (i.e., after the perturbation). For these log fold changes natural logarithms are applied, although any logarithmic base could be used. All log fold changes are presented as domains of possible values, which are intervals with lower and upper bounds. A subset of these intervals is constrained to known bounds, while the remaining intervals have unknown lower and upper bounds that are set to a common value of −4 or +4, respectively, which corresponds to fold decreases or increases of around 55, an arbitrary value very far from the expected changes. Secondly, control coefficients are assumed to be fixed parameters; therefore, Eqs. (5) and (6) can be treated as linear equations, where log fold changes in concentrations, reaction fluxes, and enzyme activities are the variables. Taking these equations into account, an interval reduction problem can be formulated that solves two LP problems – maximization and minimization, one at time, for each one of the variables of interest (de Atauri et al., 2021),

$$
\begin{array}{lll}
\text{maximize (and minimize)} & z = \Delta\log x_i, \Delta\log J_j, \Delta\log v_k & \forall i \in N, \\
 & & \forall j,k \in M \\
\text{subject to} & -\Delta\log x_i + \sum_{k=1}^{m} C^{x_i}_{v_k} \times \Delta\log v_k = 0 & \forall i \\
 & -\Delta\log J_j + \sum_{k=1}^{m} C^{J_j}_{v_k} \times \Delta\log v_k = 0 & \forall j \\
 & lb^{in}_{x_i} \le \Delta\log x_i \le ub^{in}_{x_i} & \forall i \\
 & lb^{in}_{J_j} \le \Delta\log J_j \le ub^{in}_{J_j} & \forall j \\
N = 1,\cdots,n & lb^{in}_{v_k} \le \Delta\log v_k \le ub^{in}_{v_k} & \forall k \\
M = 1,\cdots,m & \Delta\log x_i, \Delta\log J_j, \Delta\log v_k \in \mathbb{R} &
\end{array}
\tag{7}
$$

where, by maximizing and minimizing each variable, the lower (*lb*) and upper (*ub*) bounds of the initial domain (*in*) of possible values for this variable, $lb^{in}_{variable}$and $ub^{in}_{variable}$, respectively, is contracted or reduced to a final domain (*fi*), $lb^{fi}_{variable}$and $ub^{fi}_{variable}$, that satisfy all implicit constraints ($lb^{in}_{variable} \le lb^{fi}_{variable}$ and $ub^{fi}_{variable} \le ub^{in}_{variable}$). By maximizing and minimizing each one of variables of interest ($\Delta\log x_i$, $\Delta\log J_j$, and $\Delta\log v_k$), the method returns the ranges of feasible log fold changes for each of the m systemic fluxes ($\Delta\log J_j$), n concentrations ($\Delta\log x_i$), and m molecular activities ($\Delta\log v_k$).

In our previous work, we demonstrated the effectiveness of this approach in identifying targets within a highly uncertain environment (de Atauri et al., 2021). We used measured fold changes for nearly all systemic fluxes, a few concentrations, and some restrictions on molecular activities. Instead of employing a single fixed set of control coefficients, we generated multiple sets through sampling, while maintaining the same stoichiometry and regulatory networks.

In the present paper, a single set of constant control coefficients is applied, therefore a unique application of the algorithm for interval reduction is done accounting for the minimum and maximum values of each variable. In this application, all fold changes for enzyme activities are unknowns, while all fold changes for concentrations and fluxes are known. For the unknown variables, the lower and upper bounds in the initial domains are kept unknown (±4). For the known variables, there is always a certain degree of uncertainty associated, which will propagate to solution spaces. Thus, their lower and upper bounds can be derived from measured errors or other measures of uncertainty. In this study, a percentage δ of uncertainty is assigned around the values obtained by model simulations for all concentrations ($x_{iinitial} \pm \delta\%$ *and* $x_{ifinal} \pm \delta\%$) and reaction fluxes ($J_{j_{initial}} \pm \delta\%$ *and* $J_{j_{final}} \pm \delta\%$). Therefore, the reduced domains for the log fold changes of concentrations and fluxes are calculated as follows:

$$\Delta\log x_i = \left[\log\frac{x_{ifinal} - \delta\%}{x_{iinitial} + \delta\%}, \log\frac{x_{ifinal} + \delta\%}{x_{iinitial} - \delta\%}\right] \tag{8}$$

$$\Delta\log J_j = \left[\log\frac{J_{jfinal} - \delta\%}{J_{jinitial} + \delta\%}, \log\frac{J_{jfinal} + \delta\%}{J_{jinitial} - \delta\%}\right] \tag{9}$$

In the application of the approach performed in this study, we assume that we know all the fold changes for fluxes and concentrations, while all fold changes for enzyme activities are unknown. However, the approach is very flexible and does not require knowing the fold change of all fluxes and concentrations. These could be unknowns, just like the fold change of enzyme activities (examples will be provided later as part of the analyses). Additionally, known fold changes in enzyme activities could be introduced. Finally, although in all the perturbations applied in this study, the sign of the fluxes was maintained, Eq. (9) cannot be applied when net fluxes switch from positive to negative values. In each pair $J_{j_{initial}}$ and $J_{j_{final}}$, both are positive or both are negative. This is a limitation of the proposed method, which can be circumvented by splitting the net fluxes into forward and reverse fluxes, where apparent control coefficients for the forward and reverse reactions can be applied. Provided that reverse and forward reactions are not independent, a constraint must be added, as fold changes in the corresponding enzyme activities should be equal. Alternatively, the removal of the corresponding linear constraint involving this flux should have an impact on the solution space, which would depend on the relevance of the affected reaction in the metabolic network.

2.5. Calculations

All calculations were performed using 'Wolfram Mathematica 12' (www.wolfram.com). To solve steady-state simulations before and after perturbations, with kinetic models represented as a system of ordinary differential equations (ODEs), deterministic methods were used. The model simulations were solved as initial value problems. For solving control coefficients, first, elasticities are estimated by direct derivation of the model reaction-rate laws around the solved steady states, and then any of the matrix formulations to solve control coefficients can be applied (Cascante et al., 1989a,b; Fell and Sauro, 1985; Westerhoff and Kell, 1987; Reder, 1988; Small and Fell, 1989). For solving inference problems, a Mathematica notebook is available on Zenodo (de Atauri et al., 2021b).

3. Results

3.1. How does control distribution depend on determinants others than the stoichiometric structure?

Fig. 1 provides a detailed analysis of Model I-A. As shown in Fig. 1b, the steady state concentrations and fluxes and the control distribution of the initial (non-modified) model were first estimated by model simulation and matrix methods, respectively, and are presented in the first column of the four panels (labelled as "none"). Subsequently, all enzyme activities were either decreased or increased by a factor of 10, either alone or simultaneously. For each perturbation, the modified model was allowed to evolve via simulation to a new steady state and control distribution, which are estimated and presented in one of the subsequent columns (labelled with the modified activity or activities). The comparison with the first column shows how control coefficients have been redistributed.

The extent of the resulting control redistribution depends on both the degree of control possessed by the affected activities and the magnitude of the alteration made. As expected, Fig. 1 illustrates how modifying enzymes with significant control, such as HK, PFK, and aldolase (ALD), leads to a substantial impact on control redistribution following changes in steady-state concentrations and fluxes. Conversely, modifying an enzyme with very low control, such as glucose-phosphate isomerase (GPI), does not have a significant effect on control redistribution. This does not mean that it cannot have some degree of control. As can be seen in the figure (6th column in the two right panels), simultaneously increasing the activities of HK, PFK, and ALD results in a slight but significant increase in the control of GPI. This effect could become dominant with higher increases in the activities of these enzymes. Consistently, the same pattern is observed when only GPI is decreased (3rd column in the two left panels).

For reactions that are close to equilibrium, such as the reaction catalysed by GPI, the elasticities are determined mainly by the degree of displacement of the reaction from equilibrium (Fell, 1997; Cornish--Bowden et al., 2012). Thus, the elasticity for GPI, which follows a reversible Michaelis-Menten rate equation, can be rewritten as (Westerhoff et al., 1984; Rohwer and Hofmeyr, 2010; Hofmeyr, 1995):

$$\epsilon^{v}_{G6P} = \frac{1}{1-\rho} - \frac{G6P/Kms}{1 + G6P/Kms + F6P/Kmp} \tag{10}$$

$$\epsilon^{v}_{F6P} = -\frac{\rho}{1-\rho} - \frac{F6P/Kmp}{1 + G6P/Kms + F6P/Kmp} \tag{11}$$

where G6P and F6P are the substrate and the product, respectively, and ρ is the disequilibrium ratio, which is the ratio of the mass-action ratio Γ to the equilibrium constant K (ρ = Γ/K). While the second right-hand term is the fractional saturation, the first term depends on the distance to the equilibrium, going to infinity at equilibrium. Therefore, in situations close to equilibrium, this value will be increasingly higher, which will contribute to giving low values for control coefficients, such as those seen for GPI. This intrinsic thermodynamic property of the reaction catalysed by GPI has an impact on the ability of GPI activity to alter both steady-state values and control distribution, as depicted in Fig. 1.

Both the distance to equilibrium and the degree of enzyme activity saturation provide us with two examples to illustrate control determinants that should be taken into consideration. The heterogeneous list of control determinants also should include regulatory circuits and the presence of isoenzymes. Fig. 2 illustrates the distribution of control in four different model variants that differ in the metabolite responsible for product inhibition on HK, as well as the degree of inhibition. In the pathway lacking any feedback inhibition (Model I-D), it can be observed that all control shifts to the initial step catalysed by HK. Despite PFK and ALD being reactions that are far from equilibrium or irreversible in Model I-A, they lose all control. When the metabolic intermediary involved in the feedback inhibition of the first step catalysed by HK is moved along the metabolic pathway (Models I-A, I–B, and I–C), control is redistributed. As previously established for linear reaction chains (Heinrich and Schuster, 1996), the enzyme activities that become involved inside the loop lose all control, regardless of their previous level of control in Model I-A. Also, in Fig. 2, another component of the regulatory circuits, the magnitude of the inhibition, is explored. Decreasing the inhibition constant (*Ki*), the control moves from HK to PFK and ALD. This, in turn, can result in significant differences in the resulting outcomes, changes in fluxes and concentrations, as illustrated in the figure by reducing the activities of the four enzymes.

Concerning isoenzymes, Fig. 3 shows that adding isoenzymes for the first enzyme, with Kms covering a wide range of values, is an example of a design that can maintain control distribution, as well as steady-state concentrations and fluxes, when faced with variable concentrations of glucose, the substrate of the pathway. In this case, the two HKs in Model I-E with *Kms* of 0.04 mM (HK Type I) and 0.4 mM (HK Type II) enable control to be maintained over a wider range of values compared to the single HK in Model I-A, which has a *Km* of 0.4 mM. This is especially relevant when designing drug interventions that target specific isoenzymes that are overexpressed in a particular subtype of tumours. Indeed, one of the mechanisms of escaping drug interventions is through the existence of isoenzymes that catalyse the same reactions, which can compensate for the inhibition of a specific isoenzyme. An interesting example is mutant isocitrate dehydrogenases (IDH1 and IDH2), which are the main source of the oncometabolite 2-hydroxyglutarate (2HG), contributing to oncogenesis. It has been described that the expression of IDH1 and IDH2 mutants is tightly orchestrated by Polo-like kinase 1 (PlK1) during mitosis to ensure optimal production of 2HG (Saikiran Reddy et al., 2022). Furthermore, cancer cells switch between mutant IDH1 or IDH2 to acquire resistance to IDH inhibitors, ensuring 2HG production (Harding et al., 2018). Another illustrative example is the HK isoenzymes expression profile, which is described to change in cancerous transformation (Perrin-Cocon et al., 2021; Tseng et al., 2018) and to be linked to poor prognosis and acquisition of resistance to chemotherapy (Varghese et al., 2020; Marcucci and Rumio, 2021).

Lastly, another illustrative example of a control determinant is a regulatory motif, commonly known as the Universal Method or Multisite Modulation (Kacser and Acerenza, 1993; Fell and Thomas, 1995), which is independent of all the determinants discussed thus far. This method enables the increase or decrease of a flux without affecting steady-state concentrations and control distribution. As seen in the last columns of the four panels in Fig. 1, a simultaneous proportional decrease or increase in all enzymatic activities will result in a proportional decrease or increase in the flux, without affecting concentrations or control coefficients.

3.2. To what extent does the redistribution of control affect the predictions based on control coefficients?

As discussed above, the control coefficients have a predictive value, although this is only entirely accurate for small changes. The consequence of the redistribution of control associated with large changes is, in turn, the loss of predictive value for the original control coefficients. While sacrificing predictive power, approximations made for large changes can still capture the main trends of the variations, which can be sufficient for applications such as the one proposed for identifying key drivers of metabolic adaptations. To consider the effect of large changes, different approaches have explored the sensitivity of metabolic variables to finite changes. These approaches utilize adaptations of control coefficients that are based on different assumptions and can deal with large changes in enzyme activities (Small and Kacser, 1993a,b; Acerenza, 2000; Acerenza and Ortega, 2007; Ortega and Acerenza, 2011; Acerenza et al., 2015).

In our analyses, applied to mechanism-unknown perturbations, we use the inference strategy designed to identify key drivers. It is summarized in Eq. (7) and is based on the assumptions underlying Eqs. (5) and (6). At large changes, the performance in the identification of key drivers is due to the level of approximation implicit in these expressions. Given that control coefficients are assumed to be constant parameters and since control should be redistributed after a large perturbation, therefore, the accuracy of these expressions depends on the degree of control redistribution. Using kinetic Model I-A, as shown in Fig. 1a, we simulated large perturbations by increasing and decreasing the activity of each driver enzyme, which generated two steady state conditions for each perturbation (one before and one after the modification). Our objective was to assess how well the inference strategy identifies the driver and how this identification is affected by control redistributions. The procedure was applied for the same tenfold decreases and increases of enzyme activities applied in Fig. 1, and the result is illustrated in Fig. 4 for the tenfold decrease of the enzyme activity of PFK in the Model I-A: 1) exact values for initial steady-state concentrations, fluxes, and control coefficients (initial control coefficients) are solved (Fig. 1b, 1st column in all panels); 2) a large perturbation is introduced in the original model (decreasing the enzyme concentration of PFK ten times) and a new set of exact values for steady-state concentrations, fluxes, and control coefficients (final control coefficients) is solved (Fig. 1b, 4th column of all panels); and 3) changes in enzyme activities (drivers) required to explain the changes in concentrations and fluxes are inferred by applying the inference strategy. For this last step, 1) lower and upper bounds were estimated for the initial domains of log fold changes of the final steady state with respect to the initial steady state for the previously solved concentrations and reaction fluxes; and 2) all control coefficients were fixed. The three upper panels and the three bottom panels show the same simulations, but with different levels of uncertainty. In the lower set of panels, a percentage of uncertainty (δ) equal to 15% was assigned to the exact values for fluxes and concentrations. In addition, in the upper set of panels, to represent the final domains as points on the figures and clearly display unexpected deviations, the uncertainty was reduced to a much smaller value (δ = 0.000001%), which allowed the final domains to be represented as very short intervals displayable as points. These points were applied to estimate the cumulative deviation from expected changes. As seen in the upper panels in Fig. 4, to compare the cumulative deviation from expected changes, a performance measure denoted by η was applied in each analysis. Each performance measure quantifies the accumulated relative deviations in the predicted versus expected fold changes for all enzyme activities:

$$\eta = \left| \frac{\sum_k^m \left|\Delta \log v_{k_{predicted}}\right| - \sum_k^m \left|\Delta \log v_{k_{expected}}\right|}{\sum_k^m \left|\Delta \log v_{k_{expected}}\right|} \right| \quad (12)$$

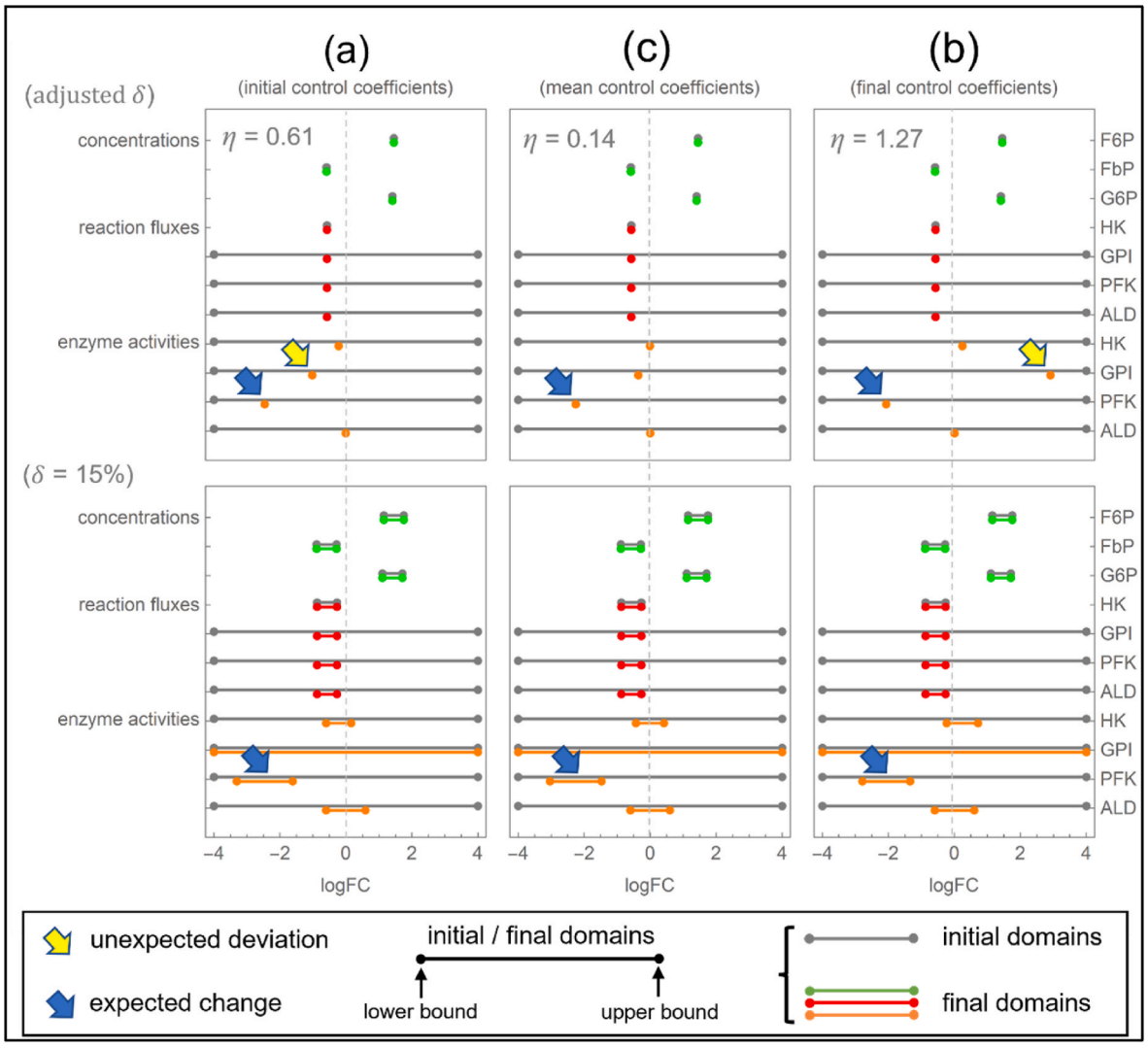


Fig. 4. Identification of key drivers in Model I-A. PFK decrease. Initial domains (grey) and final domains for log fold changes (logFC) in systemic concentrations (x_i; $i = 1, \ldots,n$) (green) and reaction fluxes (J_j; $j = 1, \ldots,m$) (red), as well as in molecular enzyme activities (v_k; $k = 1, \ldots,m$) (orange). The fixed control coefficients applied are those estimated before the perturbation (initial control coefficients) (a), after the perturbation (final control coefficients) (b) and using the mean values (mean control coefficients) (c). Each η denotes the cumulative deviation (upper panels) and the blue arrow indicates the expected change, while the yellow arrow shows the more relevant unexpected deviation. A tenfold decrease in the level of PFK activity is expected based on the pattern of fold changes measured in all concentrations and fluxes. Each measured fold change in concentrations and fluxes is estimated as a log fold change of the value after the perturbation with respect to the value before the perturbation, with a correction factor δ that involves associating a degree of uncertainty around each calculated value (±δ), which is used to estimate lower and upper bounds in initial domains. Two groups of problems are solved: 1) with δ = 15% (bottom set of panels) and 2) with an adjusted δ (δ = 0.000001%) (upper set of panels). All fold changes in fluxes were measured, but only the lower and upper bounds of the initial domain for changes in HK were defined. The remaining reaction fluxes for GPI, PFK and ALD do not require to be restricted to the measured values because are interdependent ($J_{HK} = J_{GPI} = J_{PFK} = J_{ALD}$). These remaining fluxes and all enzyme activities were restricted to a range of −4 to +4. See Table S1 for detailed initial and final domains.

Although these points displayed in the upper panels could be directly solved without the need to use the inference approach based on LP, as shown in the lower panels, this approach is efficient and permits to analyse the sensitivity of the predictions to the uncertainty around these points, therefore providing a more complete understanding. While the upper panels tell us about the deviation from expected changes, the lower panels inform us about the degree of propagation of the uncertainty associated with each measured flux or concentration. A higher uncertainty propagation should be due to a higher sensitivity of the predictions to uncertainty.

In the left and right panels (Fig. 4b), fixed control coefficients correspond to the initial control coefficients (before perturbation) and final control coefficients (after perturbation). In the central panels, an intermediate pattern is used (mean control coefficients), which is the result of calculating for each control coefficient the mean between the initial control coefficient and the final control coefficient:

$$\overline{C}^{x_i}_{v_k} = \left(C^{x_i}_{v_k} initial + C^{x_i}_{v_k} final\right) \Big/ 2 \tag{13}$$

$$\overline{C}^{J_j}_{v_k} = \left(C^{J_j}_{v_k} initial + C^{J_j}_{v_k} final\right) \Big/ 2 \tag{14}$$

Results obtained using Model I-A as a case example and perturbing only PFK enzyme activity showed that either initial control coefficients (Fig. 4a) or final control coefficients (Fig. 4b) permit to successfully identify PFK as the expected driver of the metabolic adaptation (blue arrow). Consistently, an approximate tenfold decrease is observed. The expected pattern of changes should be exactly zero for the rest of the enzyme activities. However, a slight deviation for HK and a strong deviation for GPI (yellow arrow) is observed. Additionally, a higher propagation of uncertainty is observed for GPI activity in both scenarios,

when the uncertainty (δ) is set to 15% for concentrations and fluxes. The consequence is a higher difficulty (requiring a reduction of the value of δ) in displaying the estimated fold change as a point. We hypothesize that these deviations and the increased propagation of uncertainty, which are consequences of control redistribution during the nonlinear adaptation to the perturbation, can be avoided using mean control coefficients. According to this hypothesis, using mean control coefficients (Fig. 4c), we successfully identified PFK as the only driver (blue arrow) and significantly reduced the cumulative deviation (η).

Although the degree of deviation is lower, these observations are generally consistent in the analyses for identification of key drivers in Figs. S1 and S2, where the effects of changing the concentrations of HK, GPI, and ALD are described. The lower degree of deviation is for GPI perturbation, which has very low flux control coefficients (Fig. 1) and therefore exhibits a lower degree of control redistribution. This situation is taken to the extreme when we consider a proportional decrease or increase in all enzymatic activities at the same time, as shown in Fig. S3. Multisite modulation of all enzyme activities is associated with reduced control redistribution, which is null when all activities are changed in the same proportion (Kacser and Acerenza, 1993; Fell and Thomas, 1995).

Finally, we repeated the same analyses using the larger and more complex Model II, which covers the central metabolism of *Escherichia coli* (Millard et al., 2017). These analyses produced results that are consistent with those obtained using the simpler Model I. As shown in Fig. 5 and S6, both a tenfold increase in pyruvate kinase (PYK) activity and a simultaneous twofold decrease in the levels of PFK, PYK, pyruvate dehydrogenase (PDH) and glutamate dehydrogenase (GDH) activities, respectively, were consistently inferred (blue arrows). However, unexpected deviations (yellow arrows), along with a high propagation of uncertainty, are also observed. As noted in the figures, where the control coefficients are ordered according to their magnitudes, this is a tendency for activities with lower flux control coefficients. Although a slightly different order is obtained for concentration control coefficients, the same conclusion can be drawn (Figs. S5 and S7). All of this suggests that in practical exercises for identifying key drivers, a good strategy might be, as a first step, to rule out those enzyme activities associated with low control distribution, such as GPI in Model I-A. As shown in Fig. 5c, S5c, S6c and S7c, this successfully avoids any situation of deviation or high propagation of uncertainty, improving the quality of the analysis.

4. Discussion

GSMMs (Ebrahim et al., 2013; Wang et al., 2018; Heirendt et al., 2019) comprise a family of models that account mainly for stoichiometry. Although they can integrate various omics data, they do not take into account kinetic laws and regulation, unlike kinetic models such as Models I and II. GSMMs have succeeded in identifying gene knock-out targets, in addition to up- or downregulations, required as part of metabolic manipulation strategies (Zhang and Hua, 2015; Gu et al., 2019; Fang et al., 2020; Razaghi-Moghadam and Nikoloski, 2021; Maranas and Zomorrodi, 2016). In addition, they have even been applied to unveil the important role of post-transcriptional events, such as regulation of enzyme activity by covalent modifications (e.g., phosphorylation, methylation, acetylation), beyond changes in enzyme expression (Katzir et al., 2019). Computation of metabolic control coefficients, although restricted to a subnetwork, can be complementary to GSMMs to unveil the effect of these manipulations and to quantify the control. Indeed, MCA has demonstrated that the most effective approach is to target enzymes with high control coefficients, because less concentrations of inhibitors will be required to achieve the desired effect (Cascante et al., 2002; Moreno-Sanchez et al., 2008). For example, using MCA, glutathione S-transferase has been identified as an attractive co-target for enhancing the effectiveness of Sorafenib, an approved drug for treating hepatocellular carcinoma (Mishra et al., 2018). However, as analysed in this paper, when facing metabolic adaptations to substantial changes, such as those resulting from drug administration, precautions must be taken. On one hand, we have control redistribution, which becomes more pronounced as the disturbance becomes larger and more complex. On the other hand, perhaps more importantly, there is the partial knowledge of control determinants involved in metabolic regulation, such as negative feedback regulation circuits or the presence of isoenzymes.

Regarding the degree of control redistribution, as explored in Fig. 1, the accuracy of control coefficients in predicting changes is limited in some cases to relatively small changes in enzyme concentrations, as significant enzyme changes result in the redistribution of control coefficient values along the metabolic pathway. Accordingly, as explored first in the simpler Model I-A (Fig. 4, S1, S2 and S3), and also in the more complex Model II (Fig. 5, S5, S6 and S7), the subsequent identification of key drivers sustaining metabolic adaptations may encounter unexpected deviations, along with a high propagation of uncertainty, in particular for activities associated with low control coefficients, as emphasized in figures for Model II. As mentioned before, this significant propagation of uncertainty is likely due to a higher sensitivity of the predictions to uncertainty. To address these bottlenecks, we have developed a successful workflow that involves the use of mean control coefficients and ruling out enzyme activities that are generally associated with low control coefficients, such as epimerases and isomerases, which are often near equilibrium. This makes sense considering that these are low-value control coefficients, and therefore it is unlikely that they have been the drivers of a mechanism-unknown perturbation. The use of mean control coefficients should account for the system's behaviour during the transient state between the initial and final stages, as described by the initial and final control coefficients, respectively. Assuming that both the initial and final control coefficients are available, this approach should extend the applicability of the proposed workflow beyond simply considering the initial or final set of control coefficients. The applicability of the workflow also depends on having available measurements of fluxes and concentrations at two steady states, along with fixed values for control coefficients. This application of the workflow should be viewed as a proof of concept, as we derive the constant control coefficients and known fold changes from complete kinetic models, which are assumed to accurately represent the actual system, with all kinetic control determinants implicitly contained in the rate laws and regulatory circuits of the model. As long as we have a kinetic model capable of adjusting maximum velocities (i.e., enzyme activities) to match the measured concentrations and fluxes, our approach may not offer additional benefits when the goal is to detect changes in enzyme activities. It was designed for use in scenarios of uncertainty about the kinetic mechanisms, where sets of control coefficients were estimated using sampling methods (de Atauri et al., 2021), and also in scenarios of uncertainty about measurements of fluxes, concentrations or enzyme activities. The approach applied in this study is applicable even if some of the fold changes required for a single solution are not strictly defined. A partial set of fold changes might suffice to identify some of the unknown variables. As shown in Fig. S4, the decrease in PYK activity continues to be detected, even when the analysis in Fig. 1 is repeated with the fold change for G6P ignored. In terms of the sensitivity of the predictions to uncertainty, the analysis of Model II demonstrates that the propagation of uncertainty is low enough to permit the application of the workflow in central carbon metabolism scale models. In an ideal context, emerging technologies may allow for the experimental measurement of control coefficients. On the one hand, metabolomics methods facilitate large-scale measurement of changes in the concentrations of metabolites (Johnson et al., 2016). On the other hand, continuous advancements in ^{13}C-based Metabolic Flux Analysis could provide us with reliable flux measurements (Antoniewicz, 2021). Although the scale is currently far from that of GSMMs, these data could be employed to provide a direct measurement of control coefficients by applying established strategies in MCA, such as co-response analysis (Hofmeyr and Cornish-Bowden, 1996). This approach does not require information about the kinetic

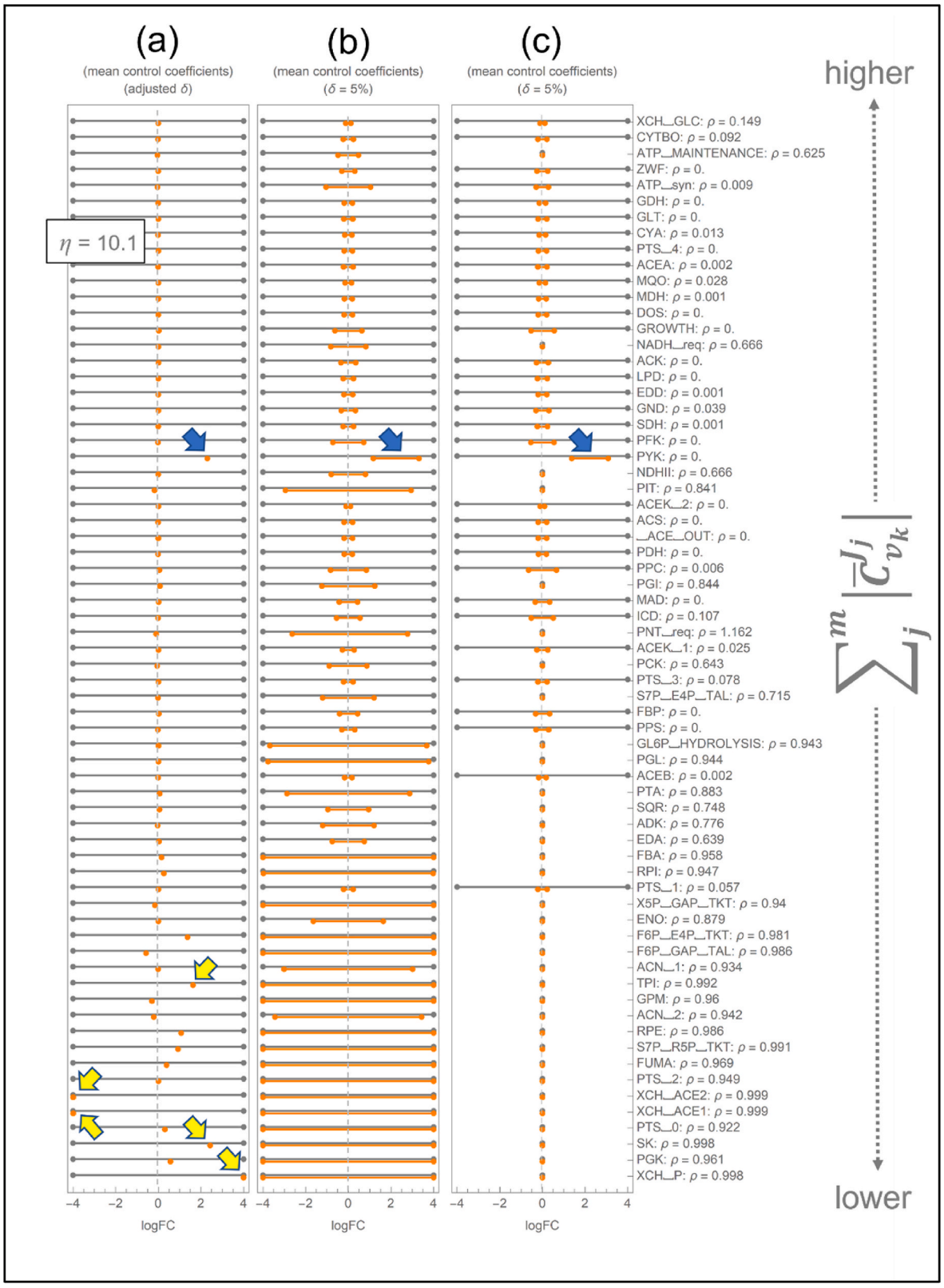


Fig. 5. Identification of key drivers in Model II. PYK increase. Initial domains (grey) and final domains for log fold changes (logFC) in molecular enzyme activities (v_k; $k = 1, \ldots, m$) (orange). The fixed control coefficients applied are the mean values (mean control coefficients). The enzyme activities are ordered according to the magnitudes of the flux control coefficients. η denotes the cumulative deviation (a) and the blue arrow indicates the expected change, while yellow arrows show the more relevant unexpected deviations. A tenfold increase in the level of PYK activity is expected based on the pattern of fold changes measured in all concentrations and fluxes. Three scenarios are presented: 1) with an adjusted δ (δ = 0.32567242%) (a), 2) with δ = 5% (b), and 3) with δ = 5% and the fold changes for enzyme activities catalysing reactions close to equilibrium (ρ > 0.5) having lower and upper bounds set to zero in the initial domains (c). Detailed initial and final domains for enzyme activities, along with those for concentrations and reaction fluxes, can be found in Table S2.

mechanisms and is based on experimental measurements. As an example, Model I was originally developed to compare calculated control coefficients from those derived experimentally by co-response analysis (Puigjaner et al., 1997).

Regarding the partial knowledge of control determinants in GSMMs, we have analysed the effects of a heterogeneous list of determinants, including the distance to the equilibrium, the degree of enzyme activity saturation, metabolite feedback inhibition loops, presence of isoenzymes, and multisite modulations, using Figs. 1–3. Other analyses may serve to illustrate the importance of accurately identifying these control determinants. Considering additional input and output reactions can have a significant impact on the control distribution (de Atauri et al., 2002), especially if they involve irreversible non-regulated steps, such as exchange fluxes with fixed values. Therefore, an appropriate selection of inputs and outputs is required. Millard et al. (2017) developed his model (Model II) to investigate the impact of metabolite-enzyme interactions on the regulation of metabolism, revealing that kinetic characteristics, such as modulation of enzyme activity by allosteric effectors and feedback/feedforward loops of metabolite inhibition/activation, may be just as important as network stoichiometry in determining the key players in metabolic regulation. Another example concerns metabolic adaptations that involve sensitive gene regulatory loops and hormone-based regulations. There are enzymes that acquire additional non-metabolic functions (known as moonlighting functions) that are not accounted for in the existing metabolic models. In this context, it has been proposed that a feedback mechanism exists between cellular metabolism and gene expression, whereby glycolytic and gluconeogenic enzymes in tumour cells acquire additional functions and directly regulate gene expression (Bian et al., 2022).

On the contrary, some mechanisms could serve to maintain control, as we have seen with isoenzymes. Here, we hypothesize that the pattern of isoenzymes expressed in a cell may result from evolutionary optimization under the selective pressure of maintaining metabolic flexibility and adapting to different environments with varying nutrient availability. For instance, we used Model I-A (which includes only HK II isoenzyme) and Model I-E (which includes HK I and HK II isoenzymes catalysing the same reaction) to explore the role of isoenzyme expression patterns in maintaining sustained control over metabolic pathway flux. The results obtained, illustrated in Fig. 3, reveal that HK isoenzymes could play a pivotal role in maintaining control coefficient distribution, intermediate metabolite homeostasis, and sustained glycolytic flux over a wide range of glucose concentrations. This would ensure optimal adaptation to different environmental challenges and resistance to drug interventions such as those based on 2-deoxyglucose. Finally, there are also enzymatic kinetic mechanisms that have been proposed to help maintain control in the face of perturbations, such as the so-called paradoxical or sustained control. This is associated with enzymes that exhibit allosteric kinetics, like glucokinase (de Atauri et al., 2001; Ortega et al., 2008).

In summary, we demonstrate that identification of kinetic determinants is essential for the inference of key enzymes that drive metabolic adaptations. Moreover, our MCA-based analysis also demonstrated that the inference of these putative drug targets can be dramatically improved using mean control coefficients and excluding those enzyme activities that are associated with low control coefficients. We envision that by incorporating MCA into GSMM workflows, even with MCA focused on only one subpart of the metabolism analysed by GSMM, researchers can obtain a more comprehensive understanding of metabolic regulation. We anticipate that this combined approach will enable a more thorough exploration of biological systems, boosting the development of effective treatments and advancing the fields of systems medicine and biotechnology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Funding: This work was supported by grant PID 2020-115051RB-I00 funded by MCIN/AEI/10.13039/501100011033, the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR, Generalitat de Catalunya) [2021SGR00350] and CIBERehd (ISCIII. Spain) [CB17/04/00023]. M.C. also received support through the prize "ICREA Academia" for excellence in research, funded by the ICREA foundation–Generalitat de Catalunya. C.F. is funded by a BHF Programme Grant [RG/18/13/33946].

Appendix

Equations and parameters in model I-A

$$1\ v_{HK} = Vmax\frac{\frac{Glc}{Ki}}{Glc\left(1+\frac{G6P}{Ki}\right)+Km}, Vmax = 63\ nmol\ min^{-1}mg\ prot^{-1}, Glc = 10\ mM, Km = 0.4019\ mM, Ki = 0.111\ mM.$$

$$2\ v_{GPI} = \frac{\frac{Vmax_f\times G6P}{Kms}-\frac{Vmax_r\times F6P}{Kmp}}{1+\frac{G6P}{Kms}+\frac{F6P}{Kmp}}, Vmax_f = 12474\ nmol\ min^{-1}mg\ prot^{-1}, Vmax_r = 18125\ nmol\ min^{-1}mg\ prot^{-1}, Kms = 0.48\ mM, Kmp = 0.272\ mM, .$$

$$3\ v_{PFK} = \frac{Vmax_f\ \frac{F6P}{K0.5s}\left(1-\frac{FbP}{F6P\times K}\right)\left(\frac{F6P}{K0.5s}+\frac{FbP}{K0.5p}\right)^{h-1}}{1+\left(\frac{F6P}{K0.5s}+\frac{FbP}{K0.5p}\right)^{h}}, Vmax_f = 1115.53\ nmol\ min^{-1}mg\ prot^{-1}, K0.5s = 0.061\ mM, K0.5p = 0.000804\ mM, h = 1.4744, K = 21724, .$$

$$4\ v_{ALD} = Vmax\frac{\frac{Glc}{Ki}}{Glc\left(1+\frac{G6P}{Ki}\right)+Km}, Vmax = 1500\ nmol\ min^{-1}mg\ prot^{-1}, Km = 0.1297\ mM.$$

Equation for the first step in model I–B

$$1\ v_{HK} = Vmax\frac{\frac{Glc}{Ki}}{Glc\left(1+\frac{F6P}{Ki}\right)+Km}, Vmax = 63\ nmol\ min^{-1}mg\ prot^{-1}, Glc = 10\ mM, Km = 0.4019\ mM, Ki = 0.0410609\ mM.$$

Ki is adjusted to maintain the steady-state ratio of *[inhibitor]* to *Ki*, while also keeping the apparent *Km* and *Vmax* values the same as those in the original Model I-A.

Equation for the first step in model I–C

$$1\ v_{HK} = Vmax\frac{\frac{Glc}{Ki}}{Glc\left(1+\frac{FbP}{Ki}\right)+Km}, Vmax = 63\ nmol\ min^{-1}mg\ prot^{-1}, Glc = 10\ mM, Km = 0.4019\ mM, Ki = 0.0116912\ mM.$$

Ki is adjusted to maintain the steady-state ratio of *[inhibitor]* to *Ki*, while also keeping the apparent *Km* and *Vmax* values the same as those in the original Model I-A.

Equation for the first step in model I-D

1 $v_{HK} = Vmax\frac{Glc}{Glc+Km}, Vmax = 46.7878\ nmol\ min^{-1} mg\ prot^{-1}, Glc = 10\ mM, Km = 0.298477\ mM.$

Km and *Vmax* are adjusted to match the same apparent values as in Model I-A.

Equation for the first step describing the two isoenzymes in model I-E

1 $v_{HK} = 0.75 \times Vmax\frac{Glc}{Glc\ \left(1+\frac{G6P}{Kia}\right)+Kma} + 0.25 \times Vmax\frac{Glc}{Glc\ \left(1+\frac{G6P}{Kib}\right)+Kmb}, Vmax = 61.6125\ nmol\ min^{-1} mg\ prot^{-1}, Kma = 0.04\ mM, Kmb = 0.4\ mM, Kia = 0.111\ mM, Kib = 0.111\ mM.$

To produce three additional model variants – models I-A′, I–B′, and I–C' – with reduced inhibition constants (*Ki*), these are divided by 100 to *Ki* = 0.00111 mM, *Ki* = 0.000410609 mM, and *Ki* = 0.000116912 mM, respectively, therefore multiplying by 100 the original steady state *[inhibitor]/Ki* ratio in HK. To ensure that the *Vmax* and *Km* values for HK in the three variants matched those of model I-A, *Vmax* are set to 1668 nmol min^{-1} mg $prot^{-1}$, and *Km* are set to 10.6408 mM.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biosystems.2023.104984. Multimedia component 1 contains Figs. S1 to S7; Multimedia component 2 contains Table S1; and Multimedia component 3 contains Table S2.

References

Acerenza, L., 2000. Design of large metabolic responses. Constraints and sensitivity analysis. J. Theor. Biol. 207, 265–282. https://doi.org/10.1006/jtbi.2000.2173.

Acerenza, L., Ortega, F., 2007. Modular metabolic control analysis of large responses. FEBS J. 274, 188–201. https://doi.org/10.1111/j.1742-4658.2006.05575.x.

Acerenza, L., Monzon, P., Ortega, F., 2015. A modular modulation method for achieving increases in metabolite production. Biotechnol. Prog. 31, 656–667. https://doi.org/10.1002/btpr.2059.

Antolin, A.A., Cascante, M., 2021. AI delivers Michaelis constants as fuel for genome-scale metabolic models. PLoS Biol. 19, e3001415 https://doi.org/10.1371/journal.pbio.3001415.

Antoniewicz, M.R., 2021. A guide to metabolic flux analysis in metabolic engineering: methods, tools and applications. Metab. Eng. 63, 2–12. https://doi.org/10.1016/j.ymben.2020.11.002.

Bian, X., Jiang, H., Meng, Y., Li, Y., Fang, J., Lu, Z., 2022. Regulation of gene expression by glycolytic and gluconeogenic enzymes. Trends Cell Biol. 32, 786–799. https://doi.org/10.1016/j.tcb.2022.02.003.

Cascante, M., Franco, R., Canela, E.I., 1989a. Use of implicit methods from general sensitivity theory to develop a systematic approach to metabolic control. I. Unbranched pathways. Math. Biosci. 94, 271–288.

Cascante, M., Franco, R., Canela, E.I., 1989b. Use of implicit methods from general sensitivity theory to develop a systematic approach to metabolic control. II. Complex systems. Math. Biosci. 94, 289–309.

Cascante, M., Boros, L.G., Comin-Anduix, B., de Atauri, P., Centelles, J.J., Lee, P.W.-N., 2002. Metabolic control analysis in drug discovery and disease. Nat. Biotechnol. 20, 243–249. https://doi.org/10.1038/nbt0302-243.

Cornish-Bowden, A., 2012. Fundamentals of Enzyme Kinetics, fourth ed. Wiley-VCH, Weinheim.

de Atauri, P., Acerenza, L., Kholodenko, B.N., de la Iglesia, N., Guinovart, J.J., Agius, L., Cascante, M., 2001. Occurrence of paradoxical or sustained control by an enzyme when overexpressed: necessary conditions and experimental evidence with regard to hepatic glucokinase. Biochem. J. 355, 787–793. https://doi.org/10.1042/bj3550787.

de Atauri, P., Fell, D.A., Chassagnole, C., Magret, D., Mazat, J.P., Cascante, M., 2002. Dependence of control coefficient distribution on the boundaries of a metabolic system: a generalized analysis of the effects of additional input and output reactions to a linear pathway. J. Theor. Biol. 215, 239–251. https://doi.org/10.1006/jtbi.2001.2492.

de Atauri, P., Tarrado-Castellarnau, M., Tarragó-Celada, J.T., Foguet, C., Karakitsou, E., Centelles, J.J., Cascante, M., 2021. Integrating systemic and molecular levels to infer key drivers sustaining metabolic adaptations. PLoS Comput. Biol. 17, e1009234 https://doi.org/10.1371/journal.pcbi.1009234.

de Atauri, P., Tarrado-Castellarnau, M., Tarragó-Celada, J., Foguet, C., Karakitsou, E., Centelles, J.J., Cascante, M., 2021b. Mathematica Notebooks supporting "Integrating systemic and molecular levels to infer key drivers sustaining metabolic adaptations.". Zenodo. https://doi.org/10.5281/zenodo.5081161.

Domenzain, I., Sánchez, B., Anton, M., Kerkhoven, E.J., Millán-Oropeza, A., Henry, C., Siewers, V., Morrissey, J.P., Sonnenschein, N., Nielsen, J., 2022. Reconstruction of a catalogue of genome-scale metabolic models with enzymatic constraints using GECKO 2.0. Nat. Commun. 13, 3766. https://doi.org/10.1038/s41467-022-31421-1.

Ebrahim, A., Lerman, J.A., Palsson, B.O., Hyduke, D.R., 2013. COBRApy: COnstraints-based reconstruction and analysis for Python. BMC Syst. Biol. 7, 74. https://doi.org/10.1186/1752-0509-7-74.

Fang, X., Lloyd, C.J., Palsson, B.O., 2020. Reconstructing organisms in silico: genome-scale models and their emerging applications. Nat. Rev. Microbiol. 18, 731–743. https://doi.org/10.1038/s41579-020-00440-4.

Fell, D.A., 1997. Understanding the Control of Metabolism. Portland Press, London.

Fell, D.A., Sauro, H.M., 1985. Metabolic control and its analysis. Eur. J. Biochem. 148, 555–561. https://doi.org/10.1111/j.1432-1033.1985.tb08876.x.

Fell, D.A., Thomas, S., 1995. Physiological control of metabolic flux: the requirement for multisite modulation. Biochem. J. 311, 35–39.

Fernandes, P.M., Kinkead, J., McNae, I.W., Bringaud, F., Michels, P.A.M., Walkinshaw, M.D., 2019. The kinetic characteristics of human and trypanosomatid phosphofructokinases for the reverse reaction. Biochem. J. 476, 179–191. https://doi.org/10.1042/BCJ20180635.

Gu, C., Kim, G.B., Kim, W.J., Kim, H.U., Lee, S.Y., 2019. Current status and applications of genome-scale metabolic models. Genome Biol. 20, 121. https://doi.org/10.1186/s13059-019-1730-3.

Harding, J.J., Lowery, M.A., Shih, A.H., Schvartzman, J.M., Hou, S., Famulare, C., Patel, M., Roshal, M., Do, R.K., Zehir, A., You, D., Selcuklu, S.D., Viale, A., Tallman, M.S., Hyman, D.M., Reznik, E., Finley, L.W.S., Papaemmanuil, E., Tosolini, A., et al., 2018. Isoform switching as a mechanism of acquired resistance to mutant isocitrate dehydrogenase inhibition. Cancer Discov. 8, 1540–1547. https://doi.org/10.1158/2159-8290.CD-18-0877.

Heinrich, R., Rapoport, T.A., 1974. A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. Eur. J. Biochem. 42, 89–95.

Heinrich, R., Schuster, S., 1996. The Regulation of Cellular Systems. Chapman and Hall, New York.

Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S.N., Richelle, A., Heinken, A., Haraldsdóttir, H.S., Wachowiak, J., Keating, S.M., Vlasov, V., Magnusdóttir, S., Ng, C.Y., Preciat, G., Žagare, A., Chan, S.H.J., Aurich, M.K., Clancy, C.M., Modamio, J., Sauls, J.T., et al., 2019. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. Nat. Protoc. 14, 639–702. https://doi.org/10.1038/s41596-018-0098-2.

Hofmeyr, J.-H.S., 1995. Metabolic regulation: a control analytic perspective. J. Bioenerg. Biomembr. 27 (5), 479–490. https://doi.org/10.1007/BF02110188.

Hofmeyr, J.-H.S., Cornish-Bowden, A., 1991. Quantitative assessment of regulation in metabolic systems. Eur. J. Biochem. 200, 223–236. https://doi.org/10.1111/j.1432-1033.1991.tb21071.x.

Hofmeyr, J.-H.S., Cornish-Bowden, A., 1996. Co-Response analysis: a new experimental strategy for metabolic control analysis. J. Theor. Biol. 182, 371–380.

Hofmeyr, J.H., Cornish-Bowden, A., 1997. The reversible Hill equation: how to incorporate cooperative enzymes into metabolic models. CABIOS 13, 377–385. https://doi.org/10.1093/bioinformatics/13.4.377.

Johnson, C.H., Ivanisevic, J., Siuzdak, G., 2016. Metabolomics: beyond biomarkers and towards mechanisms. Nat. Rev. Mol. Cell Biol. 17, 451–459. https://doi.org/10.1038/nrm.2016.25.

Kacser, H., Acerenza, L., 1993. A universal method for achieving increases in metabolite production. Eur. J. Biochem. 216, 361–367.
Kacser, H., Burns, J.A., 1973. The control of flux. Symp. Soc. Exp. Biol. 27, 65–104.
Katzir, R., Polat, I.H., Harel, M., Katz, S., Foguet, C., Selivanov, V.A., Sabatier, P., Cascante, M., Geiger, T., Ruppin, E., 2019. The landscape of tiered regulation of breast cancer cell metabolism. Sci. Rep. 9, 17760 https://doi.org/10.1038/s41598-019-54221-y.
Kholodenko, B.N., 1988. How do external parameters control fluxes and concentrations of metabolites? An additional relationship in the theory of metabolic control. FEBS Lett. 232, 383–386. https://doi.org/10.1016/0014-5793(88)80775-0.
Kroll, A., Engqvist, M.K.M., Heckmann, D., Lercher, M.J., 2021. Deep learning allows genome-scale prediction of Michaelis constants from structural features. PLoS Biol. 19, e3001402 https://doi.org/10.1371/journal.pbio.3001402.
Maranas, C.D., Zomorrodi, A.R., 2016. Optimization Methods in Metabolic Networks. Wiley, Hoboken, New Jersey. https://doi.org/10.1002/9781119188902.
Marcucci, F., Rumio, C., 2021. Glycolysis-induced drug resistance in tumors—a response to danger signals? Neoplasia 23, 234–245. https://doi.org/10.1016/j.neo.2020.12.009.
Martín-Bernabé, A., Balcells, C., Tarragó-Celada, J., Foguet, C., Bourgoin-Voillard, S., Seve, M., Cascante, M., 2017. The importance of post-translational modifications in systems biology approaches to identify therapeutic targets in cancer metabolism. Curr. Opin. Syst. Biol. 3, 161–169. https://doi.org/10.1016/j.coisb.2017.05.011.
Millard, P., Smallbone, K., Mendes, P., 2017. Metabolic regulation is sufficient for global and robust coordination of glucose uptake, catabolism, energy production and growth in Escherichia coli. PLoS Comput. Biol. 13, e1005396 https://doi.org/10.1371/journal.pcbi.1005396.
Mishra, M., Jayal, P., Karande, A.A., Chandra, N., 2018. Identification of a co-target for enhancing efficacy of sorafenib in HCC through a quantitative modeling approach. FEBS J. 285, 3977–3992. https://doi.org/10.1111/febs.14641.
Miskovic, L., Hatzimanikatis, V., 2010. Production of biofuels and biochemicals: in need of an ORACLE. Trends Biotechnol. 28, 391–397. https://doi.org/10.1016/j.tibtech.2010.05.003.
Moreno-Sanchez, R., Saavedra, E., Rodriguez-Enriquez, S., Olin-Sandoval, V., 2008. Metabolic control analysis: a tool for designing strategies to manipulate metabolic pathways. J. Biomed. Biotechnol. 2008, 597913.
Nilsson, A., Nielsen, J., 2016. Metabolic trade-offs in yeast are caused by F1F0-ATP synthase. Sci. Rep. 6, 22264 https://doi.org/10.1038/srep22264.
Ortega, F., Acerenza, L., 2011. Modular metabolic control analysis of large responses in branched systems - application to aspartate metabolism. FEBS J. 278, 2565–2578. https://doi.org/10.1111/j.1742-4658.2011.08184.x.
Ortega, F., Cascante, M., Acerenza, L., 2008. Kinetic properties required for sustained or paradoxical control of metabolic fluxes under large changes in enzyme activities. J. Theor. Biol. 252, 569–573.
Perrin-Cocon, L., Vidalain, P.-O., Jacquemin, C., Aublin-Gex, A., Olmstead, K., Panthu, B., Rautureau, G.J.P., André, P., Nyczka, P., Hütt, M.-T., Amoedo, N., Rossignol, R., Filipp, F.V., Lotteau, V., Diaz, O., 2021. A hexokinase isoenzyme switch in human liver cancer cells promotes lipogenesis and enhances innate immunity. Commun. Biol. 4, 217. https://doi.org/10.1038/s42003-021-01749-3.
Puigjaner, J., Raïs, B., Burgos, M., Comín, B., Ovadi, J., Cascante, M., 1997. Comparison of control analysis data using different approaches: modelling and experiments with muscle extract. FEBS Lett. 418, 47–52.
Razaghi-Moghadam, Z., Nikoloski, Z., 2021. GeneReg: a constraint-based approach for design of feasible metabolic engineering strategies at the gene level. Bioinformatics 37, 1717–1723. https://doi.org/10.1093/bioinformatics/btaa996.
Reder, C., 1988. Metabolic control theory: a structural approach. J. Theor. Biol. 135, 175–201. https://doi.org/10.1016/S0022-5193(88)80073-0.
Ritov, V.B., Kelley, D.E., 2001. Hexokinase isozyme distribution in human skeletal muscle. Diabetes 50, 1253–1262. http://diabetesjournals.org/diabetes/article-pdf/50/6/1253/654477/1253.pdf.
Rohwer, J.M., Hofmeyr, J.-H., 2010. Kinetic and thermodynamic aspects of enzyme control and regulation. J. Phys. Chem. B 114, 16280–16289.
Saikiran Reddy, M., Bhattacharjee, D., Jain, N., 2022. Plk1 regulates mutant IDH1 enzyme activity and mutant IDH2 ubiquitination in mitosis. Cell. Signal. 92, 110279 https://doi.org/10.1016/j.cellsig.2022.110279.
Sánchez, B.J., Zhang, C., Nilsson, A., Lahtvee, P., Kerkhoven, E.J., Nielsen, J., 2017. Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. Mol. Syst. Biol. 13, 935. https://doi.org/10.15252/msb.20167411.
Sauro, H., 2019. Systems Biology: an Introduction to Metabolic Control Analysis. Ambrosius Publishing, Seattle.
Small, J.R., Fell, D.A., 1989. The matrix method of metabolic control analysis: its validity for complex pathway structures. J. Theor. Biol. 136, 181–197. https://doi.org/10.1016/S0022-5193(89)80225-5.
Small, J.R., Kacser, H., 1993a. Responses of metabolic systems to large changes in enzyme activities and effectors. 1. The linear treatment of unbranched chains. Eur. J. Biochem. 213, 613–624. https://doi.org/10.1111/j.1432-1033.1993.tb17801.x.
Small, J.R., Kacser, H., 1993b. Responses of metabolic systems to large changes in enzyme activities and effectors. 2. The linear treatment of branched pathways and metabolite concentrations. Assessment of the general non-linear case. Eur. J. Biochem. 213, 625–640. https://doi.org/10.1111/j.1432-1033.1993.tb17802.x.
Tseng, P.-L., Chen, C.-W., Hu, K.-H., Cheng, H.-C., Lin, Y.-H., Tsai, W.-H., Cheng, T.-J., Wu, W.-H., Yeh, C.-W., Lin, C.-C., Tsai, H.-J., Chang, H.-C., Chuang, J.-H., Shan, Y.-S., Chang, W.-T., 2018. The decrease of glycolytic enzyme hexokinase 1 accelerates tumor malignancy via deregulating energy metabolism but sensitizes cancer cells to 2-deoxyglucose inhibition. Oncotarget 9, 18949–18969. https://doi.org/10.18632/oncotarget.24855.
Varghese, E., Samuel, S.M., Líšková, A., Samec, M., Kubatka, P., Büsselberg, D., 2020. Targeting glucose metabolism to overcome resistance to anticancer chemotherapy in breast cancer. Cancers 12, 2252. https://doi.org/10.3390/cancers12082252.
Wang, H., Marcišauskas, S., Sánchez, B.J., Domenzain, I., Hermansson, D., Agren, R., Nielsen, J., Kerkhoven, E.J., 2018. Raven 2.0: a versatile toolbox for metabolic network reconstruction and a case study on Streptomyces coelicolor. PLoS Comput. Biol. 14, e1006541 https://doi.org/10.1371/journal.pcbi.1006541.
Westerhoff, H.V., Kell, D.B., 1987. Matrix method for determining steps most rate-limiting to metabolic fluxes in biotechnological processes. Biotechnol. Bioeng. 30, 101–107. https://doi.org/10.1002/bit.260300115.
Westerhoff, H.v., Groen, A.K., Wanders, R.J.A., 1984. Modern theories of metabolic control and their applications. Biosci. Rep. 4, 1–22. https://doi.org/10.1007/BF01120819.
Wilken, S.E., Besançon, M., Kratochvíl, M., Foko Kuate, C.A., Trefois, C., Gu, W., Ebenhöh, O., 2022. Interrogating the effect of enzyme kinetics on metabolism using differentiable constraint-based models. Metab. Eng. 74, 72–82. https://doi.org/10.1016/j.ymben.2022.09.002.
Wilson, J.E., 2003. Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function. J. Exp. Biol. 206, 2049–2057. https://doi.org/10.1242/jeb.00241.
Zhang, C., Hua, Q., 2015. Applications of genome-scale metabolic models in biotechnology and systems medicine. Front. Physiol. 6, 413. https://doi.org/10.3389/fphys.2015.00413.