FEBS Letters 587 (2013) 2832–2841

FEBS Letters

journal homepage: www.FEBSLetters.org

A model of yeast glycolysis based on a consistent kinetic characterisation of all its enzymes ☆

Kieran Smallbone [a,b,c,1], Hanan L. Messiha [a,d,e,1], Kathleen M. Carroll [a,d,2], Catherine L. Winder [a,d,2], Naglis Malys [a,e,g,2], Warwick B. Dunn [a,d,f], Ettore Murabito [a,h], Neil Swainston [a,b,d], Joseph O. Dada [a,b], Farid Khan [a,d], Pınar Pir [a,e,i,s], Evangelos Simeonidis [a,h,j,k], Irena Spasić [a,b,l], Jill Wishart [a,e], Dieter Weichart [a,d,m], Neil W. Hayes [a,h,n], Daniel Jameson [a,b,h,o], David S. Broomhead [a,c], Stephen G. Oliver [a,e,i], Simon J. Gaskell [a,d,p], John E.G. McCarthy [a,d,e,h], Norman W. Paton [a,b], Hans V. Westerhoff [a,h,q], Douglas B. Kell [a,d], Pedro Mendes [a,b,r,*]

[a] Manchester Centre for Integrative Systems Biology, Manchester Institute of Biotechnology, The University of Manchester, UK
[b] School of Computer Science, The University of Manchester, UK
[c] School of Mathematics, The University of Manchester, UK
[d] School of Chemistry, The University of Manchester, UK
[e] Faculty of Life Sciences, The University of Manchester, UK
[f] Centre for Advanced Discovery and Experimental Therapeutics (CADET), Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Sciences Centre, UK
[g] School of Life Sciences, The University of Warwick, UK
[h] School of Chemical Engineering and Analytical Science, The University of Manchester, UK
[i] Cambridge Systems Biology Centre & Department of Biochemistry, University of Cambridge, UK
[j] Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Luxembourg
[k] Institute for Systems Biology, Seattle, Washington, USA
[l] School of Computer Science & Informatics, Cardiff University, UK
[m] School of Dentistry, The University of Manchester, UK
[n] Research and Knowledge Transfer, University of Exeter, UK
[o] Faculty of Medical and Human Sciences, The University of Manchester, UK
[p] Queen Mary, University of London, UK
[q] Department of Molecular Cell Physiology, Vrije Universiteit Amsterdam, The Netherlands
[r] Virginia Bioinformatics Institute, Virginia Tech, USA
[s] Babraham Institute, Babraham Research Campus, UK

ARTICLE INFO

Article history:
Received 26 April 2013
Revised 24 June 2013
Accepted 25 June 2013
Available online 4 July 2013

Edited by Athel Cornish-Bowden

Keywords:
Glycolysis
Systems biology
Enzyme kinetic
Isoenzyme
Modelling

ABSTRACT

We present an experimental and computational pipeline for the generation of kinetic models of metabolism, and demonstrate its application to glycolysis in *Saccharomyces cerevisiae*. Starting from an approximate mathematical model, we employ a "cycle of knowledge" strategy, identifying the steps with most control over flux. Kinetic parameters of the individual isoenzymes within these steps are measured experimentally under a standardised set of conditions. Experimental strategies are applied to establish a set of in vivo concentrations for isoenzymes and metabolites. The data are integrated into a mathematical model that is used to predict a new set of metabolite concentrations and reevaluate the control properties of the system. This bottom-up modelling study reveals that control over the metabolic network most directly involved in yeast glycolysis is more widely distributed than previously thought.

* Corresponding author. Address: Manchester Institute of Biotechnology, 131 Princess Street, Manchester M1 7DN, UK. Fax: +44 161 3064556.
E-mail address: pedro.mendes@manchester.ac.uk (P. Mendes).

[1] These authors contributed equally to this work.
[2] These authors contributed equally to this work.

0014-5793/$36.00 © 2013 The Authors. Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.febslet.2013.06.043

1. Introduction

A major goal of systems biology is the development of mathematical models of biological phenomena that predict their behaviour accurately, that are able to provide a quantitative explanation of their mechanisms, and that have predictive power as to the effects of changes in their parameters [1–4]. For cellular phenomena, this requires models that represent the action of multiple enzymes as contributors to the behaviour observed. Fortunately, this area of systems biology is built on the strong foundations of enzymology dating from the work of Michaelis and Menten [5], with their model of the dependency of the rate of enzymatic reactions on the concentration of their substrates. Their famous rate law, and similar ones for reversible reactions and reactions with multiple substrates and products, are the building blocks for kinetic models of metabolism.

So far, the majority of mathematical models of biochemical pathways have been developed based on data collected from different sources [6]. The enzyme kinetic data used in such models have usually been measured at each enzyme's optimal pH; however the optimal pH of each enzyme in a pathway is normally different and the physiological pH and other conditions would not match the optimal conditions for each enzyme [7]. Due to the shortage of comprehensive experimental datasets acquired under specific experimental conditions, this approach may result in a distorted view, i.e., one not fully representative of the system under study. Any unknown parameters of the model are estimated through the finding of a best fit to an available set of experimental data. However, there is potentially an enormous number of parameter values providing the same degree of fitting (whatever the fitting metric), and hence the fitted model may not be a good representation of the real system and may thereby fail to predict its behaviour under different conditions. From a computational perspective this issue has been addressed through sampling approaches [8]. From an experimental perspective the optimal experimental strategy must aim to deliver as comprehensive an information content as possible under a unified set of experimental conditions that should mimic the intracellular environment as much as possible.

In kinetic modelling, the potential for alternative isoenzymes catalysing each metabolic reaction step is generally overlooked. For prokaryotic systems this is likely to be relatively unimportant since a given activity is more commonly controlled by a single gene/protein. However, in eukaryotes, there are often multiple gene/protein isoforms that may act separately and with different activities under different environmental conditions, or may occur in different cellular compartments. For example, Gu et al. [9] have estimated that there are 530, 674, and 1219 duplicate protein families in *Saccharomyces cerevisiae*, *Drosophila melanogaster* and *Caenorhabditis elegans*, respectively. These isoenzymes may play an especially important role in *S. cerevisiae* whose metabolism has evolved through whole-genome duplication and subsequent gene loss [10,11] or divergence [12]. Representation of the distinct isoenzyme activities has not been addressed in any significant kinetic models of metabolic processes to date. Thus, enzymes with identical or similar activities have mostly been integrated together (as in the model of glycolysis of Teusink et al. [13]) and measured as a lumped activity without distinguishing and appreciating the individual characteristics of each isoform. In practice, we can expect differential activities of individual isoenzymes based on changes in the environmental conditions or growth cycle.

In this paper, we introduce a novel strategy that addresses the above sources of heterogeneity by producing complete sets of experimentally-derived data using the same parental strain of *S. cerevisiae* and identical growth conditions. Furthermore, we characterise individual isoforms of enzymes involved in the biochemistry of central carbon metabolism and their differential roles.

As with any system, parameters are distinguished from variables. In a given metabolic network, the parameters are the concentrations of enzymes and their kinetic properties, while the variables are represented by fluxes and concentrations [14]. One strategy therefore isolates the individual components of a system, measures their properties (parameters) in vitro, and uses knowledge of these to reconstruct the network as a mathematical model. This model can then be (and is) used to describe the time-dependent and steady-state concentrations and fluxes in the network. Separate measurements of those variables allow one to test the precision of the model. The model can also serve to highlight potential sources of error in experimental measurements that can then be re-evaluated (e.g., Ref. [15]).

Our strategy and workflow [2] for producing and testing systems biology models is summarised in Fig. 1. We have chosen to construct mathematical models of distinct and important areas of metabolism (in terms of the amount of flux carried) and have used *S. cerevisiae* as a model system to develop and validate our approach. We stress a number of features of this approach:

1 The external conditions for the biological system (the cell culture) are well defined, and these act as a source for the metabolites that are taken up by the system.
2 A topological description of the network is then constructed, in which all metabolic reactions are defined and associated to all isoenzymes known to catalyse the reaction. This process is performed by extracting a sub-network from larger, genome-scale metabolic networks, or by compiling literature data [16–20].
3 A kinetic model is built upon the metabolic map defined in point 2. Rate laws and parameter values are either taken from pre-existing models or estimated from published data.
4 The model is evaluated through software packages such as COPASI [21] and used to predict fluxes and metabolite concentrations at steady-state.
5 The control properties of the system are evaluated computationally and the reaction step exerting the major control over the system fluxes is identified.
6 The kinetic properties of the most controlling reaction step are experimentally measured. More particularly, the turnover number (k_{cat}), concentration and affinity constants of each of the corresponding isoenzymes are measured and their value is used to refine the model.
7 Points 3–5 are repeated until all the reactions in the model have been fully characterised.
8 Once this has been performed for all the elements in the network, the predicted steady-state values of the fluxes and metabolite concentrations are compared with their corresponding experimental measurements.
9 The model's parameters are then adjusted as appropriate to minimise the mismatch between predicted and measured quantities.
10 All of the relevant data and models are deposited in suitable databases.

As with any new strategy for approaching biochemical (or other) problems, it is appropriate to calibrate or validate it by applying it to a well-understood system. Glycolysis represents an excellent test-case [13,22–24]. It is a universal metabolic pathway conserved across all three domains of life (Archaea, Bacteria and Eukarya, though with some variations) consuming glucose to produce free energy in the form of ATP and reducing equivalents in the form of NADH, and precursors that are used in other cellular metabolic processes. Glycolysis is also of applied interest for the purposes of biomass production in the baking industry [25], as a catalyst for a variety of biotransformations [26–28] and in ethanol production [29–30].