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

CrossMark

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

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.
© 2013 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

* 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.

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].

S. cerevisiae is also a particularly suitable organism in which to perform such a study from a number of other perspectives: it is
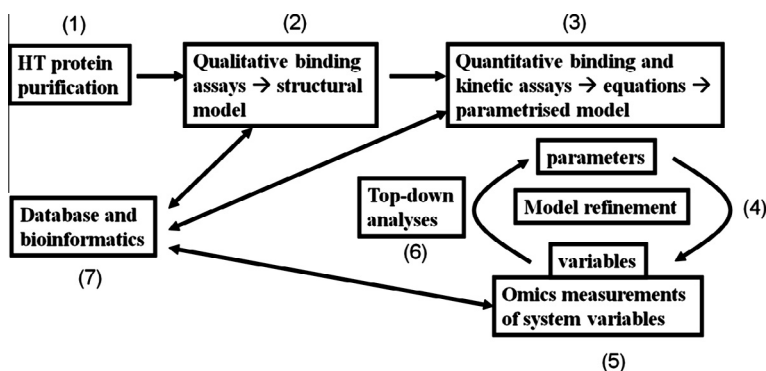
Fig. 1. The strategy for bottom-up systems biology used. The meaning of the numbers is described in the text.

well-characterised genetically, and it can be grown under conditions of continuous culture in which its metabolism can adopt a steady state [31–33]. These conditions allow multiple proteome and metabolome samples to be acquired, and ensure a high reproducibility of measurements and hence fidelity of the final results.

This article describes the overall strategy set out above and illustrates the accuracy of systems biology modelling that is currently achievable with such a strategy. This is also the first time that the individual contributions by different isoenzymes have been accounted for in such detail, whether in a model or experimentally.

2. Materials and methods

2.1. Enzyme production and purification

Enzymes were expressed in *S. cerevisiae* strains that contain either an overexpression plasmid with the open reading frame of interest, which was under control of the *GAL1* inducible promoter and in fusion to a multiple tag at the C-terminus (Yeast ORF Collection, [34]) and N-terminus (Yeast GST-Tagged Collection [35]) or a chromosomally integrated gene with a tag fusion (TAP collection [36]). All collections are available from Open Biosystems (http://www.thermoscientificbio.com/openbiosystems/). The expressions and purification of proteins was performed as described by Malys et al. [37–38]. The success of each step of purification for each protein was assessed by analysing samples from the intermediate and final protein preparations using SDS–PAGE. The amount and concentration of the purified enzyme was determined using QuantiPro™ BCA Assay Kit (Sigma–Aldrich) according to the manufacturer's recommendations. The quality of the enzyme preparation was further assessed using a 2100 Bioanalyzer (Agilent Technologies).

Although the overexpression of proteins with tags provides significant advantages by increasing throughput and protein recovery from the cell extract, it also brings some limitations related to the function and structure of enzymes. The position of the tag on the protein may interfere with its folding, multimeric complex formation, or functionality. Therefore, in some cases, as for example, enolase 2 (ENO2, EC 4.2.1.11) an alternative construct with an N-terminal tag had to be used. The use of Yeast MORF and Yeast GST-Tagged collections [34–35] in combination allows the purification of 80% of all *S. cerevisiae* proteins.

Applying a standardised approach of using a single yeast expression collection did not allow the reconstitution of fully active phosphofructokinase (EC 2.7.1.11) from individual Pfk1 and Pfk2 subunits. Therefore, in the second round of the experimental cycle, *PFK1* and *PFK2* were co-expressed and their protein products co-purified as a heterooctamer from the Yeast

TAP-Tagged collection. The activity of this complex was significantly higher than that of each isoenzyme purified alone.

2.2. Enzyme kinetic assays

To determine the kinetic parameters for individual enzymatic reactions, spectrophotometric assays were performed for the glycolytic isoenzymes measuring the consumption or production of NADH or NADPH by using one or more coupling reactions when needed. All assays were carried out in medium-throughput measurements with a BMG Labtech NOVOstar plate reader (automated fluorescence/FP/absorbance reader, Offenburg, Germany) in 384-well format plates with a 60 μl reaction volume [39]. All assays were performed in a standardised reaction buffer at 30 °C and were automated so that all reagents in the reaction buffer (including any coupling enzymes) are in 45 μl, the enzyme (to be assayed) in 5 μl and the substrate in 10 μl volumes. For each individual enzyme, the forward and the reverse reaction were assayed whenever possible. Assays for each individual enzyme were either developed or modified from previously published methodology to be compatible with the conditions of the assay reactions (e.g. pH compatibility or unavailability of commercial substrates).

All assays were coupled with enzyme(s) in which NAD(P) or NAD(P)H is a product or substrate so that its formation or consumption could be followed spectrophotometrically at 340 nm using an extinction coefficient of 6.62 mM^{-1} cm^{-1}, unless the reaction of a particular enzyme consumes or produces NADH or NADPH in which case no coupling enzymes were used. Some assays were modified by altering the concentration of coupling enzymes or other reagents to ensure that the rate measured is the rate of the reaction of interest (fully rate limiting). This is a critical step that needs special care to be taken when using coupling enzymes in assays [40].

All measurements are based on at least duplicate determination of the reaction rates at each substrate concentration. Control experiments were run in parallel to check and correct for any background activity. For each isoenzyme, the initial rates at various substrate concentrations were determined and the data obtained were analyzed by the KineticsWizard [41], fitting to Michaelis–Menten kinetics:

$$v = E\,k_{cat}S/(K_m + S).$$

Not all enzymes were found to exhibit these kinetics. As reported previously [42], triose phosphate isomerase (TPI1, EC 5.3.1.1) showed substrate inhibition in the direction of dihydroxyacetone phosphate production and was analysed using COPASI [21] according to the following rate law:

$$V = E\,k_{cat}S/(K_m + S(1 + (S/K_i)^4)).$$

2.3. Continuous cultures of S. cerevisiae

Continuous cultures of *S. cerevisiae* strain Y23925 (a heterozygous deletion derivative of the diploid BY4743: *MATa/MATα*; *his3Δ1/his3Δ1*; *ho::KanMX4/HO*; *leu2Δ0/leu2Δ0*; *LYS2/lys2Δ0*; *met15Δ0/MET15*; *ura3Δ0/ura3Δ0*) were established in a turbidostat-like approach in which the biomass was monitored by measuring the electrical capacitance of the culture [43–46]. A feedback loop was applied to regulate the addition of growth medium such that the biomass was maintained at 75% of the maximum achieved biomass yield for that medium, where a steady state was obtained (cf. [47]). The growth of the cultures was thus performed without nutrient limitation and at μmax. Samples for quantification of both protein levels and intracellular metabolite concentrations were collected from these continuous cultures as described below.

2.4. Proteomics

S. cerevisiae cells from 50 ml cultures were harvested by centrifugation for 5 min at 4000 g, and the cells mechanically disrupted using a mini bead-beater (Biospec Products Inc., Bartlesville, USA; http://www.biospec.com/) yielding the cellular cytoplasmic soluble fraction for analysis. The latter was combined with known amounts of the recombinant labelled QconCAT protein (containing diagnostic peptides for the glycolytic enzymes [48]), and co-digested to completion with trypsin. The resulting peptides were diluted and resolved over a linear incrementing solvent gradient by LC-MS using a nanoACQUITY chromatograph (Waters MS Technologies) coupled to an LTQ-Orbitrap XL (ThermoFisher Scientific). Automated data analysis and subsequent calculations were carried out using the QconCAT PrideWizard [49].

2.5. Metabolomics

For metabolomics, a quenching method was applied [50] which sampled 10 ml of culture solution into 40 ml of a 60:40 methanol/water solution stored at a temperature of −47 °C [51]. Immediate separation of cells was performed by applying centrifugation (4000 g for 5 min) followed by removal of the quenching solution. Both were stored at −80 °C prior to extraction or analysis. The intracellular metabolite pool was extracted into 1250 μl of an 80:20 water/methanol solution with cell wall disruption provided by eight freeze–thaw cycles to ensure full metabolite recovery. Samples were analysed using two analytical platforms. Fructose-1,6-bisphosphate was quantified applying UPLC-MS (Waters Acquity UPLC coupled to a ThermoFisher hybrid electrospray LTQ-Orbitrap mass spectrometer (ThermoFisher Scientific, http://www.thermofisher.com/)) and the standard addition method for accurate quantification. Other metabolites were quantified applying GC–MS (Agilent 6890 Gas Chromatograph coupled to a Leco Pegasus III electron impact-time-of-flight mass spectrometer (Leco, http://www.leco.com/)) and the external calibration method for accurate quantification. The concentrations determined in the extracted samples were reported as the number of molecules per cell.

The gas chromatographic separation of 2-phosphoglycerate and 3-phosphoglycerate is technically demanding and was not achieved during this study. Therefore, the combined concentration of both metabolites was reported rather than separate concentrations for each metabolite.

2.6. Modelling

An existing model of yeast glycolysis [23] was modified so that any fixed ('clamped') fluxes in that model were redefined as first-order in their reactants. Where possible, initial metabolite concentrations were updated according to in-house values;

otherwise recent literature data for respiro-fermentative, non-starved cells were used [24].

At each of 17 iterations, flux control coefficients [52–53] were calculated using COPASI [21]. Where possible, the uncharacterised reaction with the highest control over glucose uptake was then replaced with kinetic and proteomic data for each isoenzyme; otherwise literature data were used (Table 1). HXT parameters were taken from van Eunen et al. [24] (iteration 1); a saturative generic ATPase rate law (representing the collective of all ATP consuming processes) was taken from Teusink et al. (1998) [54] (iteration 3); the glycerol branch was replaced with the two-reaction model of Cronwright et al. [55] (iteration 4); the glycogen and trehalose branches were replaced with the five-reaction model of Smallbone et al. [56] (iteration 6); acetate and succinate branches were taken from van Eunen et al., 2012 (iteration 7); TPI function and parameters were taken from Mendes et al. [42], though these were measured in-house (iteration 17). We also modified the kinetics of PFK to take account of its thermodynamic reversibility. Subsequent fitting was performed in COPASI.

All models are provided in standard Systems Biology Markup Language format (SBML level 2 version 4; [57]) embedded with principled and computer-readable annotations [17,58–59]. The models are available from the BioModels database [60] and in the JWS Online [61], with the accession numbers given in Table 1.

2.7. Informatics infrastructure

The production of this model was supported by the development of an informatics infrastructure to aid with capture and analysis of experimental data. The toolset includes support for enzyme kinetics [41], quantitative proteomics [49], and quantitative metabolomics data [62–63]. Furthermore, workflows have been generated to automate the generation and parameterisation of kinetic models [64–66].

3. Results and discussion

3.1. A standardised pipeline from network to kinetic model

In order to describe a biological system quantitatively, the characteristics of all its components need to be established. Where

Table 1
Sequence of models generated through the cycles of experiment and modelling. At each iteration, the uncharacterised reaction with the highest control over glucose uptake was then replaced with our experimental data or literature data (see text for details). Each version of the model is available from the BioModels database [60]. For example, the final model may be accessed at http://identifiers.org/biomodels.db/MODEL1303260018. The models are also available from the JWS Online [61], where they can be simulated online at http://jjj.mib.ac.uk/database/smallbone.

Iteration	Reaction	BioModels id	JWS Online id
0		MODEL1303260000	Smallbone0
1	HXT	MODEL1303260001	Smallbone1
2	HXK	MODEL1303260002	Smallbone2
3	ATPase	MODEL1303260003	Smallbone3
4	Glycerol_branch	MODEL1303260004	Smallbone4
5	PFK	MODEL1303260005	Smallbone5
6	Glycogen_branch	MODEL1303260006	Smallbone6
7	Succinate_branch	MODEL1303260007	Smallbone7
8	PDC	MODEL1303260008	Smallbone8
9	TDH	MODEL1303260009	Smallbone9
10	FBA	MODEL1303260010	Smallbone10
11	PGI	MODEL1303260011	Smallbone11
12	ENO	MODEL1303260012	Smallbone12
13	PGK	MODEL1303260013	Smallbone13
14	ADH	MODEL1303260014	Smallbone14
15	GPM	MODEL1303260015	Smallbone15
16	PYK	MODEL1303260016	Smallbone16
17	TPI	MODEL1303260017	Smallbone17
18		MODEL1303260018	Smallbone18

possible, they should represent a system in steady state, where all measurements, even when done at different times, are performed under identical conditions. Therefore, a controlled turbidostat culture of *S. cerevisiae* strain YDL227C was grown under strictly defined conditions as described in Section 2, from which cell extracts were prepared for quantifying protein and metabolite concentrations. This pipeline approach was driven by the aim of using robust and standardised experimental strategies that can easily be transferred to the characterisation of other biochemical networks. It involved three distinct experimental platforms: enzyme kinetics, quantitative metabolomics and quantitative proteomics.

In order to characterise the kinetics of isoenzymes, proteins were expressed and purified from commercially available yeast strain collections. Kinetic constants for these enzymes were determined as described in Materials and Methods, leading to the set of parameters given in Table 2. The same table also compares these with some previously published values, although those measurements were taken from a different strain and at a different pH and under a variety of other conditions.

To establish the absolute quantities of these individual isoenzymes in the cell, the complete protein content was extracted from turbidostatic yeast cultures and quantified using the QconCAT approach, according to methods described previously [48,67] and in Section 2. These concentrations are listed in Table 3.

The complete metabolite contents were extracted from the quenched yeast cells of the same turbidostat culture and quantitatively measured by Ultra Performance Liquid Chromatography (UPLC-MS [68]) and Gas Chromatography-Mass Spectrometry (GC–MS [69]) using standard solutions of known concentration as described in Section 2. A summary of the metabolite concentrations so determined is shown in Table 4.

From a modelling perspective, the enzyme kinetic constants and protein concentrations represent the parameters of the system, and the metabolite concentrations represent the variables. Combining the protein concentration data with those for the enzyme kinetic parameters, together with the measured steady state metabolite levels, allows a mathematical model of yeast glycolysis to be produced for this system, which is provided in standard SBML [57]. A major difference in the reactions of our model when compared to previous efforts in modelling metabolism (such as [13,23]) is that isoenzymes are represented explicitly. In addition, the enzyme kinetic parameters have been measured under standard conditions designed to mimic to a greater or lesser extent the intracellular environment.

3.2. The cycle of knowledge is crucial for systems biology

Scientific advance may be seen as an iterative cycle linking knowledge and ideas to observations and data [2,70–72]. Deductive reasoning uses background knowledge to construct (by unstated means) a hypothesis that is tested experimentally to produce observations. By contrast, inductive (or abductive) reasoning is purely data-driven, generating more general (or specific) hypotheses, rather than starting with them. Because of the high dimensionality of typical data, computer-intensive methods are required to turn the data into knowledge.

This inductive–deductive "cycle of knowledge" is crucial for systems biology: predictive mathematical models are used to interpret, organise and integrate the range of available experimental data. The models produce hypotheses about the underlying complex system dynamics, which are tested experimentally. The models are then more closely validated and updated, continuing the cycle.

Many turns of the cycle of knowledge were required to produce this model. An existing model [23] was found not to predict our measured metabolite levels adequately (Table 4). Thus at each

Table 2
Enzyme kinetics parameters determined in this project and contrasted with those of [23].

Reaction	Isoenzyme	Parameter	Value	SEM (%)	Ref. [23]	Unit
ADH	Adh1p	kcat	176	1.3		s^{-1}
ADH	Adh1p	Kacald	0.462	5.3	1.11	mM
ADH	Adh5p	kcat	0	0		s^{-1}
ENO	Eno1p	kcat	7.6	1.9		s^{-1}
ENO	Eno1p	Kp2g	0.043	10	0.04	mM
ENO	Eno2p	kcat	19.9			s^{-1}
ENO	Eno2p	Kp2g	0.104		0.04	mM
FBA	Fba1p	kcat	4.14	1.5		s^{-1}
FBA	Fba1p	Kf16bp	0.451	5.3	0.3	mM
GPM	Gpm1p	kcat	400			s^{-1}
GPM	Gpm1p	Kp2g	1.41		0.08	mM
HXK	Glk1p	kcat	0.0721			s^{-1}
HXK	Glk1p	Kglc	0.0106		0.08	mM
HXK	Glk1p	Katp	0.865		0.15	mM
HXK	Hxk1p	kcat	10.2			s^{-1}
HXK	Hxk1p	Kglc	0.15		0.08	mM
HXK	Hxk1p	Katp	0.293		0.15	mM
HXK	Hxk2p	kcat	63.1			s^{-1}
HXK	Hxk2p	Kglc	0.2		0.08	mM
HXK	Hxk2p	Katp	0.195		0.15	mM
PDC	Pdc1p	kcat	12.1	3.5		s^{-1}
PDC	Pdc1p	Kpyr	8.5	10	4.33	mM
PDC	Pdc5p	kcat	10.3	2.1		s^{-1}
PDC	Pdc5p	Kpyr	7.08	5.7	4.33	mM
PDC	Pdc6p	kcat	9.21			s^{-1}
PDC	Pdc6p	Kpyr	2.92		4.33	mM
PFK	Pfk1p:Pfk2p	kcat	210	1.8		s^{-1}
PGI	Pgi1p	kcat	487	3.7		s^{-1}
PGI	Pgi1p	Kg6p	1.03	19	1.4	mM
PGI	Pgi1p	Kgf6p	0.307	6.8	0.3	mM
PGK	Pgk1p	kcat	58.6			s^{-1}
PGK	Pgk1p	Kp3g	4.58		0.53	mM
PGK	Pgk1p	Katp	1.99		0.3	mM
PGK	Pgk1p	nHadp	2			
PYK	Cdc19p	kcat	20.1	2.9		s^{-1}
PYK	Cdc19p	Kpep	0.281	12	0.14	mM
PYK	Cdc19p	Kadp	0.243	13	0.53	mM
PYK	Pyk2p	kcat	0	0		s^{-1}
TDH	Tdh1p	kcat	19.1	1.5		s^{-1}
TDH	Tdh1p	Kgap	0.495	7.5	0.21	mM
TDH	Tdh2p	kcat	8.63			s^{-1}
TDH	Tdh2p	Kgap	0.77		0.21	mM
TDH	Tdh3p	kcat	18.2	1.8		s^{-1}
TDH	Tdh3p	Kgap	0.423	9.2	0.21	mM
TDH	Tdh3p	Kbpg	0.909		0.0098	mM
TPI	Tpi1p	kcat	564	1.5		s^{-1}
TPI	Tpi1p	Kdhap	6.45	5		mM
TPI	Tpi1p	Kgap	5.25	12		mM
TPI	Tpi1p	Kigap	35.1	3.1		mM

iteration, the most important reaction to characterise was identified through its control over glucose uptake. This 'most important' reaction was then characterised experimentally, or through literature data.

For some iterations, the addition of experimental data led to a significant decrease in both fit to our data and flux through the pathway. This was because the limiting rate (V) of these reactions – a function of both protein concentration and turnover number – was now too low (i.e. was lower than the estimated in vivo flux). By contrast, using all expression data simultaneously provides substantial constraints in which non-systematic noise cancels out effectively [73]. In vitro and in vivo enzymatic activities may differ for a number of reasons such as missing or unknown effectors, cellular crowding or channelling. However, they may also differ due to inactivation of the enzyme in the purification or assay methods, or by the presence of the purification tag at the N- or C-terminus. By applying alternative purification and assay techniques to isoenzymes with low activity, new data were collected, allowing the model to be refined and improving its consistency with the

Table 3
Protein concentrations measured by QconCat, expressed as protein molecules per cell. For comparison the data determined in [36] is also displayed.

Reaction	Isoenzyme	Uniprot id	Molecules/cell	SEM (%)	Ref. [36]
ADH	Adh1p	P00330	494 000	1.1	0
ADH	Adh5p	P38113	12 800	8.1	1310
ENO	Eno1p	P00924	2 070 000	0.9	76 700
ENO	Eno2p	P00925	5 950 000	0.7	2610
FBA	Fba1p	P14540	4 030 000		1 020 000
GPM	Gpm1p	P00950	2 200 000	0.6	172 000
HXK	Glk1p	P17709	136 000	5.5	21 100
HXK	Hxk1p	P04806	50 500	1.1	40 800
HXK	Hxk2p	P04807	185 000	0.6	114 000
PDC	Pdc1p	P06169	3 220 000	0.8	8970
PDC	Pdc5p	P16467	37 200	19.1	471 000
PDC	Pdc6p	P26263	19 700	1.7	1520
PFK	Pfk1p	P16861	141 000	0.7	89 800
PFK	Pfk2p	P16862	118 000	2.8	90 200
PGI	Pgi1p	P12709	416 000	0.6	91 600
PGK	Pgk1p	P00560	776 000	1.5	314 000
PYK	Cdc19p	P00549	6 170 000	1.3	291 000
PYK	Pyk2p	P52489	18 300	11.4	2130
TDH	Tdh1p	P00360	1 060 000	1.2	120 000
TDH	Tdh2p	P00358	0		121 000
TDH	Tdh3p	P00359	12 700 000	0.9	169 000
TPI	Tpi1p	P00942	886 000	0.8	207 000

observations. Four cycles of experiment and modelling were thus carried out.

For phosphofructokinase a k_{cat} value of 3 s^{-1} was first determined, which was well below what should be expected to match the actual glycolytic flux. However this is because the α and β subunits (products of different genes, *PFK1* and *PFK2*) were purified and assayed in isolation. A subsequent assay combined the α and β subunits, that had been purified separately, in equimolar amounts, but this still produced a low k_{cat} value. Finally the two subunits were co-expressed, co-purified and then assayed, resulting in a k_{cat} value of 200 s^{-1}.

The reaction catalysed by the enolases was also deficient in total capacity relative to the observed physiological flux: the k_{cat} of Eno1p was 0.8 s^{-1} and Eno2p was inactive. Since these enzymes had been previously purified and stored at −70 °C, the new experiment assayed them immediately after purification. In this case Eno1p had a k_{cat} of 8 s^{-1} but Eno2p was still inactive, and the activity of Eno1p alone was not enough to explain the flux. A hypothesis was formulated that the tag in the C-terminus of Eno2p (which is left on in the MORPH strain proteins) could be interfering with its active site. Published evidence for the role of the C-terminus of enolase in enzyme activity and in binding Mg^{2+} were found [74–75]. Thus a new construct was obtained for *ENO2* with a tag in the N-terminus, it was expressed, purified and assayed fresh,

resulting in a k_{cat} of 20 s^{-1}, which is more than twice as active as Eno1p.

The calculated limiting rate of the reaction catalysed by the pyruvate decarboxylases, EC 4.1.1.1, was also below the physiological flux. The k_{cat} values determined were: Pdc1p 3 s^{-1}, Pdc5p 2 s^{-1}, and Pdc6p was inactive. These enzymes had been stored at −70 °C following purification. We re-purified them and performed the assay immediately after the purification. This time, the k_{cat} values for the three enzymes were all 10 s^{-1}, which leads to the conclusion that the storage had caused loss of activity.

The enzyme concentration of fructose bisphosphate aldolase (Fba1p, EC 4.1.2.13) was updated from 5×10^5–4×10^6 molecules per cell, after reinterpretation of the proteomics data. Lysine acetylation occurred on an internal lysine residue of an analyte peptide from Fba1p. This led to division of the analyte signal between the modified and unmodified forms, present in our data, and resulted in an underestimation of the total analyte signal [48]. Thus the inclusion of post-translational modifications as well as isozyme data can lead to more precise estimations of relevant kinetic parameters.

It is important to note that experimental problems such as those described above only came to light upon modelling of the entire system. Since this project included iterative modelling and experimental components, replicate experiments could be carried out when needed. This contrasts with previous modelling approaches that typically relied upon third-party experimental data. It also highlights that the iterative and integrative modelling approach followed here has an important role in quality control for the experimental data, which would not be applicable to dedicated experimental studies that do not contain a modelling element.

In Fig. 2, we present the change in the fit between model predictions and measured metabolite concentrations through each iterative characterisation. While an improvement in fit between the initial model and that fed with our parameter values is seen, there is little change in fit or flux after five iterations. As such, it is vital for model predictivity only to characterise those reactions with the highest control under standard conditions; for reactions with lower control, other sources may be used.

3.3. Effective cytoplasmic volume is an important parameter

Whilst many parameters were measured under standardised conditions, some were nonetheless estimated from the literature. The relative contribution of each unmeasured parameter to the quality of fit to data may be ranked using sensitivity analysis (see e.g., Ref. [76]). We defined a number of fitting criteria: sufficient closeness of predicted to observed metabolite concentrations, no drop in flux into glycolysis, and an energy charge of at least 0.8 [77]. If we were unable to satisfy these criteria by fitting only the

Table 4
Measured and predicted metabolite concentrations. Experimental values determined as described in Methods. Concentrations were calculated relative to an effective cytoplasmic volume of 5 fl (see main text). Predictions were calculated with the final model (iteration 18 in Table 1). Data from [23] is presented for comparison.

| ID	Name	ChEBI id	Molecules/cell ($\times 10^3$)	SEM (%)	Concentration (mM)		
					Observed	Predicted	Ref. [23]
DHAP	Dihydroxyacetone phosphate	16108	3500	18.9	1.162	1.584	1.004
F16bP	Fructose 1,6-bisphosphate	28013	13 800	20.2	4.583	2.780	6.221
F6P	Fructose 6-phosphate	16084	709	14.8	0.235	0.325	0.625
G3P	Glycerol 3-phosphate	15978	825	24.3	0.274	0.226	0.150
G6P	Glucose 6-phosphate	17665	2330	9.7	0.774	1.213	2.675
GAP	Glyceraldehyde 3-phosphate	29052	951	26.9	0.316	0.067	0.045
GLC	Glucose	4167	18 900	4.6	6.277	0.635	0.098
PEP	Phosphoenol-pyruvate	18021	1840	23.5	0.611	0.477	0.063
PYR	Pyruvate	15361	6350	16.6	2.109	3.329	1.815
P2G	Glycerate	17835	1620	17.7	0.538	0.613	1.013
P3G	Phosphates	17794					
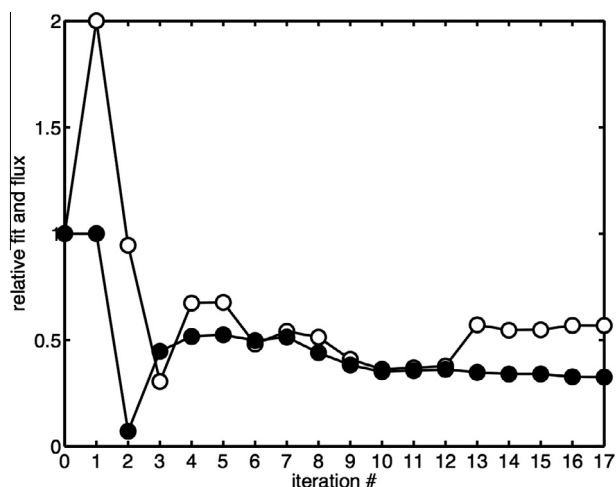
Fig. 2. Improvement in fit between model predictions and data as more reactions are characterised. The filled circles denote the fit as determined by normalised root mean square difference between the measured and experimental metabolite concentrations. Also shown is the glucose uptake rate (open circles). Both fit and flux are rescaled to take value unity for the initial model.

most important parameter, we tried using two parameters, and continued the cycle until complete.

The most important parameter was found to be the cytoplasmic volume, whose role is to convert metabolite and protein copy number to molar concentration (the units in which the Michaelis–Menten parameters are measured). Given estimates of cell volume of 40 fl [78–79], of which around 50% is cytoplasm [80], we initially estimated the cytoplasm to be 20 fl. However, following fitting, we found that a volume of 5 fl provided a much better fit to our data. This could be explained by macromolecular crowding [81], which has the effect that the space available for water and dissolved molecules is much smaller than the total space, and so the effective volume is lower. It is also important to realise that our model represents a bulk average [82–83], so the volume should reflect the average volume of cells. Actively growing cultures are dominated by new cells recently formed by budding that have not undergone division – these cells are smaller than those that have already divided. Further, older cells have large vacuoles that occupy most of the cellular space and their cytoplasmic volume is considerably smaller than the total cellular volume [84].

The other two changes in parameter values required were both related to bulked reactions: the ATPase rate was increased (V from 1.1 mM s^{-1} to 5.3 mM s^{-1}) and the succinate branch was turned off (k from 0.005 s^{-1} to 0).

3.4. Is 3-phosphoglycerate a substrate for enolase 1?

The assays performed for phosphoglycerate mutase (GPM, EC 5.4.2.1) in the direction of 3-phosphoglycerate (P3G) to 2-phosphoglycerate (P2G) revealed an interaction between the substrate P3G and the coupling enzymes system (specifically the enolase). This made it difficult to determine the kinetic properties of GPM for the forward reaction, and hence only the kinetics for the reverse reaction could be determined.

In studying this interaction, it was found that P3G acts as a substrate for Eno1p (k_{cat} = 5.07 s^{-1} ± 2.9% and K_m = 0.617 mM ± 10.2%), producing phophenolpyruvate (PEP). Multiple batches from different sources were tested and ^{31}P NMR spectra of the P3G material showed no evidence for the presence of P2G, hence this observation is not due to P2G contamination. There is some evidence [85] that ENO and GPM form a metabolic channel [86]; since our

model assumes no channelling, this was not incorporated, but is noted here as a discovery of interest.

3.5. Separating kinetic parameters from protein expression levels reveals individuality of isoenzymes

In this work, we constructed a model based on three data sources: metabolomics, proteomics and enzyme kinetics. In contrast to earlier work, we separated the limiting rates into two terms representing (the product of) the enzyme level and its maximal turnover number: V_{max} = [E]·k_{cat}. This is important because the in vivo limiting rate is near impossible to determine. Specifically, it is not possible to deconvolve the activity of different isoenzymes from cell-free extracts measurements. The strategy we followed is to determine separately [E] (by quantitative proteomics) and k_{cat} (using enzyme kinetics). Here the default assumption is that the k_{cat} of enzymes is not changed by the processes of protein purification; since we purify proteins that have tags in their C-terminus, we also assume that the k_{cat} (and other kinetic parameters) are not changed by the presence of the tag. The latter condition is more likely to be violated in some cases, so we checked for this. Indeed, we have found that Eno2p activity was affected by C-terminus tags (see above).

The benefits of this separation are twofold. First, it allows one to investigate the effects of the different isoenzymes within the system. For example glucokinase (Glk1p, EC 2.7.1.2) has lower activity than the two hexokinases (Hxk1p and Hxk2p, EC 2.7.1.1); however it also has a higher affinity for its substrate glucose (see Table 2), allowing it to operate effectively at low substrate concentrations.

Secondly, the kinetic characteristics of enzymes remain fairly constant under different growth conditions. Rather, cells respond to cues through changing enzyme levels, validating the assumptions stated above. As such, knowledge of enzyme levels under different physiological conditions may be combined with a specific set of kinetic data to create a suite of mathematical models of a system.

3.6. Metabolic control analysis

In Table 5, we present the flux-control coefficients (FCC [52]) for the initial and final models. For the initial model, the three main controllers of glucose uptake appeared to be glucose transport,

Table 5
Flux-control coefficients on the glucose input flux, for the initial and final models (iterations 0 and 18 in Table 1).

Reaction	Flux-control coefficients	
	Model 0	Model 18
HXT	0.844	1.467
HXK	0.201	1.995
ATPase	−0.076	−3.303
PFK	0.045	0.356
Glycerol_branch	−0.017	
Glycogen_branch	−0.009	
Succinate_branch	−0.008	
TDH	0.007	0.118
ADH	0.006	0.359
Trehalose_branch	−0.006	
FBA	0.005	0.102
PGI	0.005	0.023
ENO	0.003	0.016
PGK	0.000	0.172
PYK	0.000	0.063
GPM	0.000	−0.044
PDC	0.000	0.001
TPI	0.000	0.008
AK	0.000	0.000

hexokinase and ATPase. These same three also exert the most control in the final model, though through decomposition we find that it is specifically Hxk2p (FCC 1.91), rather than Hxk1p (FCC 0.08) or Glk1p (FCC 0.001) that exerts most of the flux control. Overall, we found in the original model that 95% of the total control is exerted by these three reactions plus PFK. By contrast, in the final model, control is distributed more widely over the whole system, with 95% of control distributed over ten reactions. This represents a substantial change in the picture as our understanding was refined. This is also different from distributions in other organisms; for example, in *Trypanosoma brucei*, control over glycolytic flux was found to be almost uniquely determined by the glucose transporter [87]. These results are also in agreement with the view of Kacser and Burns [88] that the control of metabolism is, in general, and necessarily, widely distributed.

4. Conclusion

For over a decade, modern systems biology has been seen as a crucial strategy and methodology for biological research. It pays particular attention to the way in which molecules interact inside cells in a quantitative way. Within systems biology there are two fundamentally distinct approaches, one uses data from large-scale measurements (transcriptomics, etc.) to reverse engineer models directly from those data; the other approach is based on combining enzyme kinetic data to build mechanistic biochemical network models. The latter approach, sometimes referred to as "bottom-up", is a descendent of the work of Michaelis and Menten, and at one level uses essentially the same methodology to assess the properties of each of the component enzymes in the system. It has thus been common to re-use, in systems biology models, the enzyme kinetic parameters previously estimated in enzymology studies (e.g., Ref. [13]). The present work improves on this practice by recognising that those enzymology studies were carried out under different conditions for each enzyme, while in cells the actual conditions are (at least, nominally) common to all enzymes of the pathway residing in a given compartment. Here, all enzymes – and importantly all isoenzymes – were assayed in standard conditions, such that the resulting model is self-consistent and physiologically relevant. The approach taken here followed the cycle of knowledge approach [2,70–72] as several iterations of experiments and modelling were performed. The final model reveals that the control of the pathway is considerably more distributed than was thought in the past based on previous models, and is in agreement with both experimental evidence and theoretical insights [88]. This study provides a clear example of how bottom-up systems biology, based on the solid foundations laid out 100 years ago by Michaelis and Menten [5], is useful in discovering and understanding new biochemistry.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2013.06.043.

References

[1] Palsson, B.Ø. (2006) Systems Biology: Properties of Reconstructed Networks, Cambridge University Press, Cambridge.

[2] Kell, D.B. (2006) Metabolomics, modelling and machine learning in systems biology: towards an understanding of the languages of cells. The 2005 Theodor Bücher lecture. FEBS J. 273, 873–894.

[3] Kell, D.B. and Knowles, J.D. (2006) The role of modeling in systems biologySystem Modeling in Cellular Biology: From Concepts to Nuts and Bolts, pp. 3–18, MIT Press, Cambridge.

[4] Westerhoff, H.V., Winder, C., Messiha, H., Simeonidis, E., Adamczyk, M., Verma, M., Bruggeman, F.J. and Dunn, W. (2009) Systems biology: the elements and principles of life. FEBS Lett. 583, 3882–3890.

[5] Michaelis, L. and Menten, M.L. (1913) Kinetik der Invertinwirkung. Biochem. Z. 49, 333–369.

[6] Lubitz, T., Schulz, M., Klipp, E. and Liebermeister, W. (2010) Parameter balancing in kinetic models of cell metabolism. J. Phys. Chem. B 114, 16298–16303.

[7] Garcia-Contreras, R., Vos, P., Westerhoff, H.V. and Boogerd, F.C. (2012) Why in vivo may not equal in vitro – new effectors revealed by measurement of enzymatic activities under the same in vivo-like assay conditions. FEBS J. 279, 4145–4159.

[8] Murabito, E., Smallbone, K., Swinton, J., Westerhoff, H.V. and Steuer, R. (2011) A probabilistic approach to identify putative drug targets in biochemical networks. J. R. Soc. Interface 8, 880–895.

[9] Gu, Z., Cavalcanti, A., Chen, F.-C., Bouman, P. and Li, W.-H. (2002) Extent of gene duplication in the genomes of *Drosophila*, nematode and yeast. Mol. Biol. Evol. 19, 256–262.

[10] Wolfe, K.H. and Shields, D.C. (1997) Molecular evidence for an ancient duplication of the entire yeast genome. Nature 387, 708–713.

[11] Kellis, M., Birren, B.W. and Lander, E.S. (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature 428, 617–624.

[12] Langkjaer, R.B., Cliften, P.F., Johnston, M. and Piskur, J. (2003) Yeast genome duplication was followed by asynchronous differentiation of duplicated genes. Nature 421, 848–852.

[13] 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.

[14] Fell, D.A. (1996) Understanding the Control of Metabolism, Portland Press, London.

[15] Iber, D., Clarkson, J., Yudkin, M.D. and Campbell, I.D. (2006) The mechanism of cell differentiation in *Bacillus subtilis*. Nature 441, 371–374.

[16] Ananiadou, S., Kell, D.B. and Tsujii, J.-I. (2006) Text mining and its potential applications in systems biology. Trends Biotechnol. 24, 571–579.

[17] Herrgård, M.J., Swainston, N., Dobson, P., Dunn, W.B., Arga, K.Y., Arvas, M., Blüthgen, N., Borger, S., Costenoble, R., Heinemann, M., Hucka, M., Le Novère, N., Li, P., Liebermeister, W., Mo, M.L., Oliveira, A.P., Petranovic, D., Pettifer, S., Simeonidis, E., Smallbone, K., Spasić, I., Weichart, D., Brent, R., Broomhead, D.S., Westerhoff, H.V., Kirdar, B., Penttilä, M., Klipp, E., Palsson, B.Ø., Sauer, U., Oliver, S.G., Mendes, P., Nielsen, J. and Kell, D.B. (2008) A consensus yeast metabolic network obtained from a community approach to systems biology. Nat. Biotechnol. 26, 1155–1160.

[18] Dobson, P.D., Smallbone, K., Jameson, D., Simeonidis, E., Lanthaler, K., Pir, P., Lu, C., Swainston, N., Dunn, W.B., Fisher, P., Hull, D., Brown, M., Oshota, O., Stanford, N.J., Kell, D.B., King, R.D., Oliver, S.G., Stevens, R.D. and Mendes, P. (2010) Further developments towards a genome-scale metabolic model of yeast. BMC Syst. Biol. 4, 145.

[19] Heavner, B.D., Smallbone, K., Barker, B., Mendes, P. and Walker, L.P. (2012) Yeast 5 – an expanded reconstruction of the *Saccharomyces cerevisiae* metabolic network. BMC Syst. Biol. 4, 55.

[20] Thiele, I., Swainston, N., Fleming, R.M.T., Hoppe, A., Sahoo, S., Aurich, M.K., Haraldsdóttír, H., Mo, M.L., Rolfsson, O., Stobbe, M.D., Thorleifsson, S.G., Agren, R., Bölling, C., Bordel, S., Chavali, A.K., Dobson, P., Dunn, W.B., Endler, L., Goryanin, I., Hala, D., Hucka, M., Hull, D., Jameson, D., Jamshidi, N., Jones, J., Jonsson, J.J., Juty, N., Keating, S., Nookaew, I., Le Novère, N., Malys, N., Mazein, A., Papin, J.A., Patel, Y., Price, N.D., Selkov Sr., E., Sigurdsson, M.I., Simeonidis, E., Sonnenschein, N., Smallbone, K., Sorokin, A., Beek, H.V., Weichart, D., Nielsen, J.B., Westerhoff, H.V., Kell, D.B., Mendes, P. and Palsson, B.Ø. (2013) A community-driven global reconstruction of human metabolism. Nat. Biotechnol. 31, 419–425.

[21] Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., Singhal, M., Xu, L., Mendes, P. and Kummer, U. (2006) COPASI: a complex pathway simulator. Bioinformatics 22, 3067–3074.

[22] Cornish-Bowden, A. (1997) New Beer in an Old Bottle: Eduard Buchner and the Growth of Biochemical Knowledge, Universidad de Valencia, Valencia.

[23] Pritchard, L. and Kell, D.B. (2002) Schemes of flux control in a model of *Saccharomyces cerevisiae* glycolysis. Eur. J. Biochem. 269, 3894–3904.

[24] van Eunen, K., Kiewiet, J.A., Westerhoff, H.V. and Bakker, B.M. (2012) Testing biochemistry revisited: how in vivo metabolism can be understood from in vitro enzyme kinetics. PLoS Comput. Biol. 8, e1002483.

[25] van Hoek, P., de Hulster, E., van Dijken, J.P. and Pronk, J.T. (2000) Fermentative capacity in high-cell-density fed-batch cultures of baker's yeast. Biotechnol. Bioeng. 68, 517–523.

[26] Nielsen, J. and Jewett, M.C. (2008) Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. FEMS Yeast Res. 8, 122–131.

[27] Nevoigt, E. (2008) Progress in metabolic engineering of *Saccharomyces cerevisiae*. Microbiol. Mol. Biol. Rev. 72, 379–412.

[28] Ward, O.P. and Young, C.S. (1990) Reductive biotransformations of organic compounds by cells or enzymes of yeast. Enzyme Microb. Technol. 12, 482–493.

[29] Österlund, T., Nookaew, I. and Nielsen, J. (2012) Fifteen years of large scale metabolic modeling of yeast: developments and impacts. Biotechnol. Adv. 30, 979–988.

[30] Zaldivar, J., Nielsen, J. and Olsson, L. (2001) Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration. Appl. Microbiol. Biotechnol. 56, 17–34.

[31] Bull, A.T. (2010) The renaissance of continuous culture in the post-genomics age. J. Ind. Microbiol. Biotechnol. 37, 993–1021.

[32] Dikicioglu, D., Karabekmez, E., Rash, B., Pir, P., Kirdar, B. and Oliver, S.G. (2011) How yeast re-programmes its transcriptional profile in response to different nutrient impulses. BMC Syst. Biol. 5, 148.

[33] Weusthuis, R.A., Visser, W., Pronk, J.T., Scheffers, W.A. and van Dijken, J.P. (1994) Effects of oxygen limitation on sugar metabolism in yeasts: a continuous-culture study of the Kluyver effect. Microbiology 140, 703–715.

[34] Gelperin, D.M., White, M.A., Wilkinson, M.L., Kon, Y., Kung, L.A., Wise, K.J., Lopez-Hoyo, N., Jiang, L., Piccirillo, S., Yu, H., Gerstein, M., Dumont, M.E., Phizicky, E.M., Snyder, M. and Grayhack, E.J. (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. Genes Dev. 19, 2816–2826.

[35] Sopko, R., Huang, D., Preston, N., Chua, G., Papp, B., Kafadar, K., Snyder, M., Oliver, S.G., Cyert, M., Hughes, T.R., Boone, C. and Andrews, B. (2006) Mapping pathways and phenotypes by systematic gene overexpression. Mol. Cell 21, 319–330.

[36] Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K. and Weissman, J.S. (2003) Global analysis of protein expression in yeast. Nature 425, 737–741.

[37] Malys, N., Carroll, K., Miyan, J., Tollervey, D. and McCarthy, J.E.G. (2004) The 'scavenger' m7G pppX pyrophosphatase activity of Dcs1 modulates nutrient-induced responses in yeast. Nucleic Acids Res. 32, 3590–3600.

[38] Malys, N., Wishart, J.A., Oliver, S.G. and McCarthy, J.E.G. (2011) Protein production in *S. cerevisiae* for systems biology studies. Methods Enzymol. 500, 197–212.

[39] Messiha, H.L., Malys, N. and Carroll, K.M. (2011) Towards a full quantitative description of yeast metabolism: a systematic approach for estimating the kinetic parameters of isoenzymes under in vivo like conditions. Methods Enzymol. 500, 215–231.

[40] Takagahara, I., Yamauti, J., Fujii, K., Yamashita, J. and Horio, T. (1983) Theoretical and experimental analyses of coupled enzyme reactions. J. Biochem. 93, 1145–1157.

[41] Swainston, N., Golebiewski, M., Messiha, H.L., Malys, N., Kania, R., Kengne, S., Krebs, O., Mir, S., Sauer-Danzwith, H., Smallbone, K., Kell, D.B., Mendes, P., Müller, W., Paton, N.W. and Rojas, I. (2010) Enzyme kinetics informatics: from instrument to browser. FEBS J. 277, 3769–3779.

[42] Mendes, P., Messiha, H., Malys, N. and Hoops, S. (2009) Enzyme kinetics and computational modeling for systems biology. Methods Enzymol. 467, 583–599.

[43] Harris, C.M., Todd, R.W., Bungard, S.J., Lovitt, R.W., Morris, J.G. and Kell, D.B. (1987) The dielectric permittivity of microbial suspensions at radio frequencies: a novel method for the estimation of microbial biomass. Enzyme Microb. Technol. 9, 181–186.

[44] Markx, G.H., Davey, C.L. and Kell, D.B. (1991) The permittistat: a novel type of turbidostat. J. Gen. Microbiol. 137, 735–743.

[45] Winder, C.L. and Lanthaler, K. (2011) The use of continuous culture in systems biology investigations. Methods Enzymol. 500, 261–275.

[46] Pir, P., Gutteridge, A., Wu, J., Rash, B., Kell, D.B., Zhang, N. and Oliver, S.G. (2012) The genetic control of growth rate: a systems biology study in yeast. BMC Syst. Biol. 6, 4.

[47] Davey, H.M., Davey, C.L., Woodward, A.M., Edmonds, A.N., Lee, A.W. and Kell, D.B. (1996) Oscillatory, stochastic and chaotic growth rate fluctuations in permittistatically-controlled yeast cultures. BioSystems 39, 43–61.

[48] Carroll, K.M., Simpson, D.M., Eyers, C.E., Knight, C.G., Brownridge, P., Dunn, W.B., Winder, C.L., Lanthaler, K., Pir, P., Malys, N., Kell, D.B., Oliver, S.G., Gaskell, S.J. and Beynon, R.J. (2011) Absolute quantification of the glycolytic pathway in yeast: deployment of a complete QconCAT approach. Mol. Cell. Proteomics 10. M111.007633.

[49] Swainston, N., Jameson, D. and Carroll, K. (2011) A QconCAT informatics pipeline for the analysis, visualization and sharing of absolute quantitative proteomics data. Proteomics 11, 329–333.

[50] Dunn, W.B. and Winder, C.L. (2011) Sample preparation related to the intracellular metabolome of yeast methods for quenching, extraction, and metabolite quantitation. Methods Enzymol. 500, 277–297.

[51] Martins, A.M., Sha, W., Evans, C., Martino-Catt, S., Mendes, P. and Shulaev, V. (2007) Comparison of sampling techniques for parallel analysis of transcript and metabolite levels in *Saccharomyces cerevisiae*. Yeast 24, 181–188.

[52] Burns, J.A., Cornish-Bowden, A., Groen, A.K., Heinrich, R., Kacser, H., Porteous, J.W., Rapoport, S.M., Rapoport, T.A., Stucki, J.W., Tager, J.M., Wanders, R.J.A. and Westerhoff, H.V. (1985) Control analysis of metabolic systems. Trends Biochem. Sci. 10, 16.

[53] Kacser, H. and Burns, J.A. (1973) The control of flux. Symp. Soc. Exp. Biol. 27, 65–104.

[54] Teusink, B., Walsh, M.C., van Dam, K. and Westerhoff, H.V. (1998) The danger of metabolic pathways with turbo design. Trends Biochem. Sci. 23, 162–169.

[55] Cronwright, G.R., Rohwer, J.M. and Prior, B.A. (2002) Metabolic control analysis of glycerol synthesis in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. 68, 4448–4456.

[56] Smallbone, K., Malys, N., Messiha, H.L., Wishart, J.A. and Simeonidis, E. (2012) Building a kinetic model of trehalose biosynthesis in *Saccharomyces cerevisiae*. Methods Enzymol. 500, 355–370.

[57] Hucka, M., Finney, A., Sauro, H.M., Bolouri, H., Doyle, J.C., Kitano, H., Arkin, A.P., Bornstein, B.J., Bray, D., Cornish-Bowden, A., Cuellar, A.A., Dronov, S., Gilles, E.D., Ginkel, M., Gor, V., Goryanin, I.I., Hedley, W.J., Hodgman, T.C., Hofmeyr, J.H., Hunter, P.J., Juty, N.S., Kasberger, J.L., Kremling, A., Kummer, U., Le Novère, N., Loew, L.M., Lucio, D., Mendes, P., Minch, E., Mjolsness, E.D., Nakayama, Y., Nelson, M.R., Nielsen, P.F., Sakurada, T., Schaff, J.C., Shapiro, B.E., Shimizu, T.S., Spence, H.D., Stelling, J., Takahashi, K., Tomita, M., Wagner, J., Wang, J. and Forum, S.B.M.L. (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 19, 524–531.

[58] Le Novère, N., Finney, A., Hucka, M., Bhalla, U.S., Campagne, F., Collado-Vides, J., Crampin, E.J., Halstead, M., Klipp, E., Mendes, P., Nielsen, P., Sauro, H., Shapiro, B., Snoep, J.L., Spence, H.D. and Wanner, B.L. (2005) Minimum information requested in the annotation of biochemical models (MIRIAM). Nat. Biotechnol. 23, 1509–1515.

[59] Courtot, M., Juty, N., Knüpfer, C., Waltemath, D., Zhukova, A., Dräger, A., Dumontier, M., Finney, A., Golebiewski, M., Hastings, J., Hoops, S., Keating, S., Kell, D.B., Kerrien, S., Lawson, J., Lister, A., Lu, J., Machne, R., Mendes, P., Pocock, M., Rodriguez, N., Villéger, A., Wilkinson, D.J., Wimalaratne, S., Laibe, C., Hucka, M. and Le Novère, N. (2011) Controlled vocabularies and semantics in systems biology. Mol. Syst. Biol. 7, 543.

[60] Li, C., Donizelli, M., Rodriguez, N., Dharuri, H., Endler, L., Chelliah, V., Li, L., He, E., Henry, A., Stefan, M.I., Snoep, J.L., Hucka, M., Le Novère, N. and Laibe, C. (2010) BioModels Database: an enhanced, curated and annotated resource for published quantitative kinetic models. BMC Syst. Biol. 4, 92.

[61] Olivier, B.G. and Snoep, J.L. (2004) Web-based kinetic modelling using JWS online. Bioinformatics 20, 2143–2144.

[62] Brown, M., Dunn, W.B., Dobson, P., Patel, Y., Winder, C.L., Francis-McIntyre, S., Begley, P., Carroll, K., Broadhurst, D., Tseng, A., Swainston, N., Spasic, I., Goodacre, R. and Kell, D.B. (2009) Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. Analyst 134, 1322–1332.

[63] Spasić, I., Dunn, W.B., Velarde, G., Tseng, A., Jenkins, H., Hardy, N., Oliver, S.G. and Kell, D.B. (2006) MeMo: a hybrid SQL/XML approach to metabolomic data management for functional genomics. BMC Bioinformatics 7, 281.

[64] Li, P., Dada, J.O., Jameson, D., Spasic, I., Swainston, N., Carroll, K., Dunn, W., Khan, F., Messiha, H.L., Simeonidis, E., Weichart, D., Winder, C., Broomhead, D., Goble, C.A., Gaskell, S.J., Kell, D.B., Westerhoff, H.V., Mendes, P. and Paton, N.W. (2010) Systematic integration of experimental data and models in systems biology. BMC Bioinformatics 11, 582.

[65] Swainston, N. and Mendes, P. (2009) LibAnnotationSBML: a library for exploiting SBML annotations. Bioinformatics 25, 2292–2293.

[66] Swainston, N., Jameson, D., Li, P., Spasic, I., Mendes, P. and Paton, N.W. (2010) Integrative information management for systems biology. Lect. Notes Comput. Sci. 6254, 164–178.

[67] Beynon, R.J., Doherty, M.K., Pratt, J.M. and Gaskell, S.J. (2005) Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. Nat. Methods 2, 587–589.

[68] Zelena, E., Dunn, W., Broadhurst, D., Francis-McIntyre, S., Carroll, K., Begley, P., O'Hagan, S., Knowles, J.D., Halsall, A., HUSERMET Consortium, Wilson, I. and Kell, D.B. (2009) Development of a robust and repeatable UPLC-MS method for the long-term metabolomic study of human serum. Anal. Chem. 81, 1357–1364.

[69] Begley, P., Francis-McIntyre, S., Dunn, W.B., Broadhurst, D.I., Halsall, A., Tseng, A., Knowles, J., HUSERMET consortium, Goodacre, R. and Kell, D.B. (2009) Development and performance of a gas chromatography-time-of-flight mass spectrometry analysis for large-scale non-targeted metabolomic studies of human serum. Anal. Chem. 81, 7038–7046.

[70] Kell, D.B. (2004) Metabolomics and systems biology: making sense of the soup. Curr. Opin. Microbiol. 7, 296–307.

[71] Kell, D.B. and Oliver, S.G. (2004) Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. BioEssays 26, 99–105.

[72] Elliott, K.C. (2012) Epistemic and methodological iteration in scientific research. Stud. Hist. Philos. Sci. 43, 376–382.

[73] Lee, D., Smallbone, K., Dunn, W.B., Murabito, E., Winder, C.L., Kell, D.B., Mendes, P. and Swainston, N. (2012) Improving metabolic flux predictions using absolute gene expression data. BMC Syst. Biol. 6, 73.

[74] Lebioda, L. and Stec, B. (1991) Mechanism of enolase: the crystal structure of enolase-Mg^{2+}-2-phosphoglycerate/phosphoenolpyrvate complex resolution. Biochemistry 30, 2817–2822.

[75] Pancholi, V. (2001) Multifunctional α-enolase: its role in diseases. Cell. Mol. Life Sci. 58, 902–920.

[76] White, T.A. and Kell, D.B. (2004) Comparative genomic assessment of novel broad-spectrum targets for antibacterial drugs. Comp. Funct. Genomics 5, 304–327.

[77] Berg, J., Tymoczko, J. and Stryer, L. (2007) Biochemistry, W.H. Freeman and Co., New York.

[78] Adams, J. (1977) The interrelationship of cell growth and division in haploid and diploid cells of Saccharomyces cerevisiae. Exp. Cell Res. 106, 267–275.

[79] Weiss, R.L., Kukora, J.R. and Adams, J. (1975) The relationship between enzyme activity, cell geometry, and fitness in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U.S.A. 72, 794–798.

[80] Biswas, S.K., Yamaguchi, M., Naoe, N., Takashima, T. and Takeo, K. (2003) Quantitative three-dimensional structural analysis of Exophiala dermatitidis yeast cells by freeze-substitution and serial ultrathin sectioning. J. Electron Microsc. 52, 133–143.

[81] Minton, A.P. (2001) The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. J. Biol. Chem. 276, 10577–10580.

[82] Kell, D.B., Ryder, H.M., Kaprelyants, A.S. and Westerhoff, H.V. (1991) Quantifying heterogeneity: flow cytometry of bacterial cultures. Antonie Van Leeuwenhoek 60, 145–158.

[83] Davey, H.M. and Kell, D.B. (1996) Flow cytometry and cell sorting of heterogeneous microbial populations: the importance of single-cell analysis. Microbiol. Rev. 60, 641–696.

[84] Jorgensen, P., Edgington, N.P., Schneider, B.L., Rupes, I., Tyers, M. and Futcher, B. (2007) The size of the nucleus increases as yeast cells grow. Mol. Biol. Cell 18, 3523–3532.

[85] Hauf, J., Zimmermann, F.K. and Müller, S. (2000) Simultaneous genomic overexpression of seven glycolytic enzymes in the yeast Saccharomyces cerevisiae. Enzyme Microb. Technol. 26, 688–698.

[86] Mendes, P., Kell, D.B. and Welch, G.R. (1995) Metabolic channeling in organized enzyme systems: experiments and modelsEnzymology in Vivo, pp. 1–19, JAI Press, London.

[87] Albert, M.A., Haanstra, J.R., Hannaert, V., Van Roy, J., Opperdoes, F.R., Bakker, B.M. and Michels, P.A. (2005) Experimental and in silico analyses of glycolytic flux control in bloodstream form Trypanosoma brucei. J. Biol. Chem. 280, 28306–28315.

[88] Kacser, H. and Burns, J.A. (1981) The molecular basis of dominance. Genetics 97, 636–639.