

TOWARDS A FULL GENOME-SCALE MODEL OF YEAST  
METABOLISM

A THESIS  
SUBMITTED TO THE UNIVERSITY OF MANCHESTER  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY (PHD)  
IN THE FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

NATALIE JANE STANFORD

SCHOOL OF CHEMICAL ENGINEERING AND ANALYTICAL SCIENCE  
2011

# CONTENTS

<b>Abstract</b>	<b>8</b>
<b>Declaration</b>	<b>9</b>
<b>Copyright</b>	<b>10</b>
<b>Acknowledgements</b>	<b>11</b>
<b>1 Introduction</b>	<b>12</b>
1.1 Systems Biology . . . . .	12
1.2 A perspective on modelling in cellular biology . . . . .	13
1.3 Flux balance analysis . . . . .	18
1.4 Metabolite profiling and quantification . . . . .	21
<b>2 Letting the flux define the kinetics: using a single steady state to predict network behaviour under diverse stress conditions</b>	<b>23</b>
2.1 Introduction . . . . .	23
2.2 Method . . . . .	29
2.2.1 Parameterising using First Approximation Methods (FAM) . . . . .	29
2.2.2 Stage 1 data collection . . . . .	32
2.2.3 Stage 2 data collection . . . . .	33
2.3 Results and discussion . . . . .	35
2.3.1 Stage 1 results . . . . .	35
2.3.2 Stage 2 results . . . . .	41
2.4 Conclusions . . . . .	42
<b>3 Avoiding perpetual motion models: a method for ensuring complex metabolic networks are not in violation of the first law of thermodynamics</b>	<b>44</b>
3.1 Introduction . . . . .	44
3.2 Testing on real reaction networks . . . . .	49
3.3 Conclusions . . . . .	52
<b>4 Generating thermodynamically-consistent and experimentally founded kinetic models at the genome-scale</b>	<b>54</b>
4.1 Introduction . . . . .	54
4.2 Method . . . . .	57
4.2.1 Metabolic control analysis . . . . .	65
4.3 Results and discussion . . . . .	66
4.3.1 Data input [Data A to D] . . . . .	66

4.3.2	Processing steps of the pipeline . . . . .	67
4.4	Conclusions . . . . .	71
<b>5</b>	<b>New insights from old data: emergent properties of the kinetically enhanced consensus yeast network built using non-homogenous data sources</b>	<b>73</b>
5.1	Introduction . . . . .	73
5.2	Method . . . . .	76
5.2.1	Combining glycolysis and trehalose . . . . .	76
5.2.2	Altering the network . . . . .	80
5.3	Results and discussion . . . . .	84
5.4	Conclusions . . . . .	95
<b>6</b>	<b>Dynamic flux experiments offer further insight into cellular behaviour</b>	<b>96</b>
6.1	Introduction . . . . .	96
6.1.1	Determining sampling times . . . . .	100
6.2	Method . . . . .	100
6.2.1	Fermentation and sampling . . . . .	100
6.2.2	Mass Spectrometry (MS) analysis . . . . .	101
6.2.3	Computational Analysis . . . . .	103
6.3	Results and Discussion . . . . .	104
6.3.1	Experimental results . . . . .	104
6.3.2	Computational results and discussion . . . . .	110
6.4	Conclusions . . . . .	114
6.5	Acknowledgements . . . . .	115
<b>7</b>	<b>Summary and future directions</b>	<b>116</b>
7.1	Summary . . . . .	116
7.2	Future directions - the thesis . . . . .	118
7.3	Future directions - the field of study . . . . .	119
7.4	Final thoughts . . . . .	120
<b>A</b>	<b>Generating thermodynamically-consistent and experimentally founded kinetic models at the genome-scale - supplementary information</b>	<b>138</b>
<b>B</b>	<b>New insights from old data: emergent properties from kinetic transformation of the consensus yeast network built using non-homogenous data sources - supplementary information</b>	<b>153</b>
<b>C</b>	<b>Dynamic flux experiments offer further insight into cellular behaviour - supplementary information</b>	<b>168</b>

## LIST OF FIGURES

1.1	The cycle of knowledge . . . . .	15
1.2	FBA example network . . . . .	18
2.1	Flux profile behaviour . . . . .	29
2.2	Control behaviour in first approximation models . . . . .	38
2.3	Control behaviour in Pritchard and Kell model . . . . .	39
2.4	Residual errors for stage 1 and stage 2 results . . . . .	40
3.1	Reaction networks . . . . .	46
3.2	Schematic of reactions outlined in Table 3.2 . . . . .	50
3.3	linearly dependent reactions in <i>Methylobacterium extorquens</i> . . . . .	51
4.1	Model building work-flow . . . . .	58
5.1	Granularity differences in top-down and bottom-up modelling . . . . .	75
5.2	Flux distribution through central carbon metabolism . . . . .	90
5.3	effect of extracellular glucose perturbation of 50mM and 100mM on D-glucose . . . . .	94
5.4	effect of extracellular glucose perturbation of 50mM and 100mM on D-fructose 1,6-bisphosphate . . . . .	94
5.5	effect of extracellular glucose perturbation of 50mM and 100mM on pyruvate . . . . .	94
5.6	effect of extracellular glucose perturbation of 50mM and 100mM on ethanol . . . . .	95
6.1	Metabolite exchange Vs Atom exchange . . . . .	99
6.2	Reduced reaction network for modelling . . . . .	104
6.3	Time points where <sup>13</sup> C concentration was highest in each metabolic pool . . . . .	105
6.4	<sup>13</sup> C saturation in extracellular glucose . . . . .	106
6.5	Pathway dynamic flux data . . . . .	107
6.7	Mapping <sup>13</sup> C saturation trend . . . . .	109
6.6	<sup>13</sup> C saturation through select amino acid pools . . . . .	109
6.8	Glycolysis branch point . . . . .	110
6.9	Network flux results . . . . .	112
C.1	Percentage <sup>13</sup> C in extracellular glucose . . . . .	168
C.2	Percentage <sup>13</sup> C in intracellular glucose . . . . .	169
C.3	Percentage <sup>13</sup> C in glucose 6-phosphate . . . . .	169
C.4	Percentage <sup>13</sup> C in fructose 6-phosphate . . . . .	170
C.5	Percentage <sup>13</sup> C in 6-phosphogluconate . . . . .	170
C.6	Percentage <sup>13</sup> C in ribose 5-phosphate . . . . .	171

C.7	Percentage $^{13}\text{C}$ in erythrose 4-phosphate . . . . .	171
C.8	Percentage $^{13}\text{C}$ in glycerol . . . . .	172
C.9	Percentage $^{13}\text{C}$ in citrate . . . . .	172

## LIST OF TABLES

2.1	Steady state fluxes . . . . .	32
2.2	Steady state concentrations . . . . .	33
2.3	Stage 1 residual errors . . . . .	35
2.4	Stage 2 residual errors . . . . .	41
3.1	Equilibrium constants associated with reaction Figure 3.1 network (b) . . .	46
3.2	Linearly dependent reactions in the erythrocyte model by Holzhütter [1].	50
3.3	Linearly dependent reactions in the <i>Methylobacterium extorquens</i> . . . . .	52
4.1	Steady state fluxes calculated from yeast specific models in the biomodels database that use glucose as the primary carbon source . . . . .	59
4.2	Equilibrium constants taken from yeast specific models in the biomodels database that use glucose as the primary carbon source. . . . .	60
4.3	Yeast specific intracellular metabolite concentrations taken from yeast specific models in the biomodels database that use glucose as the primary carbon source . . . . .	62
4.4	Extracellular metabolite concentrations taken from extracellular metabolome measurements . . . . .	63
4.5	Kegg pathways represented in the model . . . . .	69
4.6	Reactions with high total flux control . . . . .	71
5.1	Intracellular metabolite concentrations . . . . .	78
5.2	Extracellular metabolite concentrations . . . . .	79
5.3	Steady state fluxes . . . . .	80
5.4	Network adaptations to match the yeast strain knockouts . . . . .	81
5.5	Extracellular metabolites known to have low concentrations . . . . .	82
5.6	<i>in vitro</i> measured equilibrium constants . . . . .	83
5.7	Modified extracellular concentrations . . . . .	85
5.8	Flux results for known fluxes . . . . .	87
5.9	Kegg pathways represented in the model . . . . .	88
5.10	Metabolite names . . . . .	91
5.11	Reactions with high total flux control . . . . .	92
6.1	Peak area associated with highest percentage <sup>13</sup> C levels . . . . .	106
6.2	Quantified flux results . . . . .	113
6.3	Comparison of measured concentrations and steady state concentrations .	114
A.1	Full network reaction fluxes . . . . .	138
A.2	Full network metabolite concentrations . . . . .	145

B.1	Full network reaction fluxes . . . . .	153
B.2	Full network metabolite concentrations . . . . .	160

## ABSTRACT

Gaining a quantitative understanding of metabolic behaviour has long been a major scientific goal. Beginning with crude mass balance experiments and progressing through enzyme kinetics, single-pathway models and collaborative efforts such as a community-based yeast reconstruction and onwards to the digital human. The primary goal of this research was to generate a large-scale kinetic metabolic model of yeast metabolism. As a community our ability to produce large-scale dynamic metabolic models has typically been limited by the time and cost involved in obtaining exact measurements of all relevant kinetic parameters. Attempts have been made to bring about a greater understanding by using computational approaches such as flux balance analysis, and also laboratory approaches such as metabolic profiling. Unfortunately these approaches alone do not go far enough to allow for a rich understanding of the metabolic behaviour.

Methods were developed that allowed known data such as fluxes, equilibrium constants and metabolite concentrations to be used in first-approximation strategies. These made possible the construction of a thermodynamically consistent model that was reflective of the organism and growth conditions under which the known data were measured. Efforts were made to improve the strategy by developing already known dynamic flux measurement techniques so they were more reflective of the type of data required for constructing the metabolic model.

The model constructed, using data from a specific yeast strain in a continuous culture environment, and included 284 reactions. The model showed a reasonable reproduction of system behaviour after perturbations of extracellular glucose above and below the operating conditions, after identification and substitution of just two exact rate laws of reactions that showed high control over the system.

The methods developed require little knowledge beyond the stoichiometric matrix in the first instance, and as such, are applicable to any organism that has a reasonably comprehensive network reconstruction available.

# DECLARATION

**The University of Manchester**  
*Traditional format PhD Candidate Declaration*

**Candidate Name:** Natalie Jane Stanford

**Faculty:** Engineering and Physical Sciences

**Thesis Title:** Towards a full genome-scale model of yeast metabolism

**Declaration to be completed by the candidate:**

I declare that no portion of this work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Signed:

Date: November 27, 2011

## COPYRIGHT

The author of this thesis (including any appendices and/or schedules to this thesis) owns any copyright in it (the "Copyright")<sup>1</sup> and s/he has given The University of Manchester the right to use such Copyright for any administrative, promotional, educational and/or teaching purposes.

Copies of this thesis, either in full or in extracts, may be made only in accordance with the regulations of the John Rylands University Library of Manchester. Details of these regulations may be obtained from the Librarian. This page must form part of any such copies made.

The ownership of any patents, designs, trade marks and any and all other intellectual property rights except for the Copyright (the "Intellectual Property Rights") and any reproductions of copyright works, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.

Further information on the conditions under which disclosure, publication and exploitation of this thesis, the Copyright and any Intellectual Property Rights and/or Reproductions described in it may take place is available from the Head of School of Chemical Engineering and Analytical Science (or the Vice-President) and the Dean of the Faculty of Engineering and Physical Sciences, for Faculty of Engineering and Physical Sciences candidates.

---

<sup>1</sup>This excludes material already printed in academic journals, for which the copyright belongs to said journal and publisher. Pages for which the author does not own the copyright are numbered differently from the rest of the thesis.

## ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisors, Pedro Mendes and Kieran Smallbone. Pedro for showing me that the history of science is just as exciting as new research (in some cases, more so...) and Kieran for constantly pushing me beyond the borders of science that I am comfortable with. Between them they have given me encouragement and freedom to explore the areas of science that interest me the most, whilst also managing to keep me on track with the tight schedule for the PhD. A testament to their skills as scientists, and as people.

Second big thanks go to Cate Winder and Rick Dunn. Cate for being brave enough to train a dry lab scientist in the wet lab, and for being a constant stream of biological information over the years. Rick, for dedicating the time to process the Mass Spectrometry samples.

To the MCISB, and the Theoretical Biophysics group in Berlin, I thank you for being great scientists and great people. I can't imagine a more entertaining bunch of folk to have spent three years with.

For my friends at the DTC, back home and in far away lands: many thanks for keeping me sane over the years.

Huge thanks go to my family and best friend Laura, who have supported me through everything, with a passion only family (and surrogate family) could muster. My Mum for constantly being there with encouragement, Sunday lunches and hugs to keep me going through the hard days. Big Sis, for all the patience required to keep me entertained over the years (stair tobogganing anyone?). My Dad for finding a new hobby every 6 months, persuading me that you can learn to do anything if you have enough enthusiasm for it. My Nana for being a strong and open-minded lady, with an immense lust for life (if I can still line dance when I'm your age I'll be happy). Laura (my 'wife'), for being an amazing person and making me feel like I can achieve anything I set my mind to. I love you all, and you all make me a better person each and every day.

Lastly, I dedicate my thesis to my late Granddad. He was a great man, and a great role model, and I will miss him always.

## INTRODUCTION

### **1.1 Systems Biology**

Systems biology can be viewed as a philosophical approach to scientific study or an area of research in itself. Disambiguating the two is difficult. The philosophy of systems biology refers more specifically to viewing the system as a whole, in contrast to the methodologies employed in, for example, molecular biology, which focus on profiling and cataloging the single genes, enzymes and metabolites [2] and their physico-chemical behaviour. As an analogy, the systems approach is to understand the structure and behaviour of the house, without having to have an exact description of each brick.

As an area of research, systems biology focusses on using computational approaches to improve understanding on how the components of the whole system fit together and how they interact within the new, complex, environment. Some authors argue that it is a modern approach to physiology [3], but it looks at more than just how functions are carried out within an organism. It combines physiology with large components of systems engineering, self-organisation theory [4] and emergence theory [5]. What appears more clear is that systems biology is primarily not seen as a paradigm shift in conventional scientific methodologies, it is viewed more as a borrowed paradigm from neighbouring fields [4, 3].

In order to investigate a system computationally the ‘parts’ must be available to use within the ‘whole’. Systems biology is not the anti-thesis to reductionism, it is more

a necessary complement that ensures the most insight is gained out of complex data sets. This is where systems biology becomes an integration tool. There needs to be an investment in organising current knowledge so it is more accessible, whilst generating techniques to fill in the knowledge gaps. This process may require laboratory data of a different variety than has been collected so far. This integrating technique is what helps to drive understanding.

## 1.2 A perspective on modelling in cellular biology

Models of systems are commonplace in biology. Traditional models include schematics of DNA chains, structures of biochemical molecules and even the cellular structure itself [6]. They allow us more accessible insights into the behaviour of a system based upon our current knowledge. These types of models are, however, not the most useful for hypothesis generation. In the past decade, modelling in biology has progressed significantly, driven by a desire to understand the complex molecular interactions that occur to allow the living system to survive. Basic maps of molecular interactions have progressed to kinetic descriptions of interactions. These kinetic interaction data have then increased in scope to cover larger areas of cellular behaviour.

The biomodels database [7] is a repository for such models. It contains nearly 700 models coded in SBML, which is an XML markup language designed for community sharing of mathematical models. These models cover, amongst others: the cell cycle, signalling pathways, metabolic pathways, telomere behaviour and circadian rhythms. These pathways are modelled for a range of organisms. They frequently focus on very specific phenomena, and model how a finite area of the cell network of a given organism reacts to a defined set of perturbations, or growth conditions. There is little scope for a ‘general case’ of behaviour within the highly specific design, although there have been attempts to correlate data into more general representations of networks [8].

The early stages of building a model are difficult. The data needed are primarily limited to areas of interest from the past. As an example, there are 12 yeast specific metabolic models in the biomodels database, but they only cover trehalose [9], glycolysis [10, 11, 12, 13, 14, 15, 16, 17, 18, 19], glycerol synthesis [20], and a general

representation of aerobic respiration [19]. This demonstrates the degree of data overlap present, with glycolysis being highly coveted and reproduced many times whilst other metabolic pathways are neglected. This situation is common among models and experimental data. In situations where there is a lot of data, it is often disparate and not suitable for cross-comparison, which also makes it difficult to use when building models. This is due to the data being collected under different specific conditions with no way of combining them in a biologically feasible way.

An obvious, and perhaps ideal, way to overcome these issues would be to generate a full new set of laboratory data. A typical bottom-up modelling approach would require the measurement of metabolite concentrations (profiled [21, 22] and quantified ) using metabolomic methods. Protein levels would be measured using quantitative proteomic techniques [23]. Flux distributions [24, 25, 26, 27, 28, 29], and in some instances fluxes would be quantified [30, 31, 32] using fluxomic techniques. The rate laws of the reactions would be determined by performing analysis on the purified enzyme, these data would then be compiled together into a kinetic model. This is a time consuming process, and requires a huge range of specific expertise and resources. For the time-scale of many projects, and indeed many careers, generating models that reach into wider areas of the metabolism would be very hard. Therefore, a balance must be struck between what is ideal and what is functional. A forward thinking approach that aims for continual refinement of models and experiments is very accessible. It allows for collation and improvement of knowledge at each stage, meaning both experimental and modelling approaches drive each other in order to identify the most important areas of cellular behaviour to investigate. This has been referred to as the 'cycle of knowledge' [33], it is illustrated in Figure 1.1.

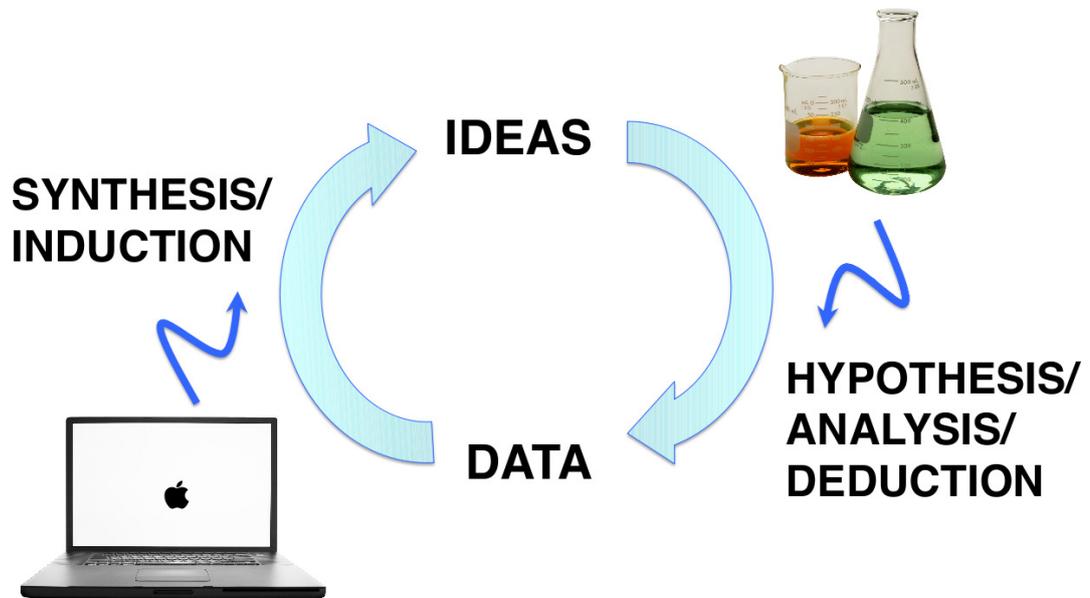


Figure 1.1: The cycle of knowledge, adapted from [33]

This thesis puts together techniques and approaches to understand metabolic systems from a wide range of fields including systems theory, thermodynamics, metabolomics, and micro-biology, and integrates them together, along with developing new techniques, to broaden the areas of metabolism which can be studied. The approach is developed using the budding yeast *Saccharomyces cerevisiae* but is designed such that it would be applicable to any organism with a genome-scale metabolic network reconstruction, and adaptable ways of culturing the cells at steady state. The methods outlined in each chapter use data and techniques from a broad range of fields, therefore, each chapter contains a review of the relevant background information. Three techniques, routinely applied to study biological systems, span the chapters. As such it is appropriate to outline these approaches here.

## Metabolic control analysis (MCA)

The viability of cells is heavily related to the ability to perform necessary metabolic processes under a wide variety of external stresses. A way of achieving this is to tightly control the metabolic processes in maintaining homeostatic conditions. Understanding how the cell achieves this is difficult, not only is the metabolic network large and complex, but the actual metabolic behaviour is related more to parametric control of the reactions

rather than the metabolic network structure itself.

Understanding of the metabolic behaviours requires quantitative approaches that allow assessment of the system in response to parametric changes within the network, and how this might help to regulate the homeostatic environment [34] (known as the ‘response coefficient’). There have been two key approaches that aim to do this, Biochemical Systems Theory (BST) [35, 36, 37] and MCA [38, 39, 40, 41]. It has been argued that the underlying mathematics to these two approaches is much the same [42]. Where the two techniques differ is in the parameters used to determine their sensitivities. BST uses the ‘rate constants’ and MCA uses the enzyme concentrations/activity [34]. The less defined relationship between rate constants and enzyme activity results in the response sensitivities being more complicated to interpret in the BST approach. The implementation of MCA in systems has also been shown to have more simplicity in its application for studying response sensitivities [43]. It is possible this is caused by the parameters and variables used within the analysis being more intuitive for analysing metabolic changes. This is further aided by the use of the response coefficient that allows the technique to be used for looking at very specific parameters, but relying on the same underlying principles.

Coupled to the advantages mentioned, MCA is also much more widely used within the literature than BST [34]. As such MCA will be the method of choice used within this thesis for analysing network behaviour.

The principles of MCA were first introduced by Kacser and Burns in 1973 [38] and independently by Heinrich and Rapoport in 1974 [39]. It is a technique for looking at how system variables such as flux and concentrations are affected by changes in system parameters, primarily the enzyme concentration. The original treatments of MCA focus on how a small change in the enzyme concentration of a reaction affect the flux ( $v_i$ ) through each reaction in the network. All non-equilibrium reactions within the network, and the metabolites associated with these reactions can show a change in response to the enzyme concentration change, the magnitude of these changes comprise the analysis.

The flux and concentration responses of the entire network, in response to a small change in one reaction are called control coefficients and are split into two categories,

flux control coefficients  $C_{v_i}^J$  and concentration control coefficients  $C_{v_i}^S$ . They are defined as follows:

$$C_{v_i}^J = \frac{d \ln J}{d \ln v_i} \quad (1.2.1)$$

$$C_{v_i}^S = \frac{d \ln S}{d \ln v_i} \quad (1.2.2)$$

The development of MCA and the control coefficients was vital for showing that previous ideas of ‘bottle-neck’ reactions and system ‘regulators’ were a fallacy. The control of a metabolic network is distributed across all reactions. This stems from the summation theorems:

$$\sum_i C_{v_i}^J = 1 \quad (1.2.3)$$

$$\sum_i C_{v_i}^S = 0 \quad (1.2.4)$$

This does not preclude that a single reaction may demonstrate a high level of control over the system, but where this occurs there will be lower control distributed between other reaction fluxes. It is also noted that the reaction can not control the system in all conditions. Reactions that are held far from equilibrium within the network, and reactions with comparatively low enzyme concentrations tend to show a higher level of control over system variables than their counterparts that are closer to equilibrium or with higher enzyme concentrations [38]. Increasing the concentration of an enzyme that demonstrates a high level of control reduces the control that particular enzyme has over the rest of the metabolic system whilst increasing that of others.

MCA has involved numerous advances which have extended its applicability to much more complex networks. These include extensions for branching pathways and moiety conservations [44] which Reder generalised into a comprehensive matrix method [45]. A more fundamental advance has been the extension to look at supply and demand of the metabolic network [46]. This extension borrows theory from traditional economics and applies it to the behaviour of the system using an MCA-based methodology. Given the

similarities between economic supply chains and the metabolic network, this leaves scope for further applications that could be used to extend analysis techniques of the network.

Throughout, this thesis focusses on traditional MCA approaches for network analysis, using COPASI [47] to obtain the values. The implementation of MCA in COPASI follows the methods of Reder [45].

### 1.3 Flux balance analysis

FBA is a linear programming technique used to calculate the flux distribution through a metabolic network. The calculations require the knowledge of the stoichiometric matrix, a set of constraints, and a biologically feasible objective for the system. Kauffman and colleagues [48] cite the most common objectives as ATP production [49], biomass production [50, 51, 52] or the rate of production of particular product [53]. Coupled to this any known fluxes through the system can be incorporated within the minimum and maximum flux bounds for each reaction. Assuming homeostatic behaviour, FBA takes the stoichiometric matrix and calculates a feasible flux distribution through the network whilst adhering to the constraints imposed.

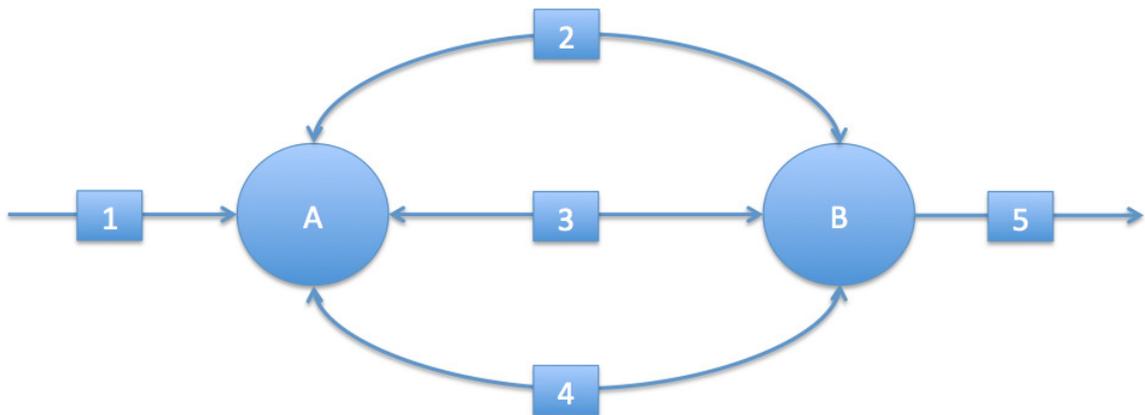


Figure 1.2: Example network adapted from [54]

Taking Figure 1.2 as an example, to maximise the flux through reaction 5, the problem must be formulated as follows:

maximise

$$Z = f^T v \quad (1.3.1)$$

subject to

$$Sv = 0 \quad (1.3.2)$$

where

$$f = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad S^T = \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ -1 & 1 \\ -1 & 1 \\ 0 & 1 \end{pmatrix} \quad v^{min} = \begin{pmatrix} 0 \\ -\infty \\ -\infty \\ -\infty \\ 0 \end{pmatrix} \quad v^{max} = \begin{pmatrix} 1 \\ \infty \\ \infty \\ \infty \\ \infty \end{pmatrix} \quad (1.3.3)$$

As this small network highlights, there is much degeneracy in traditional FBA solutions. To maximise  $Z$ , the optimal solution,  $Z^*$ , must equal 1, but this can be achieved with any of the following:  $v = (1, 1, 0, 0, 1)$ ,  $v = (1, 1, 10000, -10000, 1)$ ,  $v = (1, 0, 1, 0, 1)$ ,  $v = (1, 0, 2, -1, 1)$  or indeed an infinite number of other combinations.

In 2009 a more conservative, and arguably more biologically feasible, way of addressing this degeneracy was proposed by Smallbone and Simeonidis [54]. Here the idea was to minimize the total flux through the network, required to maximise reaction 5 (Figure 1.2), and calculate a central solution. In this case leading to  $v = (1, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 1)$ . This central solution is complex to find in larger systems and can be computationally expensive. As such, the authors developed an algorithm that uses a geometric reduction of solution space, in iterative cycles, to identify the central solution. This algorithm is known as ‘Geometric FBA’ (gFBA), and is available in the Cobra toolbox [55] for Matlab, along with traditional FBA solutions. By minimising the flux and distributing it centrally across the pathways, the method also allows for the removal of futile cycles. However if the constraints of the network are not carefully identified, these futile cycles can be forcibly introduced.

The central solution found using gFBA can be viewed as advantageous because it allows the ability to identify all potential pathways that can be used to optimise the growth of the cell under a given set of growth conditions. What the method does not allow is a way of identifying which flux carrying reactions are most vital for the cell to use.

A way of investigating the importance of different flux carrying reactions, beyond a simplistic gene knockout assessment, is to use Flux Variability Analysis (FVA) [56]. FVA finds the minimum and maximum fluxes possible through each reaction, whilst retaining a pre-defined percentage of the systems objective function. The most stringent test is to find the minimum and maximum fluxes of each reaction that still maintain 100% of the objective function. Reactions that show smaller differences between the minimum and maximum possible fluxes are deemed to be most vital for the optimisation of the objective function. Examples where FVA has been used successfully include identifying the key reactions that are needed for production of actinorhodin in *Streptomyces coelicolor* [57], and storage synthesis in developing oilseed rape embryos [58].

As the FBA technique becomes more widely used, there is an increasing pressure to identify key areas of improvement so the results are more biochemically relevant. A huge area of improvement has been the increasing accuracy of network reconstructions, something that is vital for improving the accuracy of the results. Some dynamic information can be gained from FBA techniques with the application of dynamic FBA (dFBA) [51, 59, 60]. This method uses iterative cycles of FBA, followed by updating the initial conditions (and in some cases, objective functions) so the flux information over a set of changing conditions can be traced.

More recently there has been a drive to improve the thermodynamic correctness of the models produced using FBA. One of the first implementations was to use a way of producing thermodynamically consistent concentration sets, along with the fluxes based around Gibbs free energies [61]. The next was to look at reducing futile cycles from standard FBA runs [62].

The focus of this thesis is to generate a large-scale kinetic model of yeast, therefore only gFBA will be used so a steady state flux can be identified. The thermodynamics of the approach, dynamic behaviour of the network and identification of high control reactions will all be analysed on the kinetic model, leaving little need for other techniques to be used.

## 1.4 Metabolite profiling and quantification

Metabolomics is the non-biased quantification of all metabolites present in biological systems [63]. There are two distinct approaches to data acquisition: a non-targeted approach, known as profiling and a targeted approach which is used to quantify the metabolites in a sample [64]. Quantification techniques are only possible when an authentic standard of the metabolite is available. For both approaches there are a range of different instrumentation that can be used such as Nuclear Magnetic Resonance (NMR), Infa Red (IR) and Raman spectroscopy, and Gas Chromatography or Liquid Chromatography (GC-, LC-) followed by Mass Spectrometry (MS). Here the focus will be on GC-MS techniques. GC-MS has been regarded as the 'gold standard' for quantification of metabolites, although LC-MS is increasing in popularity, whilst IR and Raman spectroscopy would not be suitable for this approach.

To profile or quantify the intracellular metabolites the cells and media solution have to be aliquoted into 60% methanol at  $-48^{\circ}$ , in order to suspend the metabolism [21] before extraction of the metabolites is carried out. This is known as quenching. The quenched cultures are separated from the supernatant prior to metabolite extraction from the biomass. If the cultures have been grown in complex media they are also washed. The supernatant that is removed can be sampled in order to ensure that there has been minimal leakage from the cells during the process. For profiling or quantifying the exo-metabolome, the cultures are not quenched, but sterile filtered immediately after collection and the filtrate analysed.

For intracellular profiling or quantification the collected cells must be lysed so that the metabolites can be extracted from the biomass for analysis. A number of methods are used for extracting the intracellular metabolites. Common extraction techniques include hot water (HW) [65, 66, 67, 68], boiling ethanol (BE) [69, 70], chloroform-methanol (CM) [71, 72, 73] and freezing-thawing in methanol (FTM) [22, 72]. All of the techniques rely on stopping enzyme activity within the cell, before disrupting the cell wall and separating the intracellular metabolites from the rest of the cell matter, *via* vortexing. HW and BE both use protein denaturation to prevent further metabolic activity, followed by vortexing and pooling of the supernatant in order to collect the intracellular

metabolites. CM and FTM rely on retaining the cells at a low enough temperature to prevent further proliferation of the metabolism, the cell wall is then broken over iterative cycles of vortexing (and in the case of FTM, repeat freeze/thaw cycles). Once the first sample of supernatant is taken, the cells are resuspended and the process repeated in order to ensure an extraction that is as exhaustive as possible.

The process preferred in the community is FTM. This is not the preferred method cited by Canelas and colleagues [64], due to metabolite leakage, however, the routine exo-metabolome testing ensures that any leakage can be identified. In practice this leakage is minimal. To increase the metabolites detected by this approach, the sample are derivatised to make the more more volatile, followed by chromatographic separation (either in gas phase or liquid phase) before the mass spectrometry analysis [74]. This reduces the complexity of the sample.

The metabolomic data collected for this thesis uses FTM, with GC-MS to profile and quantify the yeast metabolome. The techniques are primarily used so metabolite pools can be quantified accurately for modelling, and the active pathways in the metabolism can be identified using a mixture of metabolic profiling, with the addition of heavy carbon label. All of the experiments are performed on cells grown in a turbidostat at steady state.

LETTING THE FLUX DEFINE THE KINETICS: USING A  
SINGLE STEADY STATE TO PREDICT NETWORK  
BEHAVIOUR UNDER DIVERSE STRESS CONDITIONS

## 2.1 Introduction

The iterative interplay between mathematical modelling and laboratory experiments to generate biological understanding is not a new concept [75], but it is one that is gaining high esteem in the modern approaches to understanding biological systems [76]. A widely used approach is kinetic modelling, where the mathematical model is constructed using experimentally determined kinetic rate laws corresponding to each unique reaction in the system. These rate laws are usually non-linear in nature. In order to determine kinetic rate laws, *in vitro* data of turnover rates are collected using enzyme assays, performed on purified enzyme preparations. The numeric values of the determined parameters (e.g  $K_m$ ,  $V$ ,  $K_i$ ,  $K_a$ ) depend on the experimental conditions used, such as pH, temperature and presence of modifiers. Determination of these mechanistic rate laws and their parameter values is a time consuming process and is dependent on the availability of purified enzymes and reaction substrates. These measurements also depend on appropriate detection methods for the reactants and products, such as spectrophotometry or fluorescence, and, in situations where this is not possible, coupled enzyme assays have to be used. Beyond the effects of the reactants and products, a

kinetic model of the pathway also requires determination of the kinetic effect of each known modifier of the system (inhibitors, allosteric modulators, etc); modifiers that are unknown cannot be included, of course, and will be a source of error. The mechanistic rate law is determined from the data of such *in vitro* experiments by nonlinear regression ('fitting') to their ODEs or their integrated versions. Taking *in vitro* measurements is also problematic given the *in vitro* behaviour will not accurately represent *in vivo* behaviour. This is because the environmental conditions are different, particularly in terms of protein concentration. *In vitro* measurements are usually carried out in dilute enzyme concentrations, and *in vivo* in very high concentrations that lead to tight packing and possible enzyme-enzyme interactions [77, 78].

A network reconstruction of known molecular interactions in the budding yeast, *Saccharomyces cerevisiae*, shows that for metabolic processes alone there are more than 1000 reactions [79, 80]. The exact kinetic determination of all these reactions would be very costly and time consuming, and in some cases not even feasible. The need for consistent sets of data across all reactions precludes the use of a lot of published data because the conditions under which the data were collected vary drastically, particularly with regards to the pH of the assays, and this leads to thermodynamic inconsistencies. The results are models that do not attain a chemical equilibrium state when isolated from their surroundings [81] a phenomenon known as perpetual motion, which is at odds with the laws of thermodynamics, and impossible.

Previous research has aimed to simplify the process of generating a kinetic rate law by designing generic 'building blocks' that can be adapted and used to generate a first approximation of a model of the global dynamics of a system. Rate laws that can be used in this context include saturation rate laws, such as 'convenience kinetics', which are based on the random order binding Michaelis Menten equation [82, 83] and linear, log-linear and power law rate laws [84, 85]. Variables that can be estimated include metabolite concentrations and equilibrium constants. The metabolite concentrations can be calculated using quantitative structure-property relations, using the hypothesis that certain functional groups contained within a metabolite define whether the cell maintains the metabolite at high or low concentrations [86]. The concentrations can

also be measured using targeted metabolomic techniques [33], which will be more accurate, but will be limited to the metabolites that are detectable and quantifiable. The equilibrium constants can be calculated based on Gibbs free energies, calculated from structural inference [87]. In this instance an algorithm is used to calculate the energy associated with the bonds in the reactant metabolites and product metabolites. The algorithm identifies functional groupings that provide extra stability, such as conjugation, then estimates the energy change associated with the breaking of the reactant bonds and groupings, and the formation of the product bonds and groupings. Whilst useful, these building blocks do not provide all the information necessary for construction of a kinetic model.

Some initial validation for the use of such building blocks has been conducted, an example being an investigation by Hadlich and colleagues [88]. In this paper a reference model is used that shows accurate predictions of a set of stimulus response data. They then use an algorithm that picks one of any single kinetic reaction terms within the model and replaces it with the given approximate term. The new equation is then fitted to the stimulus response data using a linear regression to global least squares. The algorithm continues until no more equations can be substituted without losing predictability of the model. This study showed that good fits were achieved for complete translation into linlog and convenience kinetics rate laws. This demonstrates that a universal rate law could feasibly be used for large-scale modelling. One of the stumbling blocks of this method was the reliance on a large amount of already available data, this still leaves the question of how to implement these rate laws when there is little biological information available.

Methodologies have been developed to help produce a first approximation system. The first is the constraint-based linlog approach to modelling [89]. This methodology uses information that can be obtained from the stoichiometric matrix to generate a first approximation model. The second was an approach and associated software that used enzyme concentrations and convenience kinetics for model generation [90]. This latter approach requires large amounts of time-course data of fluxes and concentrations in order to fit the model parameters and produce a prediction for a single stepwise change in

growth conditions. This amount of data is, arguably, difficult to obtain from experiments. Whilst both experiments demonstrate a first approximation, performance differences between the approaches were not examined.

Overall, the generation of a kinetic model of a large metabolic network is constrained by the complex and time consuming nature of *in vitro* kinetic assays, or limited by a lack of information about which first approximation strategy would be best to use.

In order to address these issues a new method has been developed, based on extending existing first-approximation strategies in order to generate models based on generic rate laws. The models are fitted to a single steady state data set. The models are then extrapolated to a range of different steady states and their ability to predict the steady state fluxes and concentrations present in a range of growth conditions is assessed. Following this, flux control data is used to identify which reactions within the network are the most important, which indicate reactions that should be measured in the lab, and then substituted into the model with the goal of improving predictability. The predictability of the model is then re-assessed. The aim is to see whether a single steady state is enough to approximate a wide-range of steady state system behaviour and also which first-approximation methodology is most suited to such an approach.

### **The rate laws**

#### **Linlog kinetics**

In the linlog rate law (Equation 2.1.1), the flux is represented as a linear combination of logarithmic terms, producing linear behaviour in logarithmic space, as can be seen in figure 2.1 (c). Its use in many investigations show it is mostly suitable for approximating hyperbolic metabolite tendencies. This behaviour is similar to that in an enzyme-catalysed reaction during the transition between the mass action phase and the saturation phase. In theory this means that the linlog models should only demonstrate system behaviour at points close to the reference state. The reference state used in this investigation is the steady state to which the model was fitted. This rate law would therefore be expected to show a poor ability to extrapolate well beyond the initial steady state.

$$V = V_o \left( 1 + \sum_i \epsilon_{s_i} \log \left( \frac{S_i}{S_{o_i}} \right) + \sum_j \epsilon_{p_j} \log \left( \frac{P_j}{P_{o_j}} \right) \right) \quad (2.1.1)$$

S	= Substrate concentration	P	= Product concentration
$n$	= Stoichiometry	$V$	= Flux
$V_o$	= Rate constant	$S_o$	= Reference concentration
$P_o$	= Reference product	$\epsilon_s$	= Substrate elasticity
$\epsilon_p$	= Product elasticity		

### Convenience rate law

The convenience rate law (Equation 2.1.2) is a saturation rate law that is inspired by a random order binding mechanism. This allows each equation to produce a mass action phase, a hyperbolic phase and a saturation phase. At reactant concentrations of 0 the rate is 0 and at metabolite concentrations approaching infinity the rate law reaches a saturable maximum (see figure 2.1 (a)). The similarity between this and a real enzyme catalysed reaction suggests it should provide the best fit. It also has the advantage that the parameters are easy to interpret in a biological sense, given that all parameters can be attributed to specific functions within the enzyme catalysed reaction.

$$V = V_m \frac{\prod_i S_i^{n_i} - K_{eq}^{-1} \prod_j P_j^{n_j}}{\prod_i \left( 1 + \frac{S_i}{K_{s_i}} \right) + \prod_j \left( 1 + \frac{P_j}{K_{p_j}} \right) - 1} \quad (2.1.2)$$

S	= Substrate concentration	P	= Product concentration
$K_{eq}$	= Equilibrium constant	$V_m$	= Rate constant
$K_s$	= Substrate Michaelis constant	$K_p$	= Product Michaelis constant

### Hybrid rate law

The hybrid rate law (Equation 2.1.3) is a simplified version of the convenience rate law. It retains the mass action behaviour in the numerator, but the denominator is reduced to just additive terms to represent the binding behaviour of substrates and products to the enzyme. This rate law should display similar behaviour to the convenience rate law.

$$V = V_m \frac{\prod_i S_i^{n_i} - K_{eq}^{-1} \prod_j P_j^{n_j}}{1 + \sum_i \frac{S_i}{K_{s_i}} + \sum_j \frac{P_j}{K_{p_j}}} \quad (2.1.3)$$

S	= Substrate concentration	P	= Product concentration
$K_{eq}$	= Equilibrium constant	$V_m$	= Rate constant
$K_s$	= Substrate Michaelis constant	$K_p$	= Product Michaelis constant

### Mass action kinetics

The mass action rate law (Equation 2.1.4) is used to predict the rate at which a set of substrates will undergo a spontaneous chemical reaction into a set of products. The rate law contains only a scaling component, and lacks saturation behaviour. This results in a rate that approaches  $\infty$  as the concentration of substrate approaches  $\infty$ , much like the linlog rate law. It does confer an advantage over linlog given that when the substrate concentration is 0 the rate is 0. This makes it physically more feasible to use.

$$V = k \left( \prod_i S_i^{n_i} - K_{eq}^{-1} \prod_j P_j^{n_j} \right) \quad (2.1.4)$$

S	= Substrate concentration	P	= Product concentration
$K_{eq}$	= Equilibrium constant	$k$	= Rate constant

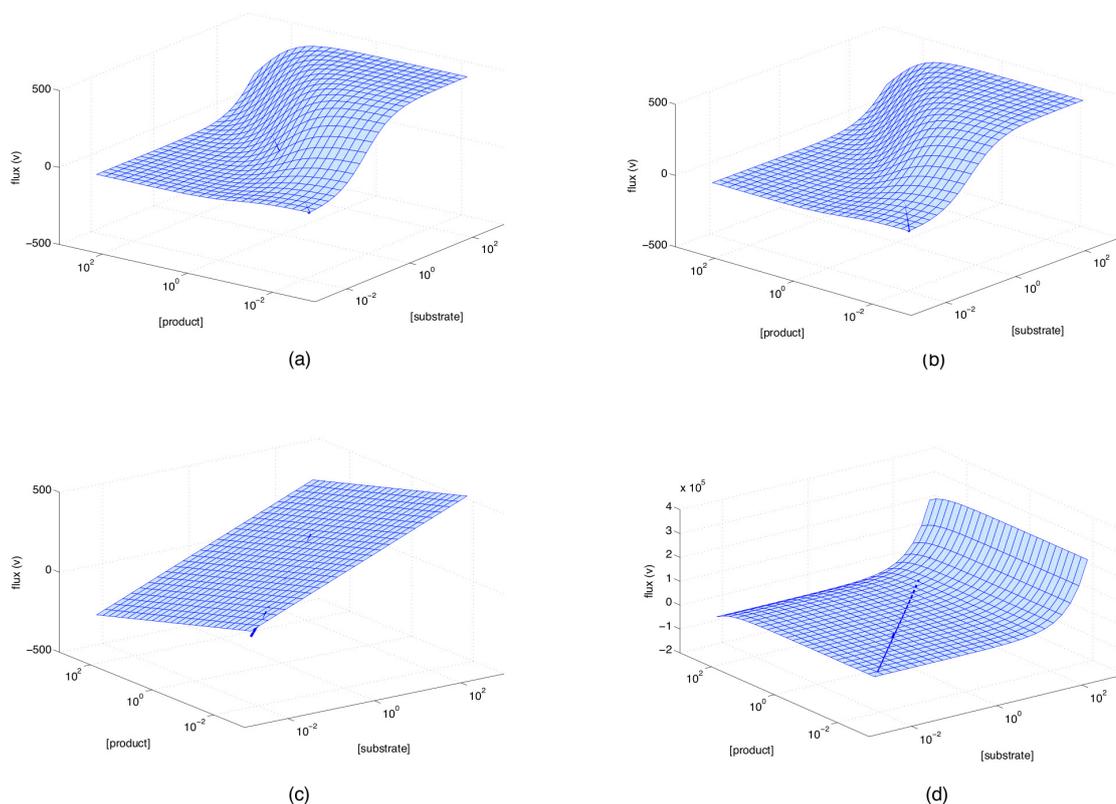


Figure 2.1: Flux behaviour of each rate law against concentrations of substrate and product. The rate laws are: a) convenience, b) hybrid, c) linlog, d) mass action. Mass action and linlog demonstrate a lack of flux saturation within the rate law as the substrate and product concentrations approach the extreme. The behaviour is for the enolase reaction in the Pritchard and Kell model [14]

## 2.2 Method

### 2.2.1 Parameterising using First Approximation Methods (FAM)

Knowledge of only one steady state set of concentrations and fluxes is assumed (see tables 2.1 and 2.2). This is to represent the data that can be obtained from metabolic and dynamic flux experiments. Each rate law has a small number of parameters that can be systematically populated with only the network structure and the flux and concentration data. The steady state data is obtained from a model of yeast glycolysis [14], note, extracellular glucose was set to 100 mM (standard operating conditions for the model). COPASI was used for the simulation [47]. The rate law for alcohol dehydrogenase was reversed so all reaction fluxes presented positive flux values.

The FAM were also tested for their prediction accuracy by populating the equations with data that is calculable from the Pritchard and Kell model. This experiment was

undertaken so a comparison of behaviour with actual data could be obtained. These are known as Data Heavy Models (DHM).

### Linlog

The FAM methodology is identical to the methodology outlined in Smallbone and Simeonidis (2007) [89]. In brief, the elasticities ( $\epsilon$ ) were taken to be the negative of the metabolites' reaction stoichiometry. The reference concentrations ( $S_0$  and  $P_0$ ) were taken to be the value of the steady state concentration of the corresponding metabolite. The rate  $V_0$  was taken to be the steady state flux concentration of the corresponding reaction. This should leave the system in a steady state matching the data.

The DHM used elasticity values calculated from the Pritchard and Kell model, using symbolic differentiation. Where elasticities were equal to  $-\infty$  or  $\infty$  (a result of unidirectional rate laws) their values were replaced with the FAM values. The reference concentrations and flux data remained the same.

### Convenience, mass action and hybrid

#### Equilibrium constant

Reaction directionality is defined by the thermodynamic component of the rate law. Taking the convenience rate law (equation 2.1.2) the thermodynamic component is the numerator, which is, in essence the mass action equation (2.1.4) without the rate scalar  $k$  (equation 2.2.1). It is only the concentrations of species and the  $K_{eq}$  that define reaction directionality,  $k$ ,  $V_o$  and  $V_m$  provide a rate scalar, and the denominator provides a scaling based on the availability of enzyme active sites for substrates to bind and react.

$$\prod_i S_i^n - \left( K_{eq}^{-1} \prod_j P_j^n \right) \quad (2.2.1)$$

In order for a reaction to proceed in the forward direction the numerator must be positive. Therefore, the following must be satisfied:

$$\prod_i S_i^n > K_{eq}^{-1} \prod_j P_j^n \quad (2.2.2)$$

and for a reverse reaction the following has to be satisfied:

$$\prod_i S_i^n < K_{eq}^{-1} \prod_j P_j^n \quad (2.2.3)$$

based on these conditions the following is true:

A forward flux reaction satisfies the following:

$$\frac{\prod_j P_j^n}{\prod_i S_i^n} < K_{eq} \quad (2.2.4)$$

a reverse flux reaction satisfies the following:

$$\frac{\prod_j P_j^n}{\prod_i S_i^n} > K_{eq} \quad (2.2.5)$$

The left hand side of equations 2.2.4 and 2.2.5 is referred to as the mass action ratio. In order to obtain the equilibrium constants the mass action ratio of each reaction was calculated using the steady state concentrations and increased by 10%. The increase ensured that the reactions were processing flux in a positive direction and not in equilibrium (ie, flux of 0) or reversed, and therefore could be scaled (explained later in this section). Transport reactions were assigned a  $K_{eq}$  value of 1, which is a value that would be realistically expected for a transport reaction that contained neutral species. Thermodynamic interdependencies of  $K_{eq}$  values were not of primary importance in this network, given that the network contained no cycles. For larger networks where there is linear dependence this should be taken into consideration when calculating the values.

### Other constants

The constants ( $K_s$  and  $K_p$ ) for the hybrid and convenience rate laws were taken to be the steady state concentrations of the corresponding metabolite, in line with the widely held view that these constants tend to be of a similar order of magnitude to their corresponding steady state metabolite concentration.

The rate constants  $k$ ,  $V_o$  and  $V_m$  were taken to be the value of the corresponding steady state flux. The initial rates of reaction were then calculated and the rate constants were

linearly scaled so the new initial rates matched the steady state flux values, which leaves each model in steady state under the initial conditions of the model. These parameters represent the effect that changing enzyme concentrations would have upon the kinetic rate law, and are a consequence of gene expression. Changing these values allowed freedom to adapt the model in a way that is consistent with cellular behaviour.

For the DHM, the procedure was the same, with the exception that the  $K_{eq}$  were calculated from the Pritchard and Kell model. The  $K_s$  and  $K_p$  values from the Pritchard and Kell model were not used because these values are dependent on the mechanistic formulation of the rate law.

Table 2.1: Steady state fluxes

Reaction	Steady state flux (mM min <sup>-1</sup> )
Glucose transport	89.3
Hexokinase	89.3
Glucose 6-phosphate isomerase	78.5
6-Phosphofructo 2-kinase	78.5
Fructosebisphosphate aldolase	78.5
Triosephosphate isomerase	59.5
Glyceraldehyde 3-phos. dehydrogenase	138
3-Phosphoglycerate kinase	138
Phosphoglyceromutase	138
Enolase	138
Pyruvate kinase	138
Pyruvate decarboxylase	138
Alcohol dehydrogenase	130
ATPase	99.7
Adenylate kinase	$1.42 \times 10^{-14}$
Glycerol 3-phosphate dehydrogenase	19.1
Glycogen branch	6.00
Trehalose branch	2.40
Succinate branch	3.81

## 2.2.2 Stage 1 data collection

### Changing extracellular glucose

The primary input of the glycolysis network is extracellular glucose. As a fixed input, alterations in its concentration cause a short term shift in metabolic behaviour. This results in the network attaining a new steady state. Each rate law has different dynamics, so this investigation gives us insight into how the rate laws deal with the characteristic steady state shifts that are associated with an increased system input. The parameter

Table 2.2: Steady state concentrations

Metabolite	Steady state concentration (mM)
D-Glucose	$9.76 \times 10^{-2}$
ATP	2.51
D-Glucose 6-phosphate	2.68
ADP	1.28
D-Fructose 6-phosphate	$6.24 \times 10^{-1}$
D-Fructose 1,6-bisphosphate	6.22
AMP	$2.93 \times 10^{-1}$
Glycerone Phosphate	1.00
D-Glyceraldehyde 3-phosphate	$4.52 \times 10^{-1}$
NAD	1.50
2,3-Disphospho D-glycerate phosphate	$7.37 \times 10^{-4}$
NADH	$8.67 \times 10^{-2}$
3-Phospho D-glycerate	$8.86 \times 10^{-1}$
D-Glycerate 2-phosphate	$1.28 \times 10^{-1}$
Phosphoenol pyruvate	$6.32 \times 10^{-2}$
Pyruvate	1.82
Acetaldehyde	$1.78 \times 10^{-1}$

scan function was used in COPASI to calculate 100 values distributed uniformly in log-concentration of extracellular glucose between 2 and 50000 mM. These extreme concentrations were used, not because of their biological feasibility, but to allow an assessment of the stability of the models over a wide range of metabolite concentrations. Concentrations of 2 mM to 1000 mM extracellular glucose would be classed as the biologically feasible range.

### 2.2.3 Stage 2 data collection

The stage 1 models were analysed using MCA. Here MCA is used to determine which enzyme/reaction has the highest control over all other reactions within the system. To do this the control coefficients ( $C_{v_i}^J$ ) associated with a small change in  $v_i$  are calculated for each reaction. The reaction with the highest control over the system is the one which causes the largest  $C_{v_i}^J$  change across all reactions within the system. The system will have both positive and negative  $C_{v_i}^J$ , it is not the directionality of the  $C_{v_i}^J$  that is important in this case, but rather the magnitude. To prevent the positive and negative control cancelling out, the overall control for each reaction is summed according to equation 2.2.6. Reactions that are at equilibrium within the system are excluded because their  $C_{v_i}^J$  is always + or -  $\infty$ .

$$C_i^* = \sqrt{\sum_J (C_{v_i}^J)^2} \quad (2.2.6)$$

The reaction with the highest summed  $C_i^*$  in each model was substituted with the corresponding equation from the Pritchard and Kell model (originally measured by Teusink and colleagues in [15]). These formed the stage 2 models.

The parameter scan was then repeated on the stage 2 models, in the same way as described in Stage 1, and MCA of the new system was taken.

## 2.3 Results and discussion

### 2.3.1 Stage 1 results

Table 2.3: Stage 1 residual errors from dynamic glucose experiments

	Convenience	Linlog	Hybrid	Mass action
FAM residual error.				
Flux	15122	-	15017	41614
Flux and concentration	128037	-	126958	907912
DHM residual error.				
Flux	7693	756	9036	26261
Flux and concentration	76548	7968	86610	994077

#### Changing extracellular glucose

A total of 100 values distributed uniformly in log-concentration of extracellular glucose between 2 to 50000 mM were simulated to a steady state. Figure 2.4 (a) shows that the convenience, hybrid and mass action models manage to attain a steady state even under extremely high extracellular glucose concentrations. The convenience and hybrid models, in particular, begin to show a curtailing of the computed error difference from the original model. This is likely related to the enzyme saturation that is mimicked within their respective rate laws. These characteristics could prove important when modelling a larger metabolic network. Further downstream metabolic pathways may differ in flux and metabolite concentrations by several orders of magnitude. An ability to process these diverse data in a biologically meaningful way is vital for the overall stability and predictability of the network.

The mass action rate law appears to demonstrate a discontinuity during the simulation at approximately 5000 mM extracellular glucose. At this point, after having a largely increasing error at concentrations higher than the initial state, the models' predictability improves dramatically for a short range of concentrations before starting to rise once again. This is likely the result of the systemic properties of the network above 5000 mM extracellular glucose.

The FAM linlog model is unable to obtain a steady state at extracellular glucose concentrations below the reference state, or above around 5000 mM extracellular glucose,

resulting in the sum of squares not being calculable. It is possible that a steady state was not achievable because of the typical flux behaviour associated with the linlog rate law. At metabolite concentrations of 0, the flux tends to  $-\infty$ , and at concentrations of substrate approaching  $\infty$  the flux tends to  $\infty$ . This gives us some indication that the stability of predictions using the linlog rate law may not be suitable, especially as the number of reactions within a network increases, and the ranges of fluxes and metabolite concentrations that are encountered also increases.

Table 2.3 shows the total error across all 100 steady state points for both the FAM and the DHM models. Given that there are 19 reactions and 17 variable metabolites, there are 3600 individual residuals. This leaves an average error per residual of around 35 units in the FAM models, which is very large considering the average value of each unit is around 1.5. There is a large improvement in predictions for the DHM, especially for the linlog model, which not only provides a full extrapolation across all 100 steady state values, but also has a unit error of around 4. This is the same order of magnitude as the average unit value, so shows that the behaviour approximations are reasonable. In order to achieve these results, it would require determination of all elasticities within the metabolic network. This is unfeasible for larger models.

Concentrations of between 2 mM and 1000 mM could be classified as biologically reasonable extracellular concentrations of glucose for a *S.cerevisiae* culture. It is clear from figure 2.4 (a) that even under these concentrations the FAM models show a poor prediction of system behaviour for all steady states apart from the reference state. In order for the FAM methodologies to be useful as an incorporation in a ‘cycle of knowledge’ [33] with other experimental techniques, they must provide better predictions of system behaviour. As such, MCA network analysis will now be looked at.

### Metabolic Control Analysis

MCA gives us an indication of which reaction  $v_i$  has the most control over the rest of the network. The idea that an increase in the biochemical data added to a model will help improve the prediction accuracy of the model is widespread within the community. The method used here involves taking the reaction flux that demonstrates most network

control and substituting from a ‘first approximation’ to its *in vitro* enzyme kinetics characterised rate law. Here it is available from the original model but in practice it would be experimentally determined. The aim is to present a strategy that yields better prediction quality from minimal laboratory data input. The point is that the model is used to indicate what is the most effective experiment that will maximally determine the model.

Figure 2.2 shows that the convenience (a), hybrid (b) and mass action (c) models all demonstrate similar behaviour. The glucose transporter shows the highest level of control over the flux within the system. Glucose is the primary input of the model, with the dominant flux direction being from glucose influx to trehalose, glycogen, succinate and ethanol efflux. Following this trend, the glycogen, trehalose and succinate branches also demonstrate elevated control over the system. This correlates strongly with the control behaviour seen in the original model, as Figure 2.3 shows.

The glucose transporter has been identified experimentally as showing a high level of control over glycolytic flux in a number of organisms [91, 92, 93]

The conditions under which the models are measured includes growth on glucose as the only carbon source. The first reaction after glucose transport is hexokinase, which demonstrates a high affinity for glucose even at low concentrations, and under physiological conditions is irreversible. Therefore it is likely that most glucose transported into the cell will be quickly committed to the glycolytic pathway. From this, small fluctuations in glucose uptake would directly affect all downstream glycolytic reactions. It is possible that this could be an explanation for the high control shown in glucose transport.

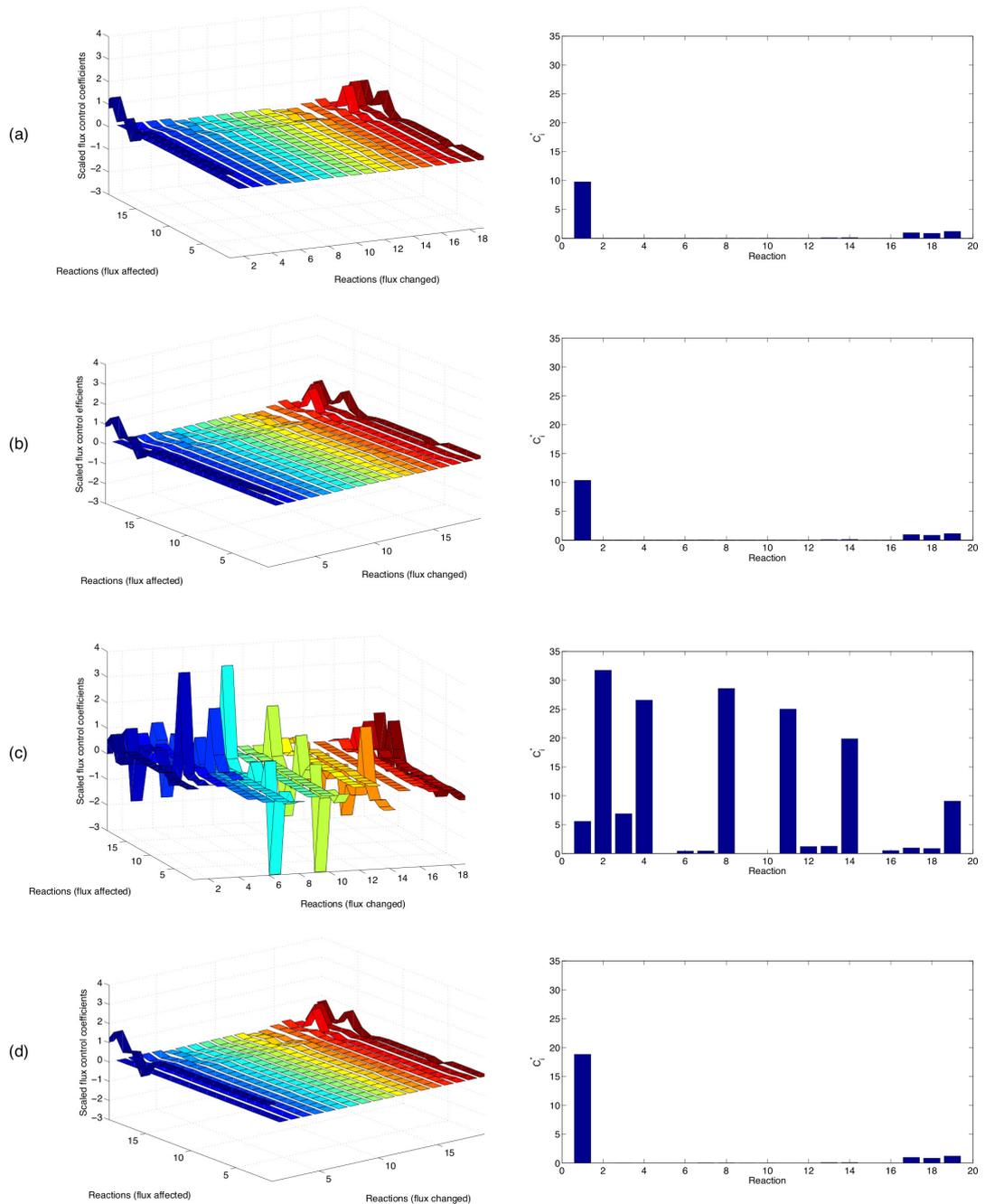


Figure 2.2: Flux control in the FAM models. (a) convenience. (b) hybrid. (c) linlog. (d) mass action. The x and y axis numbers correspond to the following: 1. glucose transport 2. hexokinase 3. phosphoglucoseisomerase 4. phosphofructokinase 5. fructosebisphosphate aldolase 6. triosephosphate isomerase 7. glyceraldehyde phosphate dehydrogenase 8. 3-phosphoglycerate kinase 9. phosphoglyceromutase 10. enolase 11. pyruvate kinase 12. pyruvate decarboxylase 13. alcohol dehydrogenase 14. ATPase 15. adenylate kinase 16. glycerol-3-phosphate dehydrogenase 17. glycogen branch 18. trehalose branch 19. succinate branch. It shows changes in glucose to transport and other transport reactions (glycogen branch, trehalose branch and succinate branch) for convenience, hybrid and mass action rate laws. The control is much more equally distributed in the the linlog model. Flux changed reactions correspond to  $v_i$  and flux affected reaction values correspond to  $C_{v_i}^J$

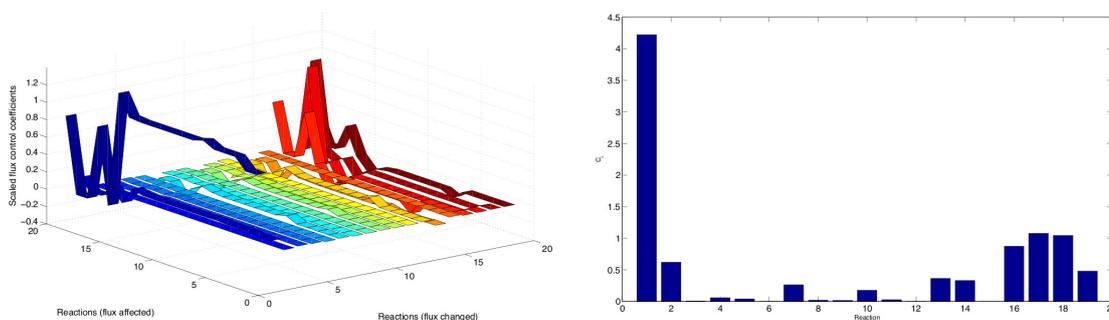


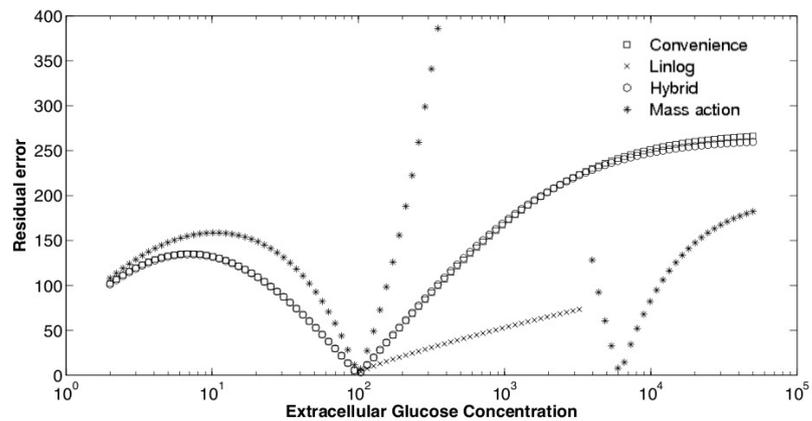
Figure 2.3: Flux control in the Pritchard and Kell model. The x and y axis numbers are labelled according to Figure 2.2

The linlog model, figure 2.2 (c), shows a more distributed control across the system, with larger positive and negative  $C_{v_i}^J$  in comparison to the other three models. The numerically highest flux value, as calculated using equation 2.2.6, was in reaction two: hexokinase, with high control also seen in reaction four: phosphofructokinase and reaction six: triosephosphate isomerase. Previous literature has outlined both hexokinase and phosphofructokinase as having high control over the glycolytic system [94], but this was based on a control analysis that only looked at intracellular metabolites, therefore transport reactions were never taken into consideration [95]. As shown above, more modern approaches to MCA which also consider transport reactions have found glucose transport to be a key regulator of the glycolytic pathway.

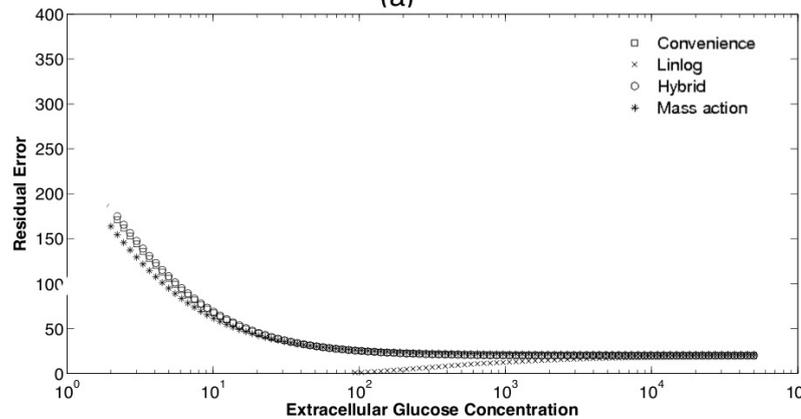
Phosphofructokinase is the only reaction in the original model that contains allosteric regulation. Information on activation and inhibition was not included in the FAM models because, in larger networks, the proportion of regulation known would make up a very small proportion of the network, thereby diluting the effects. One of the regulators associated with phosphofructokinase is AMP. AMP is only involved in one other reaction, and that is adenylate kinase (see Equation 2.3.1) which is a phosphate balancing reaction. Therefore the rate of this reaction should be controlled, to some degree by the phosphate ratio. This ratio is an important aspect of all metabolic networks as it is the driving force that keeps the metabolic behaviour away from equilibrium, at which point the cell would not be living. It is also of vital importance for recycling the depleted ADP into ATP in order to generate energy for living systems. As such it is affected by many reactions within the metabolic network.



Triosephosphate isomerase converts glyceralone phosphate reversibly into D-glyceraldehyde 3-phosphate. This reaction occurs at the branch point of glycolysis. It is believed to play a vital role in energy production (ATP) of the cell, one of the primary goals of glycolysis. Showing, once again that control, at least in the linlog model seems to be associated with reactions that are strongly related to energy production within the cell.



(a)



(b)

Figure 2.4: (a) Stage 1 residual errors. (b) Stage 2 residual errors. The errors are defined as the absolute difference between the concentration and flux values at each steady state point of the FAM models in comparison to the same data calculated from the original model.

Table 2.4: Stage 2 residual errors from dynamic glucose experiments

	Convenience	Linlog	Hybrid	Mass action
FAM residual error.				
Flux	2847	-	2882	2686
Flux and concentration	4135	-	4172	3986
DHM residual error.				
Flux	3592	453	3423	2823
Flux and concentration	4835	469	4677	4115

### 2.3.2 Stage 2 results

#### Changing extracellular glucose on stage two models

Figure 2.4 (b) shows that the convenience, hybrid and mass action models have achieved a marked improvement in predictability of system behaviour. This is most apparent in concentrations that are above the reference state, where the error is reasonably stable. There is little difference between the behaviour of all three rate laws, with mass action appearing to demonstrate the same type of saturation behaviour that is present in the convenience and hybrid rate laws. This would suggest that the new glucose transport rate law, which has been substituted in, reaches its saturable maximum at concentrations near the reference state. The saturation behaviour in this reaction then limits the flux and intracellular concentrations that can be obtained from other reactions, and hence controlling the flux through the system in a very effective way. Below the reference state, where glucose transport is not at its saturable maximum the error increases, but remains no worse than in the first FAM incarnation.

The linlog model still fails to extrapolate for all of the steady states. Figure 2.4 (b) demonstrates an improvement in the predictability of fluxes above the steady state, but the error can be seen to increase in proportion with the increase in extracellular glucose. The DHM (Table 2.4) shows much improvement, extrapolating to all steady state concentrations, and producing the best fit to the original data. In practice, obtaining all elasticities for the network before implementing this methodology would be impractical, and counter intuitive given this methodology aims at reducing the number of laboratory measurements needed for constructing kinetic models. Coupled to this elasticities are difficult to measure *in vivo* because they need to be isolated from the rest of the system to get accurate results.

Table 2.4 shows a reduction in error of an order of magnitude, over all steady states, for convenience, hybrid and mass action models when compared to stage 1 models. When this is split into error per unit, it is approximately 2 per unit. This gives an error of the same order of magnitude as the average unit value, suggesting that the errors within the model are within a suitable range to be able to use the models for predictive purposes. The DHM consistently show a slightly larger residual error for all rate laws, but given that they are of the same order of magnitude the difference in performance is negligible. The data differences between Table 2.3 and Table 2.4 show that it is more important to know the correct kinetic behaviour of the reaction with the most control, than it is to have a number of accurate data measurements for system variables, but no correctly measured kinetics.

## 2.4 Conclusions

The goal of this chapter has been to identify a suitable first approximation strategy that could be implemented in order to help generate larger metabolic models and improve our knowledge of metabolic behaviour. Four first approximation strategies have been developed, extended and tested from existing first approximation rate laws. These have then been tested using a single master strategy, centred around identifying the most important network flux. It is demonstrated that a pure first approximation strategy is not ideal for generating suitable approximations of system behaviour. By contrast, it is shown that correctly being able to identify the reaction flux with the most control over the system, using these rate laws, is of vital importance. When these points of control are identified, their true system kinetics can be measured and substituted in, resulting in a large improvement in model predictability. In reality this means that a good approximation of system behaviour can be produced using minimal experimental elucidation of parameters.

Of the rate laws, convenience, hybrid and mass action all managed to identify a reaction that had considerable control over the system. Upon substitution of the real kinetics, all showed a major improvement in system predictability. Whilst any of these first approximation strategies could be used, the application of the convenience rate law

would be advocated. The convenience rate law demonstrates a faithful representation of enzymatic behaviour, the parameters used within the rate law are easily identifiable with regards to enzymatic behaviour, it is stable with both small and large metabolite concentrations and it produces MCA results which correctly identify the reactions that demonstrate highest control in the system.

Further work will involve the application of this methodology to a larger-scale metabolic network, which will also require the development of a method for thermodynamic handling of reaction loops within the network.

AVOIDING PERPETUAL MOTION MODELS: A  
METHOD FOR ENSURING COMPLEX METABOLIC  
NETWORKS ARE NOT IN VIOLATION OF THE FIRST  
LAW OF THERMODYNAMICS

### **3.1 Introduction**

A metabolic network is a system of enzyme catalysed reactions. These reactions use controlled auto-catalytic combustion and generation of ATP to drive the formation of cellular structures that are vital for the growth and maintenance of the cell. The metabolites are organic compounds that undergo a large number of chemical conversions to transform them from starting metabolites to final downstream metabolic products.

The chemical reactions are therefore subject to the physical laws. The metabolism involves large complex pathways with many different possible routes for metabolite X to be converted to metabolite  $X_n$ . Each metabolite is a composition of bonded elements. To break these bonds the attractions that the elements have for each other have to be overcome, requiring an input of energy. Conversely, forming bonds releases energy because elements capable of bonding are more stable paired than they are as free. The calculation of energy input as bonds in X are broken, and energy output as bonds in  $X_n$  are made, is known as free energy change ( $\Delta G$ ) and can be represented as a function of

enthalpy (H), temperature (T) and entropy (S)

$$\Delta G = \Delta H - (T \cdot \Delta S) \quad (3.1.1)$$

where H is a function of internal energy (U), pressure (p) and volume (V)

$$H = U + pV \quad (3.1.2)$$

It is important that biological models conform to the physical laws. The most pertinent in this instance is the first law of thermodynamics, stating that energy must be conserved. This means that multiple pathways connecting X and X<sub>n</sub> must involve the same  $\Delta G$  in order to satisfy the first law and prevent the model being a perpetual motion machine.

Figure 3.1 can be taken as an example, it shows two different network structures, similar to what would be found in smaller metabolic models of a network. Network (a) shows the metabolites to be produced through independent reaction pathways. Network (b) shows that metabolite F can be produced from metabolite B either by  $v_5 + v_6$  or by the route  $v_2 + v_3 + v_4$ .

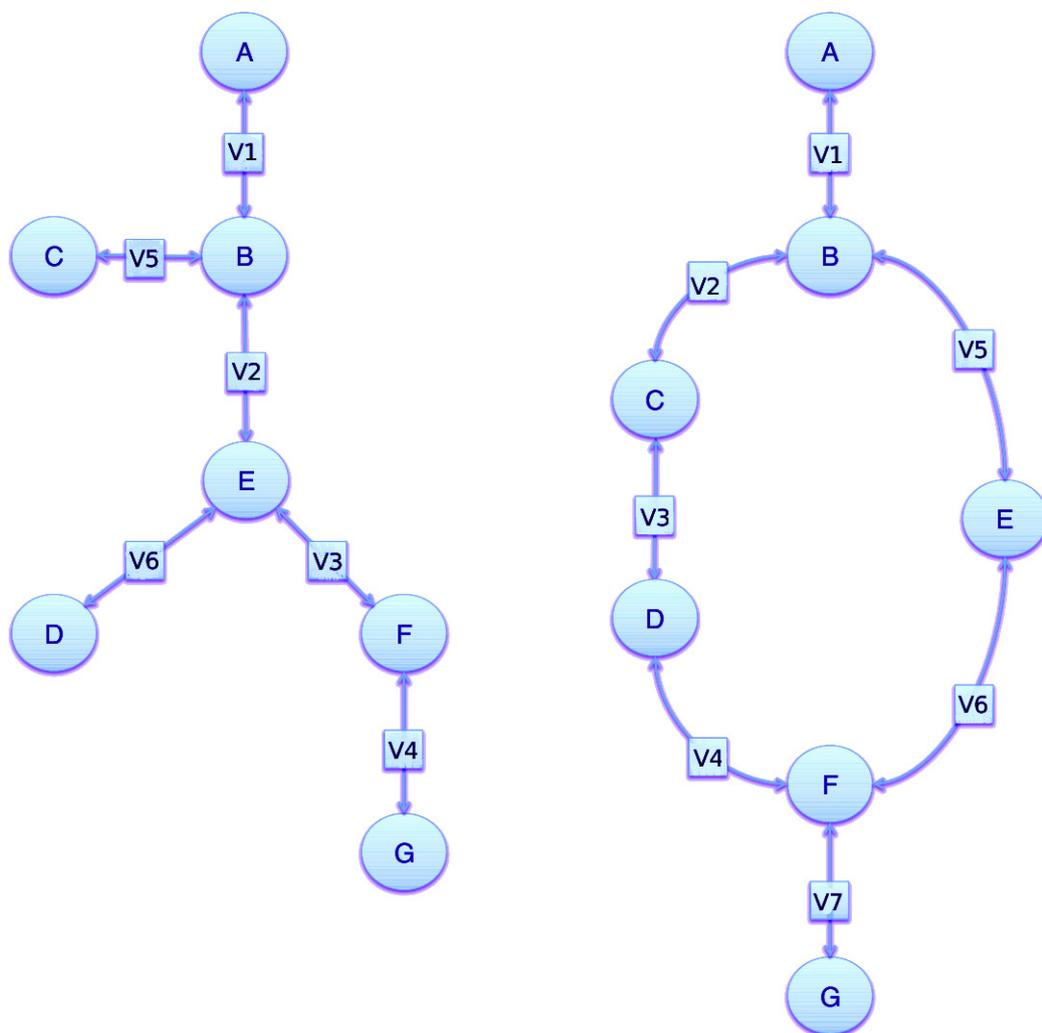


Figure 3.1: (a) Network with no linear dependency. (b) Network with linear dependency.

Table 3.1: Equilibrium constants associated with the reactions in network (b).

Reaction	$K_{eq}$
$v_1$	5
$v_2$	2
$v_3$	2
$v_4$	4
$v_5$	6
$v_6$	2
$v_7$	8

The  $\Delta G$  associated with the conversion of  $B \rightarrow F$  holds a finite value of  $x$ . This is

true irrespective of the number of reactions that separate the metabolites. For instance, if  $B \rightarrow F$  via  $v_5 + v_6$  has a  $\Delta G = -8$ , the conversion of  $B \rightarrow F$  via  $v_2 + v_3 + v_4$ , must also have a  $\Delta G = -8$ . The reverse reaction  $F \rightarrow B$  must hold an equal, opposite value, in this example a  $\Delta G = 8$ . To traverse all reactions in the ‘loop’  $B \rightarrow B$  there is no change in bonding within the compound, resulting in a  $\Delta G = 0$ . For any metabolic network, the first law of thermodynamics requires that the sum of  $\Delta G$ 's around any loop equal 0.

The rate of reaction ( $V$ ) in a metabolic model is a function of substrates ( $S$ ), products ( $P$ ), the equilibrium constant ( $K_{eq}$ ) and other constants ( $c$ )

$$V = f(S, P, K_{eq}, c) \quad (3.1.3)$$

$K_{eq}$  is related to  $\Delta G$  via

$$K_{eq} = \exp(-\Delta G/RT) \quad (3.1.4)$$

where  $R$  is the universal gas constant and  $T$  is the absolute temperature.

The relationship between  $\Delta G$  of a given metabolic transformation and the  $K_{eq}$  make it pertinent that all reactions within a network have thermodynamically consistent  $K_{eq}$  values to prevent models in perpetual motion. To make sure a metabolic model is thermodynamically consistent  $K_{eq}$  must be reverted into  $\Delta G$  so the reactions can be summed across the loops. Thus a metabolic network is thermodynamically consistent if and only if the sum of its  $\log K_{eq}$  values around any loop equals zero.

$$\Delta G \propto \log K_{eq} \quad (3.1.5)$$

The examples in Figure 3.1 are simple, making it easy to visually identify where a linear dependency exists, and to assign a solution out of the infinite combinations available. In a real network this would be much more difficult. The linearly dependent reactions can be identified by calculating the stoichiometric matrix ( $N$ ) with no reaction bounds ( $N_o$ ) and then taking the null space of the matrix.

The stoichiometric matrix for network (b) is as follows

$$N_o = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.1.6)$$

With columns representing reactions and rows representing metabolites.

leaving a basis for the null space (Null  $N_o$ )

$$\text{Null } N_o = \begin{bmatrix} 0 \\ -1 \\ -1 \\ -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad (3.1.7)$$

Vectors in the null space of  $N_o$  are internal loops and hence a thermodynamically feasible set of  $K_{eq}$  values, and must satisfy

$$\text{Null } (N_o^T) \log[K_{eq}] = 0 \quad (3.1.8)$$

The  $K_{eq}$  values of network (b) are outlined in table 3.1. They correspond to one of the infinite solutions of this network, providing thermodynamically feasible values.

$$\text{Null } N' \cdot \begin{bmatrix} 5 \\ 2 \\ 2 \\ 4 \\ 6 \\ 2 \\ 8 \end{bmatrix} = 0 \quad (3.1.9)$$

Where networks contain multiple loops there will be a vector associated with the loops generating a matrix. Each column vector must hold the relationship outlined in Equation 3.1.9 for this to be a feasible solution.

## 3.2 Testing on real reaction networks

In 2004, Holzhütter applied a method of flux minimisation during flux balance analysis to two models of metabolic networks: a model of erythrocyte metabolism and *Methylobacterium extorquens* (originally taken from [96]). The method relied on using the equilibrium constants to dictate the balance between forward and reverse fluxes of a given reaction. Given the importance of the equilibrium constants in the methodology, the thermodynamic consistency of the two networks should be tested.

### Erythrocyte metabolism

The stoichiometric matrix was calculated from the kinetic model (available from [7]), from which the null space was computed in Matlab. The  $K_{eq}$  vector was taken from values outlined in the paper. The  $K_{eq}$  values were also computed from the kinetic model by isolating each reaction in turn, simulating the reaction until the point of equilibrium and then calculating the mass action ratio. In some instances there was a difference in these values. As examples, from the linearly dependent reactions outlined in Table 3.2, bisphosphoglycerate phosphate has a  $K_{eq}$  value of 1 in the model, but 100000 in the paper. Similarly, ATPase is represented as an irreversible reaction in the kinetic model, with no associated  $K_{eq}$  value. The value used in the paper (100000) is quite large, however, and

favours the forwards direction in most cases, therefore this reaction is still representative of the model.

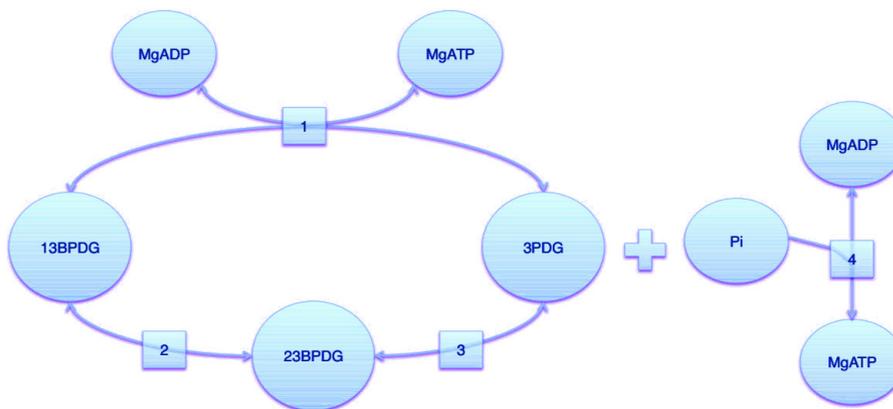


Figure 3.2: Schematic of reactions outlined in Table 3.2

Table 3.2: Linearly dependent reactions in the erythrocyte model by Holzhütter [1].

Reaction	#	Equation	Original $K_{eq}$	New $K_{eq}$
Phosphoglycerate kinase	1	MgADP + 1,3-bisphospho-D-glycerate = MgATP + 3-phospho-D-glycerate	1455	1455
Bisphosphoglycerate mutase	2	1,3-bisphospho-D-glycerate = 2,3-bisphospho-D-glycerate; MgGri23P2	100000	100000
Bisphosphoglycerate phosphatase	3	2,3-bisphospho-D-glycerate = 3-phospho-D-glycerate + phosphate; MgGri23P2	100000	$1.00e^{-5}$
ATPase	4	MgATP = Phosphate + MgADP	100000	$6.91e^{-4}$

Using the equilibrium constants in column 4 of Table 3.2, the following result was obtained:  $\text{Null } N_o^T \log \mathbf{K}_{eq} = 2.1151$ . This shows that the equilibrium values used were not a solution of the system, and were therefore in violation of the first law of thermodynamics.

Given that bisphosphoglycerate phosphatase and ATPase were the two values that were least consistent with the kinetic model these were allowed to vary in order to calculate a feasible system solution. A particle swarm optimisation [97] was used, using a least squares fit to try and compute a viable solution to the system. The results obtained are outlined in column 5 of Table 3.2. Here it can be seen that, in order to prevent a violation of the first law, the values bisphosphoglycerate phosphatase and ATPase must be many orders of magnitude smaller than the values chosen in the paper (and the kinetic model).

The weighting of flux direction in the paper were assigned by noting whether the  $K_{eq}$  values showed a strong or weak affinity for the forwards reaction, this huge change in  $K_{eq}$  value could prove to have a large impact on the flux prediction of the system.

### *Methylobacterium extorquens*

The *methylobacterium extorquens* model (outlined in Table 3 of Holzhütter (2004) [1]) is an adaption of the metabolic network outlined in Dien and Lidstrom (2002) [96]. A summary of the linearly dependent reactions within the network can be found in Table 3.3 and Figure 3.3. Here the network and  $K_{eq}$  values were analysed using the same method as above. Many of the  $K_{eq}$  values in this model were assigned according to the known behaviour of the reaction, so a reaction close to equilibrium was assigned  $K_{eq} = 1$ . Reactions believed to be far from equilibrium were assigned  $K_{eq} = 100000$ .

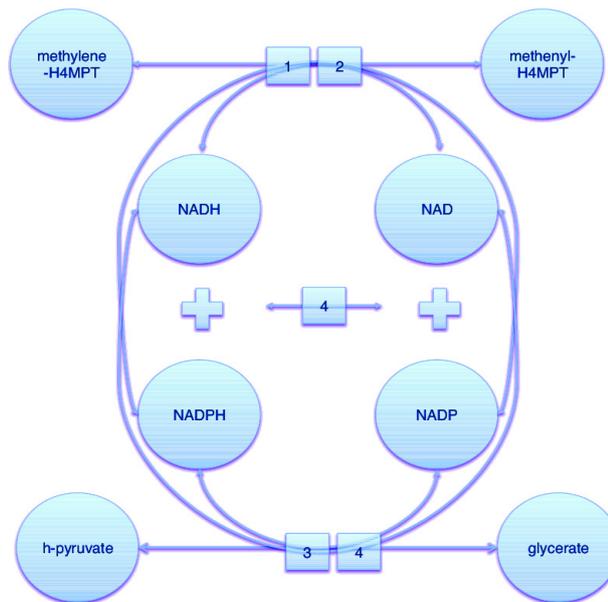


Figure 3.3: linearly dependent reactions in *Methylobacterium extorquens*

Table 3.3: Linearly dependent reactions in the *Methylobacterium extorquens* metabolic model outlined in Holtzhütter [1] modified from Dien and Lidstrom (2002) [96].

Reaction	#	Equation	Original $K_{eq}$	New Eq $K_{eq}$
Methylene H4MPT dehydrogenase (MtdA)	1	methylene-H4MPT + NADP = methenyl-H4MPT + NADPH	1	1
Methylene H4MPT dehydrogenase (MtdB)	2	methylene-H4MPT + NAD = methenyl-H4MPT + NADH	1	1
Hydroxypyruvate reductase	3	h-Pyruvate + NADH = glycerate + NAD	1	1
Hydroxypyruvate reductase 2	4	h-Pyruvate + NADPH = glycerate + NADP	1	1
Transhydrogenase	5	NADH + NADP = NADPH + NAD	1	1

The *methylobacterium extorquens* model shows the linearly dependent reactions to be thermodynamically consistent with  $\text{Null}N_o^T \log \mathbf{K}_{eq} = 0$ . For this model it is assumed that the all substrates and products in the linearly dependent reactions have the same free energy, hence no reaction has an associated  $\Delta G$ . This is certainly not an ideal way of assigning  $K_{eq}$ , whilst it does prevent violation of the first law, the biological assumptions are incorrect.

### 3.3 Conclusions

Here a method has been developed that can ensure all reactions, within a network loop, contain  $K_{eq}$  values that are consistent with the first law of thermodynamics. This is vital in metabolic modelling to ensure that the model is not in perpetual motion, a state which causes the model to be unrealistic and unstable. The method was applied to two models available from Holtzhütter (2004) [1] and showed that the erythrocyte model was not thermodynamically consistent. It was also shown that a thermodynamically consistent solution to a network could be calculated by altering only the  $K_{eq}$  values where there is less confidence associated with their measurement. This allows a greater ability to compute network solutions with input from actual  $K_{eq}$  measurements. The method also highlighted that it was possible to circumvent the violation of the first law by assigning a  $\Delta G$  value of 0 ( $K_{eq} = 1$ ), which would not be recommended as this assumes no

metabolite identity change, or little structural change, between metabolic reactions. The method also highlighted a multi-solution problem, where there are an infinite number of solutions for each loop. A solution to this issue would be to obtain as many accurately measured  $K_{eq}$  contained in the loop to constrain the solution more stringently.

GENERATING THERMODYNAMICALLY-CONSISTENT  
AND EXPERIMENTALLY FOUNDED KINETIC  
MODELS AT THE GENOME-SCALE

## 4.1 Introduction

Gaining a quantitative understanding of metabolic behaviour has long been a major scientific goal, beginning with crude mass balance experiments [98] and progressing through enzyme kinetics, single-pathway models [99, 75, 15] and collaborative efforts such as the yeast reconstruction [79, 80] and the digital human [100]. Traditionally the desire to understand metabolic behaviour has been driven by the need to understand phenotypic responses to disease [101]. Being an amplification of the genome, transcriptome and proteome, it is believed that metabolism is key to understanding the chemistry of life [98].

Yeast has notable uses in the biotechnology and brewing industries. Metabolic products produced and extracted, can be used safely for drugs and fuels. This is highly beneficial if the metabolite of interest is difficult, expensive or polluting to produce using traditional organic synthesis techniques [102]. For the brewing industry, being able to have finer control over ethanol production would also be a huge advantage.

Predominantly, most work on understanding the metabolism has focussed on bottom-up elucidation of smaller pathways such as glycolysis [15, 14] and the pentose phosphate

pathway [99]. The repositories where these models can be found [7, 103] contain relatively few metabolic pathways beyond central carbon metabolism. There are also some approaches that focus on understanding the metabolism from a top-down approach such as Flux Balance Analysis (FBA) [104, 48, 105, 59], and dynamic Flux Balance Analysis (dFBA) [51, 106, 107]. Flux balance analysis is a linear programming technique that uses flux constraints, coupled with physiochemical constraints and energy requirements in order to get an understanding of the potential flux distributions through the pathways. An iterative implementation of FBA forms the basis of dFBA. Here FBA is run and the results obtained, the initial conditions are then altered based on this result, before FBA is run again. This produces pseudo-dynamic results.

Unfortunately, flux behaviour without metabolite concentrations and some idea of kinetics tells us very little. For example, a flux of 0 can indicate no activity through the pathway, or a reaction under equilibrium conditions. The metabolite pools within a system can also increase or decrease concentration drastically with little change in flux [108]. It is therefore pertinent that a move is made to larger-scale kinetic modelling to gain a deeper understanding of metabolic behaviour.

Improvements have been made to the top-down approaches such as modifying the dynamic flux balance analysis approach [60] to include pseudo-kinetic behaviour. Perhaps the most profound advances have been made in 2010 by Smallbone and colleagues [109] and Li and colleagues [110]. Smallbone (2010), takes a first approximation methodology developed in an earlier paper [89], which uses knowledge of the stoichiometric matrix and flux balance analysis approaches, to populate the original yeast consensus model [79] with linlog rate laws. Li (2010), uses a workflow approach that takes a selected network, applies the kinetic rate-law that is stored in Sabio RK or, where none is available, a generic rate law, to the network and then populates the values with data obtained from various databases. Where no information is available it populates the parameters with a value of 1. These are both important papers in the progression of metabolic modelling and provided the foundations of this methodology.

Here I present a methodology that retains collective advantages from the two approaches, in the form of:

- Modified flux balance analysis which provides the ability to obtain a realistic and balanced distribution of network fluxes.
- The rate law used contains a small number of parameters that require values, making their calculation and assignment less complex.
- It is adaptable to any network or organism.
- The reaction kinetics used mimic actual enzymatic behaviour, making the approach biologically feasible.

whilst addressing the weaknesses in order to create a stronger methodology, which:

- Ensures that the equilibrium constants and fluxes are all thermodynamically feasible.
- Allows the potential for full integration of data taken ‘in house’ and from databases, including metabolite concentrations, dynamic flux values, equilibrium constants and full kinetic rate laws, to be integrated within the model, whilst minimising data inconsistencies.
- The design of the methodology ensures that all variables within the system are consistent and will produce a biologically plausible pre-defined steady state, and achieve an equilibrium state when isolated from its surroundings. Overcoming some of the issues that can be associated with the ‘per-reaction’ substitution of rate laws.
- The strategy has been shown in Chapter 2, to provide solid indications, through the use of metabolic control analysis, which reactions are most important to the behaviour of the system and should be correctly measured in the laboratory.

The method uses the stoichiometric network as a scaffold and always considers the full network in the computation of all fluxes, metabolite concentrations and equilibrium constants. The completed model will adhere to the following pre-defined conditions:

1. If different reactions share the same quantities (i.e. the concentration of a shared metabolite) their values must be the same for each of the reactions.

2. The fluxes must be at steady state (i.e. for every internal metabolite incoming and outgoing fluxes must balance each other).
3. The equilibrium constants must be balanced such that they adhere to the first law of thermodynamics (see Chapter 3)
4. The mass action ratio of each reaction must be consistent with the flux directionality computed with FBA.

All of which are selected to ensure consistent network behaviour.

## 4.2 Method

To construct the metabolic model, the work-flow outlined in Figure 4.1 was followed, where letters represent data input/output and numbers represent methods to process the data into a consistent steady-state model of yeast metabolism. The data structures used in the model are designed so there is a seamless integration between accurately measured data, and data calculated using first-approximation methodologies. The methods used to process the data allow a layered construction of the metabolic model, resulting in a kinetic network that holds the desired steady-state.

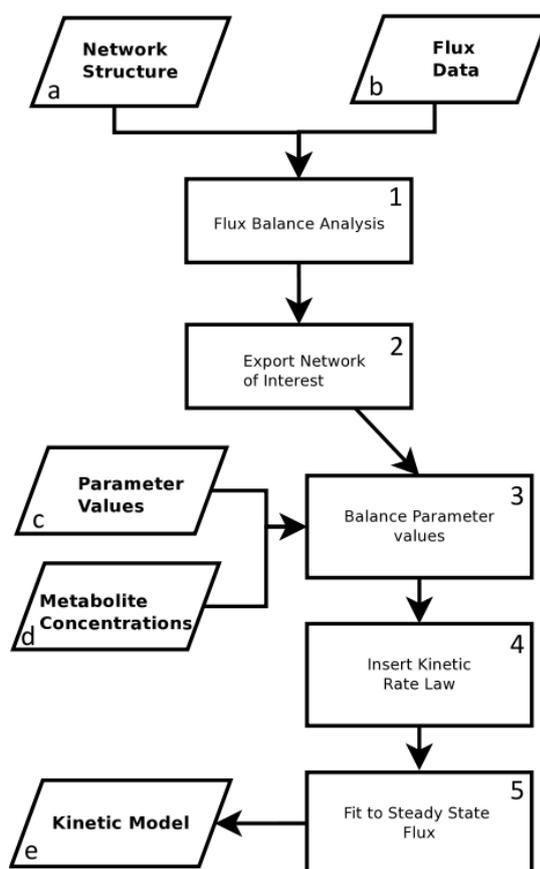


Figure 4.1: Here a schematic of the workflow is presented showing how different data sets can be integrated in order to generate the large-scale first approximation model. All data are obtained, manipulated and layered into the model systematically, allowing independent data sets to be merged effectively. The final output is a model that holds the pre-defined steady state.

### Network structure [Data a]

The network structure used was Yeast 4.0 [80]. The model is larger in size than other available yeast models, with 1102 unique metabolic reactions and 924 unique metabolites. The connectivity of the network is also high, with around 90% of the metabolites reachable from the extracellular media. The model includes an extensive proportion of lipid metabolism. Lipid metabolism is a vital component for biomass production, making the network behaviour more biochemically sound. It has been noted that this model is not as good at reproducing behaviour from gene knockout experiments as other models. Given the increased biochemical correctness of this model, and that the intended end use is to apply kinetics before further network analysis, it is believed that using this network is more justifiable than using previous reconstructions for this method. Where no flux information was available (see Section 4.2 for further information), the

flux boundaries were kept as the original values outlined in the model.

### Known flux values [Data b]

Table 4.1 shows the selection of yeast specific fluxes that were used as constraints for the FBA. These can be calculated from yeast specific models, with glucose as the primary carbon source, in the biomodels database [7]. They were set as fixed values, where possible, and for those that caused issues in convergence of the algorithm (see Section 4.2 for further information) the boundaries of the value were expanded.

Table 4.1: Steady state fluxes calculated from yeast specific models in the biomodels database that use glucose as the primary carbon source.

Reaction	Flux (mM sec <sup>-1</sup> )
alcohol dehydrogenase, reverse rxn (acetaldehyde → ethanol)	1.17
ATPase, cytosolic	.595
enolase	1.76
fructose-bisphosphate aldolase	.733
glucose-6-phosphate isomerase	.733
glyceraldehyde-3-phosphate dehydrogenase	1.06
glycerol-3-phosphatase	.051
glycerol-3-phosphate dehydrogenase (NAD)	.149
hexokinase (D-glucose:ATP)	.866
phosphofructokinase	.606
phosphoglycerate kinase	.875
phosphoglycerate mutase	1.76

It should be noted that for construction of a specific large-scale model, much care should be taken with which flux bounds are used. Models are typically constructed with the intention of answering a specific biological question. This leads to models demonstrating flux under very different biological conditions, which would not be consistent with each other. To ensure a biologically sound model, data should only be used from models and experiments that use similar growth conditions. Data calculation from models may also require flux scaling for time-scale consistency.

### Parameter Values [Data c]

The equilibrium constants ( $K_{eq}$ ), were taken from kinetic rate laws available in the yeast specific models in the biomodels database, that use glucose as the primary carbon source. All transport reactions were assigned a  $K_{eq}$  value of 1. Where values were not available they were computed using a best fit. The lower bound for each value was set at 1.1 times the value of the mass action ratio of the associated reaction. The upper bound was set as  $1 \times 10^{+6}$  (this is for linearly dependent reactions). All other values were set as 1.1 times the mass action ratio. For information regarding the concentrations used for the mass action ratios see Section 4.2.

Table 4.2: Equilibrium constants taken from yeast specific models in the biomodels database that use glucose as the primary carbon source.

Reaction	Equilibrium constant
hexokinase (D-glucose:ATP)	2000
glucose-6-phosphate isomerase	0.290
fructose-bisphosphate aldolase	0.045
glyceraldehyde-3-phosphate dehydrogenase	3200
phosphoglycerate mutase	6.700
enolase	6500

Of the models used, some reactions were set as constant rates, for these reactions it was not possible to obtain the  $K_{eq}$  values. As such, these reactions were treated as unknown. For the fitting procedure the linearly dependent reactions only contained the constraints of two equilibrium constants, this led to a fit being achieved relatively easily, given there are likely a large number potential solutions to the system. Where more information is known the number of solutions to the system will decrease. This is likely to lead to increased difficult in fitting a consistent set of equilibrium constants. Should such issues arise, modifications to the method would have to be sought.

**Metabolite concentrations [Data d]**

The intracellular metabolites were taken from yeast specific models from the biomod-els database, that use glucose as the primary carbon source (see table 4.3). Where concentrations were not known, the intracellular concentrations were taken to be the median value of the intracellular values in Table 4.3 (0.176 mM). For the extracellular concentrations extracellular metabolome measurements were used where possible, where no data was available the median value of all extracellular concentrations (11.1 mM) was used (see Table 4.4).

For the generation of a more specific large-scale metabolic model, it would be advised that concentrations are collected using quantification analysis of the intra- and exo-metabolome specific to the desired organism, and biological state. Where more metabolite data is available it may also be possible to use thermodynamic FBA to help improve the estimated values of metabolites of unknown concentration [61].

Table 4.3: Yeast specific intracellular metabolite concentrations taken from yeast specific models in the biomodels database that use glucose as the primary carbon source.

Intracellular metabolite	Concentration mM
2-phospho-D-glyceric acid	$3.70 \times 10^{-2}$
3-phospho-D-glyceric acid	$2.78 \times 10^{-1}$
3-phospho-D-glyceroyl dihydrogen phosphate	$2.75 \times 10^{-4}$
acetaldehyde	$1.70 \times 10^{-1}$
ammonium	$3.80 \times 10^{+1}$
AMP	$7.96 \times 10^{-1}$
ATP	1.13
beta-D-glucose 6-phosphate	1.02
carbon dioxide	1.00
D-fructose 1,6-bisphosphate	2.82
D-fructose 6-phosphate	$1.12 \times 10^{-1}$
D-glucose	$9.10 \times 10^{-2}$
ethanol	$5.00 \times 10^{+2}$
glyceraldehyde 3-phosphate	$6.90 \times 10^{-2}$
glycerol	2.27
glycerone phosphate	$5.90 \times 10^{-1}$
L-leucine	1.00
L-lysine	1.00
L-serine	1.00
L-threonine	1.00
L-tryptophan	1.00
L-valine	1.00
NAD(+)	1.50
NADH	$8.60 \times 10^{-2}$
phosphoenolpyruvate	$3.00 \times 10^{-2}$
pyruvate	8.36
sn-glycerol 3-phosphate	$4.57 \times 10^{-1}$

Table 4.4: Extracellular metabolite concentrations taken from extracellular metabolome measurements.

Extracellular metabolite	Concentration mM
D-Glucose	11.1
Glycerol	1.76
H+	$3.20 \times 10^{-2}$
O <sub>2</sub>	12.0
Sulfate	42.2
Median value	11.1

The assignment of many of the metabolite concentration values, through their relationship with the mass action ratio, have a direct impact on the bounds for computing the equilibrium constants, and therefore the favoured direction of the transport reactions. If these are found to be producing flux directions that are contrary to known network behaviour (either from models or experimental knowledge) or where they are causing issues with equilibrium fitting, these should be altered to more accurately reflect the known system behaviour.

### Flux Balance Analysis (FBA) [Step 1]

The flux values outlined in Table 4.1 were used as the lower and upper flux bound for their corresponding reaction in the Yeast 4 [80] network. The original flux bounds outlined in the Yeast 4 model were used for all other reaction fluxes. Geometric FBA was run on the network using the COBRA toolbox [55] in Matlab.

In some instances there may not be a solution to the system using the exact flux constraints outlined. This could be a result of experimental error in the laboratory data, or simply due to network constraints in the larger model. Alterations in fluxes are unlikely to cause large issues within smaller systems because there is a lower interdependence between the reactions. To encourage a solution close to the original fluxes the bounds of the conflicting reactions should be set as close as possible to the original.

### Exporting the network of interest [Step 2]

After the first solution had been computed, all reactions with a flux of 0 were removed to generate a reduced network. Table 4.5 shows which KEGG [111, 112, 113] pathways the reactions in the reduced model are drawn from. For a more detailed list of reactions contained in the model see Appendix A.

Reactions with negative fluxes were reversed to produce a positive flux for all reactions. This is not strictly necessary but simplifies the computation required for further stages of the methodology. The reduction of the network to a smaller kinetic version is recommended because of the limited amount of known data contained within it. A larger network produces a significantly higher dilution of the known data and could potentially compromise the behavioural output of the final model. As more data is obtained the network can be allowed to increase in size.

It is possible to reduce the network using a finite flux value cut off point (e.g all fluxes below  $1 \times 10^{-5}$ ). In these instances the FBA needs to be recalculated on the new network in order to ensure that there is: (a) still a solution to the system (it is possible that pathways producing important components of biomass may be interrupted) and (b) the solution must be adjusted to ensure it is still representative of a steady state of the system (this is vital to ensure that a steady state kinetic model can be produced from the flux scaling (see Section 4.2)).

### Balancing parameters [Step 3]

The known  $K_{eq}$  (Table 4.2) are kept as fixed within the model. All unknown  $K_{eq}$  can be calculated using the methodology outlined in Chapter 3, using the boundary conditions outlined in Section 4.2. The unknown equilibrium constants have the lower bound set such that, at the desired steady state, the metabolite concentrations and flux directions are thermodynamically consistent. Coupled to this, they are fitted such that isolated metabolic system will attain an equilibrium state and not violate the first law of thermodynamics.

The primary output that is required for a thermodynamically consistent system is outlined in Equation 4.2.1, where  $N_o$  is the stoichiometric matrix calculated with no boundary conditions, and  $\mathbf{K}_{eq}$  is the vector of equilibrium constants.

### Equilibrium Solution

$$\mathbf{N}_o^T \log \mathbf{K}_{eq} = 0 \quad (4.2.1)$$

For this method the  $K_s$  and  $K_p$  values were set to be of the same value as their corresponding metabolite concentration ( $S$  or  $P$ ).

### Kinetic rate law [Step 4]

The convenience kinetics rate law [82] seen in Equation 4.2.2, which resembles a random order binding enzymatic mechanism, has been shown in Chapter 2 to hold an advantage over rate laws such as linlog [89, 109] for reproducing desired behaviour when applied to a metabolic system. As equation 2.1.2 shows, the rate law requires knowledge of metabolite concentrations and a way of populating the parameters  $V_m$ ,  $K_{eq}$ ,  $K_s$ ,  $K_p$ .

The parameters were populated according to the methodology developed in Chapter 2, along with the flux and equilibrium extensions developed in Sections 4.2 and 4.2, respectively.

$$V = V_m \frac{\prod_i S_i^{n_i} - K_{eq}^{-1} \prod_j P_j^{n_j}}{\prod_i \left(1 + \frac{S_i}{K_s}\right) + \prod_j \left(1 + \frac{P_j}{K_p}\right) - 1} \quad (4.2.2)$$

$S$	= Substrate concentration	$P$	= Product concentration
$K_{eq}$	= Equilibrium constant	$V_m$	= Rate constant
$K_s$	= Substrate Michaelis constant	$K_p$	= Product Michaelis constant

### Fitting to steady state flux [Step 5]

The initial flux through each reaction was calculated computationally,  $V_m$  was then scaled so that the initial flux matched the FBA flux result. This process ensures that the system is at steady state at the point of generation.

#### 4.2.1 Metabolic control analysis

To identify reactions which demonstrate the most control over the system, MCA of the final model was calculated using COPASI and the resulting flux control coefficients were computed according to the method outlined in Chapter 2 Section 2.2.3.

The reactions that demonstrated a  $C_i^*$  value of more than 95% of the highest  $C_i^*$  were identified and selected as the reactions demonstrating most control over the system.

## 4.3 Results and discussion

The final network contains 285 reactions with 294 metabolites. The method results in a steady state flux distribution that is the same as the flux balance analysis result, and steady state concentrations that are the same as the concentrations input as the initial concentrations. The full set of fluxes and concentrations can be found in Appendix A.

### 4.3.1 Data input [Data A to D]

In this instance the network is a generic reconstruction that reflects the current known metabolic reactions. For testing the methodology this was a reasonable approximation. When applying the methodology for specific cases it is important that any gene-knockouts within the strain that is being studied are identified, and the flux bounds altered accordingly, so that the biochemistry of the network is as similar as possible. This would also apply to other organisms, where an organism specific reconstruction should be used, and then modified to generate a system representation that is as accurate as possible.

The concentrations and fluxes used in the model were the median values of known yeast specific fluxes and concentrations in the Biocompare database. The aim was to test the methodology to produce a generic kinetic model of the yeast metabolism, as such these data were a generic representation of the behaviour that known yeast models produce. This method could be carried out for any organism where there is a sufficiently large and biochemically sound network reconstruction for use as the scaffold. In order to further expand the methodology to more specific metabolic and yeast strain models more careful consideration of data input is vital. It would be best practice to use flux and concentration data taken from models that are constructed to investigate similar biological phenomena, and under similar stress conditions. Where possible it would be recommended to obtain real intracellular metabolite concentrations using quantitative metabolomics. It would also be useful to obtain real quantitative flux data, however this can be complex to measure

[24].

Where the equilibrium constants were known, the flux values of the system were set such that they were thermodynamically consistent. For unknown equilibrium constants, the values were set, such that they were thermodynamically consistent with the system fluxes. To achieve this the relationship between concentrations, the mass action ratio, equilibrium constants and reaction directionality were used. The output is similar to that outlined by Henry and colleagues in 2007 [61]. The difference here is that more confidence is placed in the knowledge of the fluxes and concentrations, leaving the equilibrium bounds to be defined afterwards, indirectly assigning Gibbs free energies. In the Henry paper, there is a reliance on structurally computed Gibbs free energies, which are known to have more than a 50% error rate [87]. Given the uncertainty within the values, and also that there would be only a small number of known concentrations to apply to the network it was believed that the current methodology was preferable.

### **4.3.2 Processing steps of the pipeline**

#### **FBA result, and exported network of interest [Step 1 and 2]**

Table 4.5 shows that purine and pyrimidine metabolism accounts for a large proportion of the reactions. The nucleotides (adenine, guanine, cytosine, thymine and uracil), which are required for DNA and RNA synthesis, contain purine and pyrimidine structures. This indicates that a large number of reactions are devoted to making DNA and RNA within the cell. The production of RNA in particular is vital for the cell to generate the proteins that catalyse the metabolic reactions to generate energy. Glycolysis and the citric acid cycle combined contain 23 reactions, which also covers a large proportion of reactions in the network. Glycolysis generates ATP either anaerobically via fermentation into ethanol or aerobically via the citric acid cycle. ATP is a high energy metabolite capable of transferring a phosphate group in order to make the energetics of unfavourable reactions more viable. The use of ATP to drive glycolysis is an autocatalytic process, where more ATP is generated as a result. A higher concentration of ATP can be generated using the citric acid cycle, rather than by fermenting, but most experimental conditions for growing yeast are set up such that fermentation is the primary goal.

Lipid metabolism also accounts for a large proportion of the reactions within the network. Lipid generation is vital for processing components of the cell, such as the membranes. Therefore the lipid metabolism is important in the generation of biomass of the cell, thereby accounting for its heavy presence in the reduced model.

It is possible that the pathways that appear to account for less reactions in the network may not be as important for the growth of a cell. However, this would be based on the assumption that all reactions could be exclusively defined as part of a single pathway. In reality reactions can form part of multiple pathways, which are somewhat arbitrarily ‘ring-fenced’ in KEGG and text books, so it is possible that reactions attributed to other pathways also form part of the pathways that demonstrate a lower number of network reactions.

Table 4.5: Kegg pathways represented in the model

Pathway	Number reactions
Purine and pyrimidine biosynthesis	27
Lipid metabolism	19
Valine, leucine and isoleucine biosynthesis	17
Glycolysis / gluconeogenesis	16
Nucleotide salvage pathway	15
Tyrosine, tryptophan, and phenylalanine metabolism	15
Arginine and proline metabolism	12
Carbohydrate and lipid metabolism	11
Histidine metabolism	10
Biosynthesis of secondary metabolites	9
Cofactor and prosthetic group biosynthesis	9
Alternate carbon metabolism	8
Transport, outer membrane porin	8
Citric acid cycle	7
Glycerophospholipid metabolism	7
Glycerolipid metabolism	5
Membrane lipid metabolism	5
Pentose phosphate pathway	5
Threonine and lysine metabolism	5
Biosynthesis of unsaturated fatty acids	4
Cell envelope biosynthesis	4
Cysteine and methionine metabolism	4
Glycine and serine metabolism	4
Lysine biosynthesis	4
Oxidative phosphorylation	4
Alanine and aspartate metabolism	3
Anaplerotic reactions	3
Sulfur metabolism	3
Misc	42

### Balanced parameter values and kinetic rate laws [Steps 3 and 4]

Only the equilibrium constants were balanced in this model. Leaving the model with thermodynamically consistent equilibrium constants (and therefore Gibbs free energies),

metabolite concentrations and flux distribution. The Michaelis constants were not thermodynamically balanced. It is possible that this could be improved by implementing global linear regression reliant parameter balancing, such as that outlined by Lubitz and colleagues in 2010 [114]. However, the addition of extra constraints, and system level balancing to such a large network, could result in difficulty fitting a thermodynamically consistent solution for the network. This would also preclude that ability to manipulate the Michaelis values in order to produce different dynamic reaction profiles, which is known to be advantageous in producing networks that also predict transient network behaviour. The improvement of the method for transient data is a large priority for future work. Therefore inclusion of full parameter balancing must be carefully considered.

### **MCA analysis**

The MCA result (table 4.6) shows that the flux control is heavily dominated by transport reactions within the network. A change in these parameters on a short time scale will result in a depletion of, or increase in metabolite pool concentrations across the network, depending on whether the flux is increased or decreased. Like a traffic jam, the effect can rapidly expand. Longer term changes in these parameters should alter the system so it attains a new steady state, with behaviour that balances the cells requirements with the new extracellular limitations. This shows that getting the correct kinetics for the transport reactions and having a full understanding of the extracellular concentrations within the growth media, is a vital step in producing a model that demonstrates good predictability. Exact kinetic measurement of transport reactions is very difficult, it requires using an intact membrane structure which is only possible *in vivo*.

A number of glycolysis reactions also show a high degree of control over the system. Glucose is the primary carbon source of the system and feeds directly into glycolysis, which is a branched pathway feeding to many areas of the metabolism. It appears that reactions that are close to these branching areas are the ones that show a high degree of control over the rest of the network. The glycolysis network is well understood, and therefore substituting in real network dynamics for these reactions should be plausible. Substitution of reaction would require careful scaling of the substituted reaction, such

that the model steady state was not violated.

ATP balancing reactions also show a high degree of control over all reactions. ATP is a co-metabolite for many reactions, and has to be maintained in a desirable ratio with ADP and AMP in order for the cell to remain alive. It is likely that these could be vital in ensuring a network behaves correctly given the significance of ATP within the metabolism and for the viability of the cell.

Table 4.6: Reactions with high total flux control

Reaction	Squared total scaled flux control coefficient
phosphate transport	$1.00 \times 10^{+4}$
ammonia transport	$5.01 \times 10^{+3}$
sulfate uniport	$4.55 \times 10^{+3}$
glucose transport	$3.72 \times 10^{+3}$
glyceraldehyde-3-phosphate dehydrogenase	$3.24 \times 10^{+3}$
sulfate adenyltransferase ADP	$2.34 \times 10^{+3}$
CO2 transport	$2.22 \times 10^{+3}$
phosphofructokinase	$1.86 \times 10^{+3}$
bicarbonate formation	$1.53 \times 10^{+3}$
transketolase 2	$1.44 \times 10^{+3}$
ATP synthase	$9.31 \times 10^{+2}$
pyruvate decarboxylase	$7.80 \times 10^{+2}$
pyruvate carboxylase	$5.92 \times 10^{+2}$
glutamine synthetase	$5.53 \times 10^{+2}$
transaldolase	$5.14 \times 10^{+2}$

## 4.4 Conclusions

This methodology provides a unique way of seamlessly integrating a large proportion of known data on the yeast metabolic network. The output is a large, steady state network, on which analysis such as MCA can be performed. From this information, important control reactions within the network can be identified, and more detailed laboratory study can be undertaken on the kinetic behaviour of the reaction. This leads to an iterative process of knowledge generation. This methodology is scalable to increased

data availability and also applicable to any organism where a network reconstruction is available.

NEW INSIGHTS FROM OLD DATA: EMERGENT  
PROPERTIES OF THE KINETICALLY ENHANCED  
CONSENSUS YEAST NETWORK BUILT USING  
NON-HOMOGENOUS DATA SOURCES

## **5.1 Introduction**

The previous chapters have demonstrated the development, and validation of methods to generate a large-scale metabolic model. This chapter will focus on the application of these methods to develop a model that is specific to yeast grown in a continuous culture environment. The final aim of the chapter is to have a condition-specific yeast model that accurately represents the steady state growth conditions of the culture, but is also able to transiently reproduce system perturbations. Being the first model of this scale, it will greatly enhance the ability to understand dynamic cellular responses in a larger number of metabolic pathways.

Systems biology focusses on two primary modelling strategies, ‘top-down’ and ‘bottom-up’. In a top-down strategy the models provide a broad overview of the system components that are being analysed, and they usually encompass larger subsections of the network under study. These networks can involve Boolean algebra, stoichiometric equations, or crude models of cells that often contain lumped reactions. Top-down models

tend to contain a larger number of components, but with less biochemical and biophysical detail. From these models an overall understanding of biological behaviour can be gained.

Bottom-up modelling is more detailed in its approach and involves combining detailed biochemical and biophysical data, starting with a small number of reactions that cover a specific area of molecular interactions. A number of examples of these models have been detailed throughout the thesis, including the Erythrocyte model of Holzhütter [1] used in Chapter 3, as well as several glycolysis models [10, 11, 12, 13, 14, 15, 16, 17, 18, 19], and glycerol synthesis model [20] used in Chapter 4.

The data-centric approach used for bottom-up modelling results in models that display accurate prediction of behaviours, these behaviours include steady states, dynamic fluxes, network control, and stimulus response. The scope for generalisation (using a model, defined under one set of conditions, to predict behaviour from a different set of conditions [6]) is high. This is the primary reason why the approach is popular, despite the large number of man hours required to collect the data for a relatively small network. The average size of a glycolysis model is around 19 reactions, whilst the ‘yeast 4’ stoichiometric model [80] contains 1102 unique reactions, leaving these models with a network coverage of less than 2%.

Conversely, the sparse data approach used for top-down modelling restricts the predictions that can be made with the model. Only steady state, or pseudo-dynamic (i.e. steady-state data taken for a range of conditions) data can be calculated. Using FBA as an example, this data is based only on a set of initial conditions and an objective function. These analyses have provided valuable information of potential pathway fluxes and gene knockout behaviour [115, 50, 52], but lack the dynamic behaviour that is provided by the bottom-up approach. The approaches are illustrated in Figure 5.1

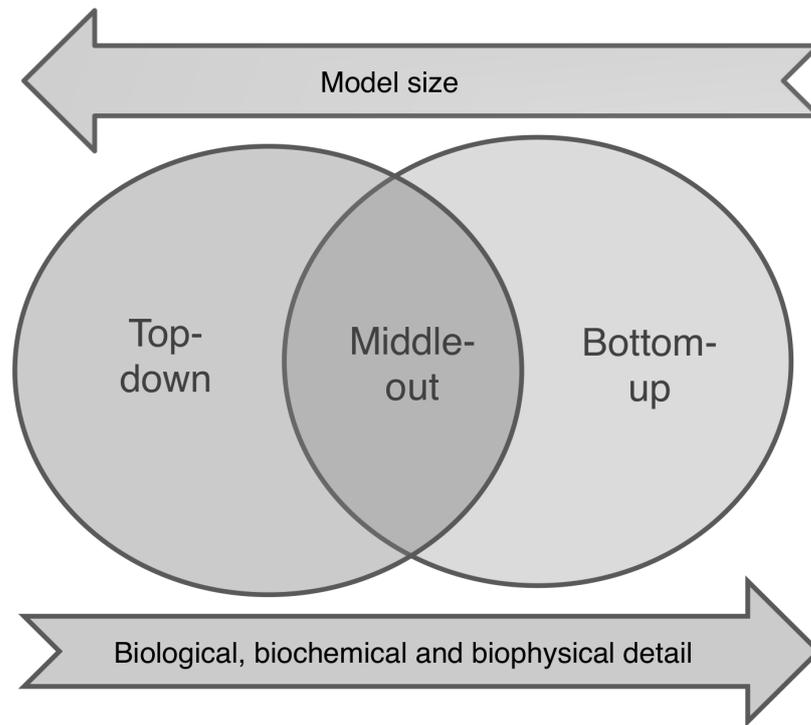


Figure 5.1: Models are typically grouped into top-down or bottom-up approaches with the level of biological, biochemical and biophysical detail contained in the model increasing as the model size decreases.

The methods generated so far in the thesis combine both approaches, in order to harnesses the simplicity of the top-down approach, but expand it such that dynamic analysis of the network is possible. The application of this approach to data collected from a specific yeast strain, grown under specific continuous culture conditions will provide more insight into the methodology. The model will be built using one steady state set of data, which corresponds to the steady state of the the yeast culture in the experimental conditions. Extensions to the previous methodology will be made in the form of stimulus response experiments. These will involve changes in extracellular glucose concentration, and will test the models ability to extrapolate to new growth conditions, something that is vital for models to be used for *in silico* hypotheses generation, which can be tested and validated in the lab. These feed into the cycle of knowledge [116] and iteratively build understanding of cellular behaviour.

This will result in a medium sized model that will provide insight into

- The reactions within the metabolic network which are active under the growth conditions.

- A more holistic view of which reactions have the most control over the system.
- How perturbations affect flux behaviour of reactions and concentrations of metabolites. This information can be used for more targeted metabolome analysis.

Two models have been combined to construct a larger model, these are glycolysis and trehalose synthesis. They were created as separate models, though they share some reactions, so they intersect in their behaviour. The first stage of generating the larger-scale model involves combining the two original models. The steady state fluxes will then be computed, along with metabolite concentrations and equilibrium constants (where not explicit). These data will be used to build the large-scale model.

## 5.2 Method

### 5.2.1 Combining glycolysis and trehalose

The glycolysis model contains 34 reactions 59 metabolites. A number of these reactions convert the same substrates and products using an isoenzyme. The trehalose model contains 8 reactions and 16 metabolites. Three of the reactions are shared (glucose transport, hexokinase and phosphoglucose isomerase), as are eight of the metabolites (ADP, ATP, fructose 6-phosphate, glucose, glucose 6-phosphate, H<sup>+</sup>, water and phosphate), and are therefore referred to as cross-over reactions and metabolites respectively. Glycolysis is the most comprehensive of the two models, it was built using data from metabolomic and proteomic quantification carried out at the Manchester Centre for Integrative Systems Biology (MCISB). It also has exact kinetic measurements for all individual isozymes within the glycolysis network. As such the glycolysis was taken to be the primary network, meaning that all information should be retained as close as possible to the original network behaviour when plugging in trehalose synthesis. The five non-crossover reactions of trehalose synthesis, and the eight non-crossover metabolites were added to the glycolysis model, including their original kinetic functions. The metabolite concentrations and  $V_{max}$  in the trehalose model were re-scaled to produce flux in mMsec<sup>-1</sup> units, like the glycolysis model. This ensured consistency between

the two models. The concentrations of the crossover metabolites were taken from the glycolysis model. ATP and ADP were among the crossover concentrations, meaning that the original ATP and ADP concentrations in the trehalose model were changed. The concentrations of the metabolites UDP and UTP in the trehalose model were set as ratios of ATP and ADP concentrations, where the ADP:UDP concentration ratio is 8.97:1, and the ATP:UTP concentration ratio is 3.89:1 [117]. The concentrations of UDP and UTP were re-calculated to reflect the new ATP and ADP concentrations. The steady state concentrations and fluxes of the new combined model can be seen in Tables 5.1, 5.2 and 5.3. The fluxes for reactions with isozymes are lumped together so a single total flux is obtained for the generic reaction. These will be used to constrain FBA and for metabolite concentrations for the larger-scale model.

Table 5.1: Intracellular metabolite concentrations from the individual models and the combined model

Intracellular metabolite	Concentration mM		
	Glycolysis	Trehalose	Combined
acetaldehyde	1.16	-	1.16
ADP	1.10	-	1.10
AMP	$1.19 \times 10^{-1}$	-	$1.19 \times 10^{-1}$
ATP	4.58	-	4.58
beta-D-glucose 6-phosphate	3.33	-	3.33
D-fructose 1,6-bisphosphate	1.68	-	1.68
D-fructose 6-phosphate	$9.08 \times 10^{-1}$	-	$9.08 \times 10^{-1}$
D-glucose	$7.69 \times 10^{-1}$	-	$7.69 \times 10^{-1}$
ethanol	$2.08 \times 10^{+2}$	-	$2.08 \times 10^{+2}$
glyceraldehyde 3-phosphate	$5.21 \times 10^{-2}$	-	$5.21 \times 10^{-2}$
glycerol	$2.50 \times 10^{+1}$	-	$2.05 \times 10^{+1}$
NAD(+)	1.42	-	1.42
NADH	$8.67 \times 10^{-2}$	-	$8.67 \times 10^{-2}$
phosphoenolpyruvate	$6.46 \times 10^{-2}$	-	$6.46 \times 10^{-2}$
pyruvate	2.65	-	2.65
sn-glycerol 3-phosphate	$1.43 \times 10^{-2}$	-	$1.43 \times 10^{-2}$
glycerone phosphate	1.24	-	1.24
uracil	$8.02 \times 10^{-1}$	-	$8.02 \times 10^{-1}$
L-leucine	1.03	-	1.03
L-histidine	$6.38 \times 10^{-1}$	-	$6.38 \times 10^{-1}$
3-phospho-D-glyceric acid	$1.22 \times 10^{-1}$	-	$1.22 \times 10^{-1}$
2-phospho-D-glyceric acid	$2.00 \times 10^{-2}$	-	$2.00 \times 10^{-2}$
3-phospho-D-glyceroyl dihydrogen phosphate	$1.19 \times 10^{-3}$	-	$1.19 \times 10^{-3}$
alpha,alpha-trehalose	-	$5.00 \times 10^{-1}$	$5.00 \times 10^{-1}$
alpha,alpha-trehalose 6-phosphate	-	$2.00 \times 10^{-1}$	$2.00 \times 10^{-1}$
UDP-D-glucose	-	$7.00 \times 10^{-1}$	$7.00 \times 10^{-1}$
UDP	-	$4.91 \times 10^{-1}$	$4.91 \times 10^{-1}$
UTP	-	1.13	1.13
D-glucose 1-phosphate	-	$1.0 \times 10^{-1}$	$1.0 \times 10^{-1}$
Median value			$7.69 \times 10^{-1}$

Table 5.2: Extracellular metabolite concentrations taken from the individual models and quantified from the extracellular media

Extracellular metabolite	Concentration mM		
	Glycolysis	Trehalose	Media
D-glucose	$7.40 \times 10^{+1}$	$7.4 \times 10^{+1}$	-
ethanol	$5.20 \times 10^{+1}$	-	-
glycerol	$1.75 \times 10^{+1}$	-	-
L-histidine	-	-	$7.09 \times 10^{-1}$
L-leucine	-	-	1.14
uracil	-	-	$8.92 \times 10^{-1}$
Median value			8.82

Table 5.3: Steady state fluxes taken from the individual models and the combined model

Reaction	Flux (mMsec <sup>-1</sup> )		
	Glycolysis	Trehalose	Combined
adenylate kinase	0	-	0
alcohol dehydrogenase	-1.6	-	-1.6
ATPase, cytosolic	3.1	-	3.1
fructose-bisphosphate aldolase	$8.4 \times 10^{-1}$	-	$8.4 \times 10^{-1}$
glucose-6-phosphate isomerase	$8.4 \times 10^{-1}$	1.6	$8.4 \times 10^{-1}$
glyceraldehyde-3-phosphate dehydrogenase	1.7	-	1.7
glycerol-3-phosphatase	$4.6 \times 10^{-4}$	-	$4.6 \times 10^{-4}$
glycerol-3-phosphate dehydrogenase (NAD)	$4.6 \times 10^{-4}$	-	$4.6 \times 10^{-4}$
hexokinase (D-glucose:ATP)	1.0	1.7	1.1
phosphofructokinase	$8.4 \times 10^{-1}$	-	$8.4 \times 10^{-1}$
phosphoglycerate kinase	-1.7	-	-1.8
phosphoglycerate mutase	-1.7	-	-1.7
pyruvate decarboxylase	1.6	-	1.6
pyruvate kinase	1.7	-	1.7
triose-phosphate isomerase	$8.4 \times 10^{-1}$	-	$8.0 \times 10^{-1}$
ethanol transport	-1.6	-	-1.6
glucose transport	$9.4 \times 10^{-1}$	1.6	1
glycerol transport via channel	$4.6 \times 10^{-4}$	-	$4.6 \times 10^{-4}$
phosphoglucomutase	-	$-5.8 \times 10^{-2}$	$-9 \times 10^{-4}$
UTP-glucose-1-phosphate uridylyltransferase	-	$5.8 \times 10^{-2}$	$9 \times 10^{-4}$
alpha, alpha-trehalose-phosphate synthase (UDP-forming)	-	$5.8 \times 10^{-2}$	$9 \times 10^{-4}$
trehalose-phosphatase	-	$5.8 \times 10^{-2}$	$9 \times 10^{-4}$
alpha, alpha-trehalase	-	$5.8 \times 10^{-2}$	$9 \times 10^{-4}$

### 5.2.2 Altering the network

The yeast 4 metabolic reconstruction [80] was used as the starting network with which the flux through the entire metabolic network would be calculated. The model is a generic representation of reactions found in the *S. cerevisiae*, based on genome annotation and literature. The strain of yeast that was cultured for measuring the metabolite concentrations and kinetics (BY4743 ho/HO,[YDL227C; MATa/ $\alpha$ MAT; his3

1/his3 1;leu2 0/leu2 0; met15 0/MET15; LYS/lys2 0; ura3 0/ura3 0]) was an auxotroph strain requiring the addition of histidine, uracil, and leucine in the culture media. In order to produce a network representation close to the strain, reactions associated with these metabolites were altered. The uptake reactions of each amino acid was then altered so that only net import of these metabolites was allowed, to ensure the cellular fluxes would be viable for cellular growth. These flux bound changes are summarised in table 5.4.

Table 5.4: Reactions in Yeast 4 where flux bound conditions were altered to make the network represent the yeast strain used to collect the data

Reaction	Original Bound		New Bound	
	lb	ub	lb	ub
L-histidine transport	$-\infty$	0	0	$\infty$
imidazoleglycerol-phosphate dehydratase	0	$\infty$	0	0
uracil transport	0	$\infty$	0	$\infty$
orotidine-5'-phosphate decarboxylase	0	$\infty$	0	0
L-leucine transport	$-\infty$	0	0	$\infty$
3-isopropylmalate dehydratase	$-\infty$	$\infty$	0	0

### Flux Balance Analysis (FBA)

The FBA technique used was the same as outlined in Chapter 4, Section 4.2, using the fluxes outlined in Table 5.3. These fluxes were calculated from in-house models of the glycolysis and trehalose pathways, which were built using *in vitro* kinetic rate laws, and a mixture of intracellular metabolomics, proteomics and exo-metabolome data. The steady state data from the models was used, which differed slightly from the raw data, to ensure that the data being used was most compatible for the steady state kinetic analysis. Where a good fit to the exact fluxes was not feasible the bounds were systematically expanded until a suitable flux distribution could be found.

The flux solution computed was taken, and all reactions displaying a flux of 0 were removed from the network. The FBA was then re-run to ensure that the solution of fluxes was still consistent with the new smaller network. The 284 reactions that remained in the network, displaying a finite flux can be found in Appendix B. Table 5.9 shows the KEGG [111, 112, 113] reaction pathways that were present in the reduced model, identifying the

likely reaction pathways that are important for the standard growth for this strain of yeast, in these conditions.

Reactions with negative fluxes were reversed for the convenience of having a positive flux for all reactions. This is not strictly necessary but simplifies the computation required in further stages of the methodology.

### Metabolite concentrations

The steady state concentrations used in the model correspond to those that were obtained for the glycolysis and trehalose models using quantitative metabolomics. Others that were quantified for studies under similar conditions were also added to the network. These intracellular concentrations can be seen in Table 5.1, the extracellular concentrations can be seen in Table 5.2. For metabolites with no associated quantitative data, the median value of all known intracellular concentrations was applied as the default concentration. For the extracellular concentrations the median value of all known extracellular values were computed and applied as the default.

A number of metabolites are known to show an extremely low extracellular concentration, these are outlined in Table 5.5. These metabolites were assigned as 0.

Table 5.5: Modified metabolites known to have a extremely low extracellular concentrations

Extracellular metabolite	Concentration mM
alpha-D-glucosamine 6-phosphate	0
carbon dioxide	0

### Balancing parameters

The same parameter balancing methodology as that outlined in Chapter 4 Section 4.2 was used. The measured equilibrium constants were taken from the in-house glycolysis and trehalose models, the values can be found in Table 5.6.

Where solutions were not found for the linearly dependent reactions, the lower or upper  $K_{eq}$  bound were altered accordingly until a suitable solution to the system was found. For reactions that required a lower bound such that the  $K_{eq} >$  mass action ratio no longer held true, the metabolite concentrations involved in the reaction were altered.

Only metabolites that had no associated concentration data (and, as such, had assigned values equal to the median of known values) were changed.

The original models do not include phosphate, diphosphate, carbon dioxide or bicarbonate directly in the rate laws. For this reason, these concentrations were not used when calculating the mass action ratios of each reaction.

Table 5.6: *in vitro* measured equilibrium constants taken from the in house glycolysis model

Reaction	Equilibrium constant		
	Glycolysis	Trehalose	Combined
alcohol dehydrogenase (ethanol)	14000	-	14000
fructose-bisphosphate aldolase	0.069	-	0.069
glucose-6-phosphate isomerase	0.29	-	0.29
glyceraldehyde-3-phosphate dehydrogenase	0.006	-	0.006
glycerol-3-phosphatase	10000	-	10000
hexokinase (D-glucose:ATP)	2000	2000	2000
phosphofructokinase	800	-	800
phosphoglycerate kinase	3200	-	3200
phosphoglycerate mutase	0.19	-	0.19
pyruvate kinase	6500	-	6500
triose-phosphate isomerase	0.045	-	0.045
ethanol transport	1	-	1
glycerol transport via channel	1	-	1
phosphoglucomutase	-	0.17	0.17
UTP-glucose-1-phosphate uridylyltransferase	-	$\infty$	$\infty$
alpha,alpha-trehalose-phosphate synthase (UDP-forming)	-	$\infty$	$\infty$
trehalose-phosphatase	-	$\infty$	$\infty$
alpha,alpha-trehalase	-	$\infty$	$\infty$

### Generating the kinetic model

The convenience rate law was used, with the same scaling methodology outlined in Chapter 4 Section 4.2 and 4.2. Phosphate, diphosphate, carbon dioxide and bicarbonate were all excluded from all rate laws, and assigned as constant concentrations within the model.

### Metabolic control analysis

The same method as that outlined in Chapter 4 Section 4.2.1 was used. The results from the MCA were used to identify which reactions from the original model shows the highest flux control over the system. Once identified the corresponding equation from the original model was substituted into the large model, simulating the measurement of that enzyme's kinetics.

To see how much improvement was brought about by adding in accurate kinetics, two perturbations were made to the new model: a reduction to 50mM extracellular glucose, and then an increase to 100mM extracellular glucose from the operating concentration of 74mM. The transient concentrations of selected metabolites (present in the original model) were tested against the transient behaviour from the same perturbations in the original model. Further iterations of the process were repeated until a reasonable approximation of the original system was achieved. This iterative method of identification of a reaction with high flux control, followed by 'measurement' of the kinetics of the associated enzyme, suggest a model-based strategy towards building large-scale kinetic models.

## 5.3 Results and discussion

The model produced contains 284 reactions and 293 metabolites. Full details of these can be found in Appendix B. Three extracellular metabolite concentrations were adjusted to balance the influx and efflux to that observed in the FBA. These metabolites concentrations are outlined in Table 5.7 along with the new associated concentrations. In order to improve the transient behaviour of the large model, ATP, ADP and AMP metabolites were assigned as fixed concentrations. This was important for transient behaviour because of their high involvement in numerous reactions, small perturbations in the model can cause their concentrations to alter rapidly, which perpetuates instability through the model.

The rate laws for lipid production, biomass and growth were assigned according to

$$K \cdot \text{D-glucose [extracellular]} \quad (5.3.1)$$

where D–glucose [extracellular] is a modifier to the reaction and corresponds to the fixed extracellular glucose input. The parameter  $K$  is scaled such that the flux through each reaction matches its corresponding FBA flux at the original concentration of extracellular glucose. In Chapter 4, these were assigned as constant flux. The complex nature of the biomass equation causes issues with extrapolation to a convenience kinetics rate law. The assignment of constant flux works well for a single steady state, but causes instabilities within the system for changes in flux. The specific growth rate of the cell should be related to the amount of glucose available in the media through sensing, by using the mass action rate law, this ensured that the model could alter growth related to changes in the extracellular glucose.

Table 5.7: Extracellular concentrations that were modified to maintain correct reaction directionality for export reactions with an equilibrium value of 1

Extracellular metabolite	new concentration (mM)	original concentration (mM)
D-glucitol	0.7	8.82
(S)-malate(2-)	0.7	8.82
succinate(2-)	0.7	8.82

### Steady state behaviour

There were some differences between the fluxes computed from the original combined model of glycolysis and trehalose, and those computed from the FBA of the yeast 4 network, detailed in Table 5.8. The flux boundaries set for the glycolysis specific reactions in the FBA (taken from the combined model) were kept close to the original values for the fitting, by only allowing the values to vary between +/- 20% of the steady state value computed from the original combined model. This follows the aims outlined at the start, where the biochemistry of the glycolysis model (believed to be the most biochemically correct) was given highest priority for close reproduction of the flux values. When similar bounds were used for the trehalose specific reactions, from the combined model, the FBA

result from the yeast 4 network did not fulfil the objective function of biomass production. This shows that the flux values were not viable when applied to the yeast 4 model. A solution was found when the flux bounds for the trehalose specific reactions (taken from the combined model) were allowed to vary between +/- 100%. The FBA flux values, computed for yeast 4, for the trehalose specific reactions, show that UTP-glucose-1-phosphate uridylyltransferase and alpha,alpha-trehalase demonstrated flux that was very different from their flux values in the combined glycolysis and trehalose model. For UTP-glucose-1-phosphate uridylyltransferase, the FBA flux brought the flux value closer to the original flux value from the original trehalose model before it was combined with glycolysis (see Table 5.3). The FBA flux for alpha,alpha-trehalase was extremely low. The differences in these flux values suggest that these reactions are involved in pathways that are not taken into account in the original combined model of glycolysis and trehalose, causing a break in the 'loop' to and from glycolysis. This also suggests that combining the models before taking the steady state is not necessary. The separate steady state fluxes could be taken, and prioritised with flux bounds to ensure the more biochemically correct data is maintained within stricter bounds. This approach is advocated for future implementations of the methodology.

Table 5.8: Flux results for the known fluxes

Reaction	Flux (mM sec <sup>-1</sup> )	
	Original	FBA
adenylate kinase	0.0	0
alcohol dehydrogenase	-1.6	-1.3
ATPase, cytosolic	3.1	2.5
fructose-bisphosphate aldolase	$8.4 \times 10^{-1}$	$6.8 \times 10^{-1}$
glucose-6-phosphate isomerase	$8.4 \times 10^{-1}$	$6.7 \times 10^{-1}$
glyceraldehyde-3-phosphate dehydrogenase	1.7	1.4
glycerol-3-phosphatase	$4.6 \times 10^{-4}$	$5.2 \times 10^{-4}$
glycerol-3-phosphate dehydrogenase (NAD)	$4.6 \times 10^{-4}$	$5.2 \times 10^{-4}$
hexokinase (D-glucose:ATP)	$9.9 \times 10^{-1}$	$7.9 \times 10^{-1}$
phosphofructokinase	$8.4 \times 10^{-1}$	$6.8 \times 10^{-1}$
phosphoglycerate kinase	-1.7	-1.4
phosphoglycerate mutase	-1.7	-1.4
pyruvate decarboxylase	1.6	1.3
pyruvate kinase	1.7	1.4
triose-phosphate isomerase	$8.4 \times 10^{-1}$	$6.8 \times 10^{-1}$
ethanol transport	-1.6	-1.3
glucose transport	1.0	$8.0 \times 10^{-1}$
glycerol transport via channel	$4.6 \times 10^{-4}$	$5.2 \times 10^{-4}$
phosphoglucomutase	$-9.0 \times 10^{-4}$	$-9.0 \times 10^{-3}$
UTP-glucose-1-phosphate uridylyltransferase	$9.0 \times 10^{-4}$	$9.0 \times 10^{-3}$
alpha,alpha-trehalose-phosphate synthase (UDP-forming)	$9.0 \times 10^{-4}$	$1.3 \times 10^{-4}$
trehalose-phosphatase	$9.0 \times 10^{-4}$	$1.3 \times 10^{-4}$
alpha,alpha-trehalase	$9.0 \times 10^{-4}$	$9.0 \times 10^{-6}$

Table 5.9: Kegg pathways represented in the model

Pathway	Number reactions
Purine metabolism	28
Lipid Metabolism	22
Glycolysis / Gluconeogenesis	21
Biosynthesis of secondary metabolites	19
Transport	17
Glycine, serine and threonine metabolism	13
Steroid biosynthesis	13
Arginine and proline metabolism	12
Pyrimidine metabolism	12
TCA cycle	12
Alanine, aspartate and glutamate metabolism	11
Phenylalanine, tyrosine and tryptophan biosynthesis	10
Valine, leucine and isoleucine biosynthesis	9
Riboflavin metabolism	7
Cysteine and methionine metabolism	6
Fatty acid biosynthesis	6
Pentose phosphate pathway	6
Starch and sucrose metabolism	5
Fructose and mannose metabolism	4
Glycerophospholipid metabolism	4
Lysine biosynthesis	4
Oxidative phosphorylation	4
Pyruvate metabolism	4
Glycerolipid metabolism	3
Inositol phosphate metabolism	3
Sulfur metabolism	3
Misc	26

The pathways contained in the model, and the number of reactions in each pathway, are outlined in Table 5.9. They follow similar trend to those in Chapter 4, where purine metabolism, lipid metabolism and glycolysis are among the most populated pathways.

The data within the model is generated from yeast cells grown in a continuous culture environment, known as a turbidostat. The cells are grown under non-nutrient

limited conditions and at the maximum growth rate. The concentration of cells in the culture is kept constant by a biomass probe which controls the addition of media into the vessel and waste is removed from the vessel at the same rate to retain a constant culture volume. The primary carbon source is glucose, and *S.cerevisiae* is known to be Crabtree positive, meaning it prefers to ferment, even under excess oxygen, therefore the majority of pyruvate produced will enter the fermentative route, to produce ethanol. Diagram 5.2 (with the full metabolite identifiers in Table 5.10) shows the main pathways and associated flux through the central carbon metabolism in the final model, including glycolysis, trehalose, Pentose Phosphate Pathway (PPP) and Tricarboxylic Acid Cycle (TCA). The primary flux through the models correlates with the growth conditions of the original yeast culture, with a large amount of the flux being directed from D-glucose to ethanol production, expected behaviour.

The model demonstrates a flux through PPP of <1%, the literature suggests that between 4% [118] and 47% of flux has been seen to be diverted through the PPP. The low flux generated in the model could be the result of the growth conditions of the culture, or it could be a facet of the data used to constrain the model. The flux through the combined model of glycolysis and trehalose heavily constrains the FBA to divert flux primarily through glycolysis, with deviations through trehalose.

The flux through the TCA has been noted to be very small in fermenting yeast cultures [118], although the culture modelled in this case had plenty of oxygen. It was also noted in the same paper that the flux through the TCA showed a reverse correlation with the cellular growth, although this was only a weak correlation. The yeast culture the model was based on was grown in a turbidostat so was not exposed to glucose limitation, as such they were allowed to attain maximum growth. Based on this, it would suggest that the flux through TCA should perhaps be lower than computed from the FBA. However, the cultures are making a large amount of biomass, given they are under maximum growth conditions. Much of the biomass production comes from reactions within the TCA, and this would suggest that the higher than expected flux through the TCA reactions are reflecting this.

There are some methods that could be used to analyse this flux behaviour further, such

as static and dynamic flux analysis, both of which used labelled glucose to accurately identify the flux patterns more accurately.

The fluxes through the TCA show that the pathway is not operating as a cycle, and it is using the glyoxylate shunt. The glyoxylate shunt is an variation on the TCA cycle that converts acetyl-CoA to succinate.

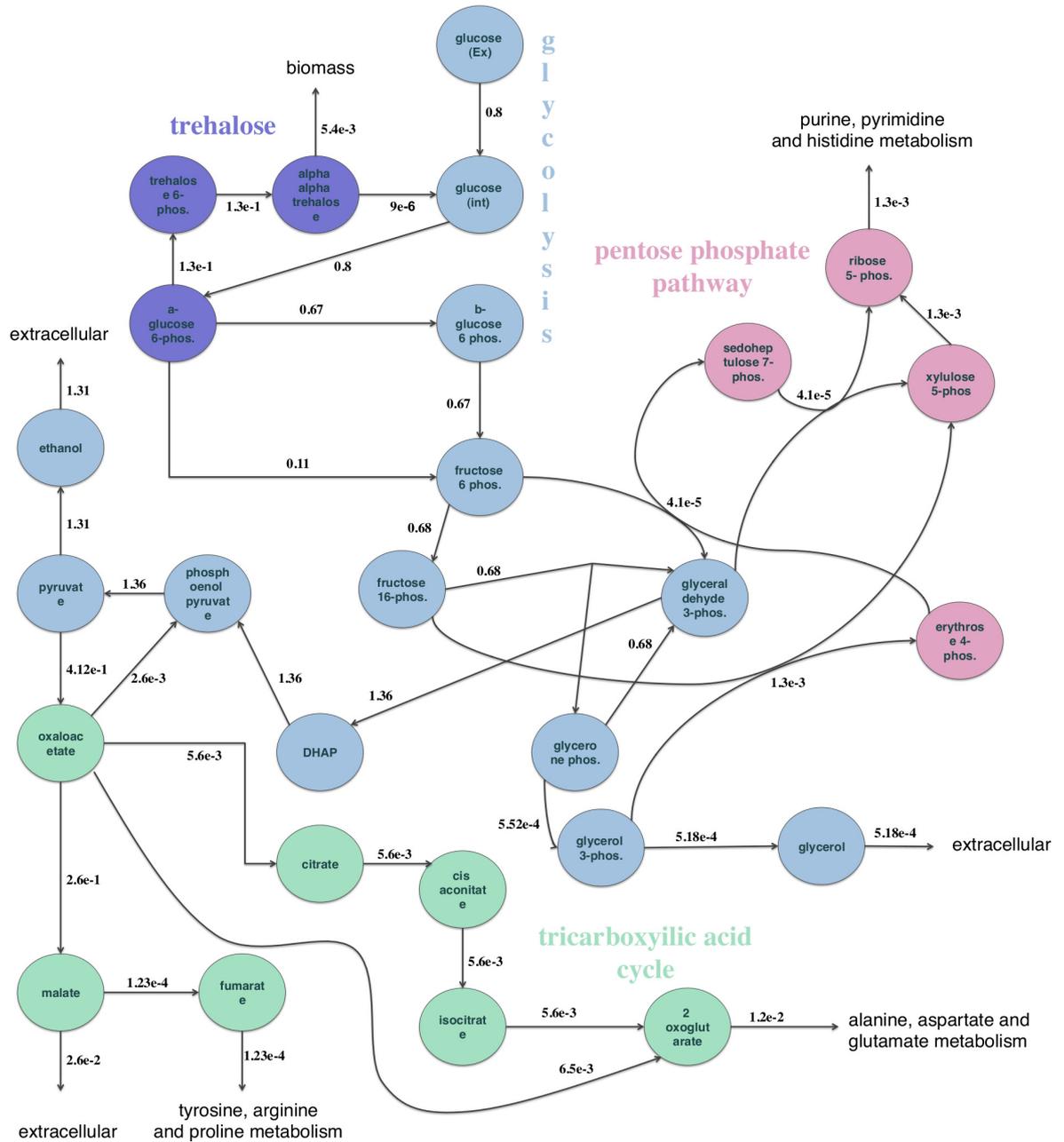


Figure 5.2: Flux distribution through central carbon metabolism. Full metabolite names are available in Table 5.10

Table 5.10: Metabolite names used in Figure 5.2

Identity in diagram	Metabolite name
glucose (ex)	D-glucose [extracellular]
glucose (int)	D-glucose [intracellular]
b-glucose 6 phos.	beta-D-glucose 6-phosphate [intracellular]
a-glucose 6 phos.	aldehydo-D-glucose 6-phosphate [intracellular]
fructose 16-phos.	D-fructose 1,6-bisphosphate [intracellular]
glyceraldehyde 3-phos.	glyceraldehyde 3-phosphate [intracellular]
glycerone phos.	glycerone phosphate [intracellular]
glycerol 3-phos.	sn-glycerol 3-phosphate [intracellular]
glycerol	glycerol [intracellular]
DHAP	3-phospho-D-glyceroyl dihydrogen phosphate [intracellular]
phosphoenolpyruvate	phosphoenolpyruvate [intracellular]
pyruvate	pyruvate [intracellular]
ethanol	ethanol [intracellular]
fructose 6-phos.	D-fructose 6-phosphate [intracellular]
trehalose 6-phos.	alpha,alpha-trehalose 6-phosphate [intracellular]
alpha alpha trehalose	alpha,alpha-trehalose [intracellular]
oxaloacetate	oxaloacetate(2-) [intracellular]
malate	(S)-malate(2-) [extracellular]
fumarate	fumarate(2-) [intracellular]
citrate	citrate(3-) [intracellular]
cis aconitate	citrate(3-) [intracellular]
isocitrate	isocitrate(3-) [intracellular]
2 oxoglutarate	2-oxoglutarate [intracellular]
erythrose 4-phos.	D-erythrose 4-phosphate(2-) [intracellular]
xylulose 5-phos	D-xylulose 5-phosphate [intracellular]
sedoheptulose 7-phos.	sedoheptulose 7-phosphate [intracellular]
ribose 5-phos.	alpha-D-ribose 5-phosphate [intracellular]

### Substituting reaction kinetics

The analysis demonstrates 12 of the reactions in the pathway were shown to have high control over the network (defined here to be reactions that show a flux control coefficient of >95% of the highest flux control value), these can be seen in Table 5.11. Of these

reaction, original kinetics are known for 6. The three known reactions with the highest control are hexokinase (D-glucose:ATP), glucose transport, phosphofructokinase. These three reactions have also been identified as having high flux control in the glycolysis pathway used in Chapter 2. Although it was glucose transport that showed up as having the highest flux control predominantly.

Compared to the ‘generic’ yeast model, built in Chapter 4 (see Table 4.6 for the ‘generic’ model flux control coefficients), the flux control in this network seems to be less strongly associated with transport reactions and does not include any ATP balancing reactions (although the lack of ATP balancing reactions is accounted for by the these metabolites being fixed within the model). It is possible that this is a result of the network being restricted to growth on a minimal media.

Table 5.11: Reactions with high total flux control

Reaction	Root squared total scaled flux control coefficient
hexokinase (D-glucose:ATP)	540
D-sorbitol reductase	380
glucose transport	342
phosphofructokinase	265
glyceraldehyde-3-phosphate dehydrogenase	262
transketolase 2	251
glucose-6-phosphate isomerase	198
phosphoglycerate kinase	193
citrate synthase	186
aspartate transaminase	184
glutamine synthetase	157
ribulose 5-phosphate 3-epimerase	138

Using the flux control data, coupled with the known reaction kinetics from the original model, the hexokinase (D-glucose:ATP) was replaced first and its transient behaviour was tested using a perturbation below (50mM) and above (100mM) the operating concentration of extracellular glucose. Details of selected metabolites from the original pathway can be seen in Figures 5.3 to 5.6. The ability of the models to predict the transient

behaviour after the first rate-law substitution was still poor, showing transient behaviour far removed from that of the original combined glycolysis and trehalose model. The model was also unable to attain a steady state after perturbations above or below the original extracellular glucose concentration. This suggests that the model is still unstable when perturbed from the steady state to which it was fitted.

In order to improve the fit of the model, a second accurate rate law substitution was made. Of the known reactions glucose transport was the second reaction showing high levels of flux control. The transient behaviour produced from the substitution of the glucose transport rate law showed a huge improvement in model behaviour. Perturbations below also show a marked improvement in being able to reproduce the transient behaviour of the metabolic pools, but still show the system may be unstable for perturbations below the operating concentration of extracellular glucose. Substitution of two *in vitro* measured rate laws means that less than 1% of the network has accurately measured kinetics, yet the reproduction of transient behaviour is already becoming very accurate. This is remarkable, and shows that the method of using models to drive which kinetics should be measured in the lab is viable for building better models, using less resources. It is possible that further rate law substitution from the original kinetics would help to improve the stability of the system.

Further improvements to the model could be made by alterations to the equations assigned for lipid metabolism, biomass production and growth. These rate laws are matched to the original system flux. When glucose input is increased or decreased, these reactions scale linearly with the extracellular glucose value. Infinitely large perturbations of extracellular glucose result in an infinitely large biomass production. In reality there would be a limit to the growth and biomass production a cell could achieve within a given time scale. More biologically feasible rate laws could be generated using dynamic FBA, where a rate law could be computed that correlates glucose input with the change in flux lipid metabolism, biomass production and growth. This should improve the transient predictability of the model.

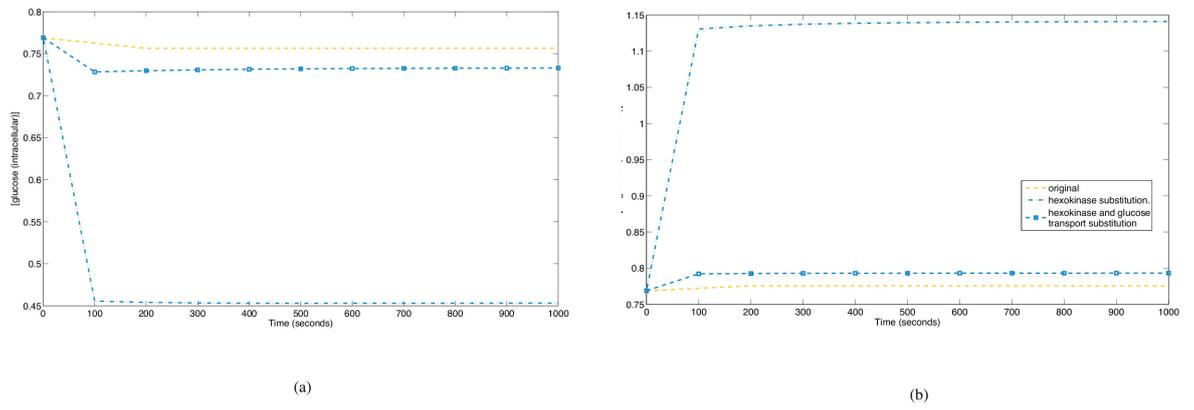


Figure 5.3: extracellular glucose perturbation of 50mM and 100mM on D-glucose. Fluxes are given in  $\text{mM sec}^{-1}$

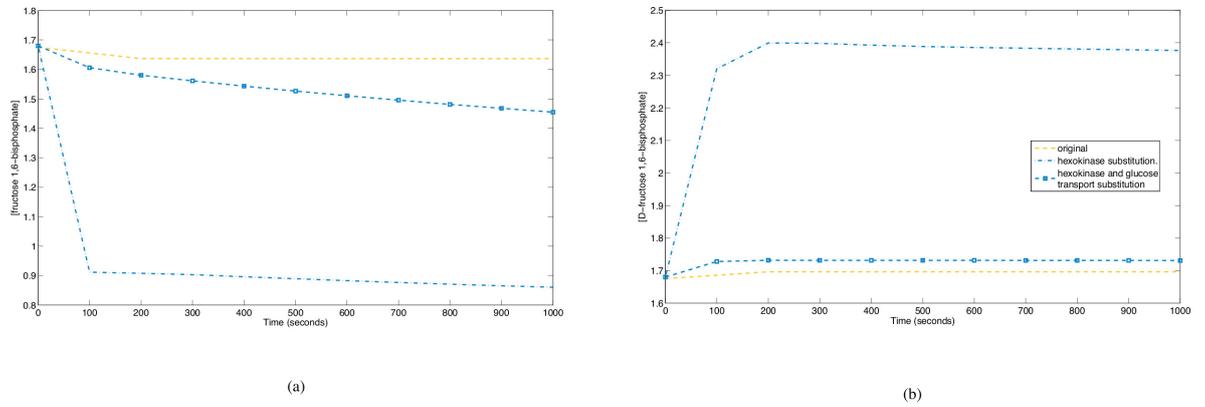


Figure 5.4: extracellular glucose perturbation of 50mM and 100mM on D-fructose 1,6-bisphosphate. Fluxes are given in  $\text{mM sec}^{-1}$

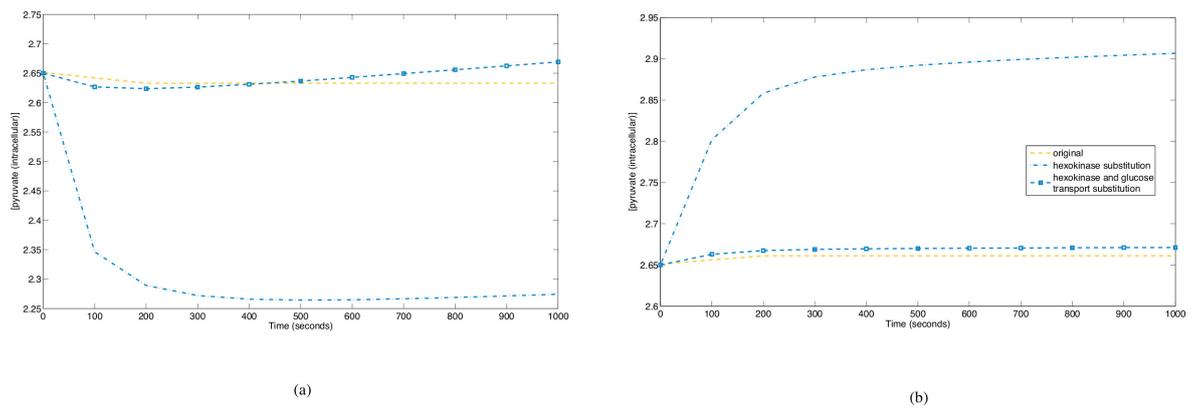


Figure 5.5: extracellular glucose perturbation of 50mM and 100mM on pyruvate. Fluxes are given in  $\text{mM sec}^{-1}$

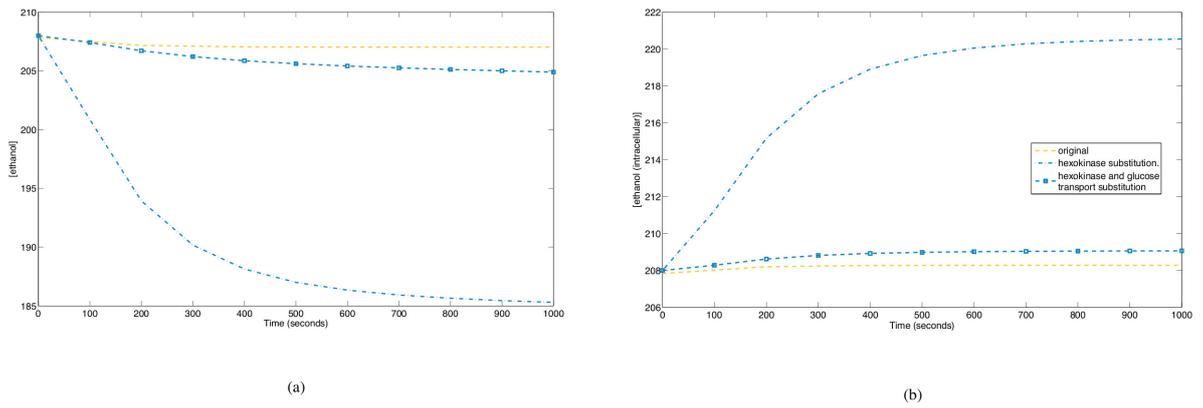


Figure 5.6: extracellular glucose perturbation of 50mM and 100mM on ethanol. Fluxes are given in  $\text{mM sec}^{-1}$

## 5.4 Conclusions

This chapter aimed to apply the methods developed in previous chapters to a specific yeast model so that specific steady state and transient behaviour of the model could be assessed and the methodology validated. Here it has been shown that the methodology shows promise at being able to integrate data easily, and reproduce behaviours of the original model to a reasonable degree. Some literature inconsistencies with the steady state flux distribution with regards to the amount of flux through the TCA and PPP were noted, and ways of addressing these should be developed for further improvement of the model.

The ability of the model to show a reasonable reproduction of transient behaviour after perturbation suggest that it could be used as a starting point for hypothesis generation, with regards to further downstream metabolic behaviour in response to perturbations. This is something that has not been possible before with models of such a size.

## DYNAMIC FLUX EXPERIMENTS OFFER FURTHER INSIGHT INTO CELLULAR BEHAVIOUR

### **6.1 Introduction**

An overarching theme of the previous chapters has been the major outputs of interest from kinetic modelling of metabolism, which is the steady state flux and concentration values and their combined phenomenological interactions. The fluxes of reactions and concentrations of metabolites within a network are independent variables [108]. At steady state it is possible to alter fluxes without alteration of concentrations, equally, changes in metabolite concentrations do not necessarily alter steady state flux through the system [108], meaning knowledge of one without the other does not determine exactly the state of the system.

Metabolic profiling and quantification are metabolomic techniques that can be used to assign phenotypic profiles of biological samples. These techniques are also useful to quantify the metabolite pools for use in bottom-up modelling techniques. Experimental flux determination has not been quite as widely used, perhaps because it is harder. Profiling of internal flux distributions through different metabolic pathways has seen the most popularity, providing information on metabolically active pathways, this is known as static Metabolic Flux Analysis (static MFA) [32, 119, 120, 121]. Quantitative flux values are more frequently calculated by combining uptake/excretion rates of extracellular

metabolites [8] with the quantitative metabolomic data and enzyme kinetic measurements [15]. It has been rare for direct flux quantification to be performed, although dynamic MFA (dynamic MFA) has been developed and used for this purpose [31].

Both static and dynamic MFA techniques allow for a large proportion of the metabolic system to be analysed simultaneously. MFA combines traditional targeted and non-targeted metabolomic techniques, using dynamic tracers, and computational data analysis to calculate the steady state fluxes of the reactions. This can be performed using a static, or a dynamic approach. Both approaches are reliant on the addition of stable isotopes, such as  $^{13}\text{C}$  [24], to the system and normalisation to natural abundance [122]. Traditional experiments have used radioactive heavy carbon ( $^{14}\text{C}$ ) [22], but for modern experimentation, this is a less popular technique because Nuclear Magnetic Resonance (NMR) and Mass Spectrometry (MS) are able to detect stable isotopes.

Static MFA has been the most popular of the two approaches [28, 26, 123], providing information on the pathway distribution of fluxes through the metabolic network. It has been used, for example, to investigate robustness of knockout mutants in yeast [121], quantifying the relative contribution of converging pathways to metabolic intermediates [26] and identifying specific metabolite production hosts [123]. The details of the technique are available from Zamboni and colleagues [124]. In brief, it requires organisms to be cultured in a continuous environment (chemostat or auxostat) until a steady state is reached, or using a pseudo steady state with batch cultures. The growth medium is then switched from a non labelled source ( $^{12}\text{C}$ ) to a 20% uniformly labelled ([U- $^{13}\text{C}$ ]) and 80%  $^{12}\text{C}$  mixture or a 100% singularly labelled at C-1 source ([1- $^{13}\text{C}$ ]) source. Both sources provide variance in the specific carbon labelling patterns of amino acids. The new carbon source is allowed to permeate through the cell until full isotopic saturation in the amino acid pools is achieved. The minimum time for this saturation to take place will be of the order of the doubling time of the cell culture. The amino acid labelling patterns are then identified using GC-MS and flux distributions quantified by calculation of possible distributions based on a reduced network of downstream metabolite carbon labelling patterns. The pathways used to make the downstream metabolite are identified based on how the individual carbon atoms are re-ordered as

they are processed by different enzymes [122]. Software such as OpenFlux [125] and FiatFlux [126] are used to aid the analysis.

Dynamic flux analysis follows a similar methodology to that outlined above: the organism is cultured to steady state in a continuous culture environment before being isotopically perturbed using a labelled carbon source. Previous experiments have used mixtures of unlabelled,  $[1-^{13}\text{C}]$  and  $[\text{U}-^{13}\text{C}]$ , with the optimal (so both dynamic and static flux experiments can be performed on the same culture) proportion believed to be 18%, 39% and 43% respectively [31]. For *E.coli*, rapid sampling on a sub-second time scale was required to ensure that the dynamic behaviour of glycolysis could be obtained. Each time-point sample must have the cellular metabolism immediately suspended in order for the *in vivo* data to be accurate, a process known as quenching. The favoured method of quenching is in cold methanol [22, 127, 71, 21]. The metabolite pools of interest are then quantified, followed by the changing distribution of carbon labelling, which is traced over the different time points. To transform these data into flux results computational analysis is then performed.

The computational method of obtaining the fluxes relies on solving the non-linear ODE system, of which the general form, taken from [31] and [128], can be seen in equation 6.1.1

$$\text{diag}(\mathbf{X}) \cdot \dot{x} = f(v, x^{\text{inp}}, x), x(0) = x_0 \quad (6.1.1)$$

where  $\mathbf{X}$  a vector of the fixed concentrations of all metabolite pools,  $x^{\text{inp}}$  is the known input label, and  $v$  is the flux through each reaction pool. The equation is solved for  $x$  which represents the transient labelling patterns present in each metabolite pool. The most complex form of these calculations involves tracing the labelling patterns of each carbon atom present in the molecule as its identity is transformed by enzymes within the system. Figure 6.1 demonstrates this increase in complexity between moving from metabolite tracing to atom tracing. Different enzymes converting reactions can cause atom exchange in different patterns, as demonstrated by reactions  $v_2$  and  $v_3$ . If these exchanges are of primary interest, then a partial labelling methodology should be used so that these changes are identifiable. A detailed development of the method can be found in the series

of papers published by Weichert and colleagues [129, 130, 131, 132]. It demonstrates how to construct matrices for atom and metabolite transitions before constraining an ODE system, based around this, to the experimental results.

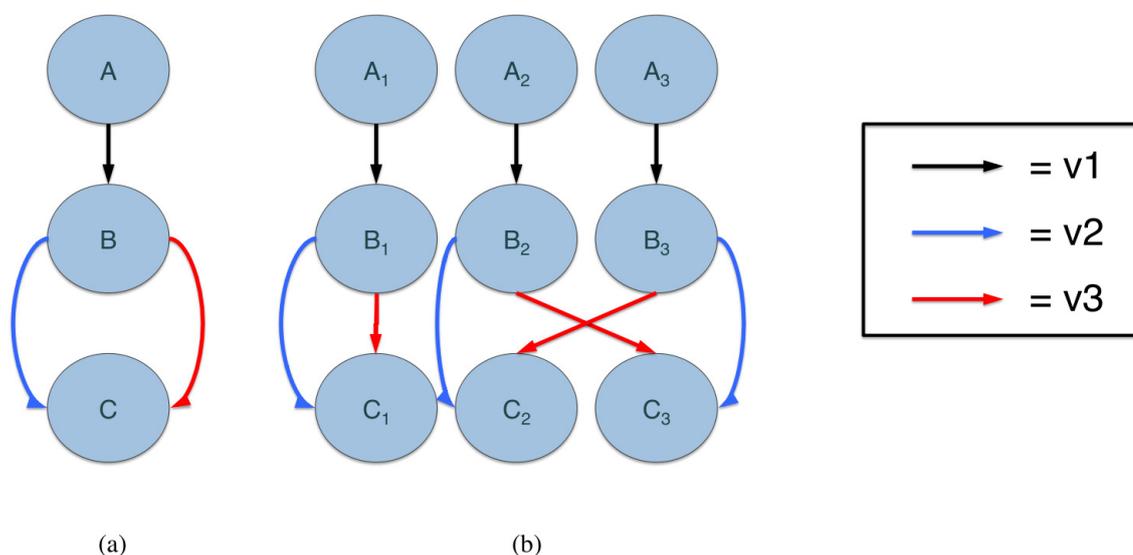


Figure 6.1: Demonstrating a flux schematic for (a) metabolite exchange and (b) Atom exchange. (a) shows the metabolite as a whole molecule, (b) separates the metabolite into the individual carbon atoms contained within the whole molecule

The kinetic modelling methodology, developed in this thesis and outlined in Chapter 4, requires knowledge of the net flux through the system. For this purpose the individual atom transfer methodology is needlessly complex. By using this as a foundation, I developed a more simplistic methodology that relied only on the metabolic exchange network. The metabolic exchange network is available from KEGG [111, 112, 113] or from metabolic reconstructions [79, 80], and is known more comprehensively for most of the metabolic network. The new method traces 100% [U-  $^{13}\text{C}$ ] through the metabolic pools of interest (in this instance, the branching points of the central carbon metabolism). Inverse modelling of the isotopic saturation patterns during the time course experiment will allow the estimation of fluxes through different groups of metabolic pools within the central carbon metabolism. The success of this experiment requires that the biochemical interactions of the metabolic exchange network to be understood in detail. The well-studied nature of central carbon metabolism (glycolysis, TCA cycle, glycerol synthesis, pentose phosphate pathway) meant that this was an ideal network on which to develop the method.

### 6.1.1 Determining sampling times

To ensure that the fast dynamics of glycolysis could be investigated, a preliminary experiment was undertaken. This used the same culturing conditions outlined in Section 6.2, but with the addition of 20% [U-  $^{13}\text{C}$ ]. The data showed that isotopic changes were observable using the maximum sampling times possible (per two minutes for up to 15 minutes, followed by a relaxation in sampling time) given the equipment limitations. To see dynamics of the glycolysis pathway, the method has typically required sampling on a sub-second time scale, however, the turbidostat conditions negate this need. The turbidostat runs in excess glucose, so full  $^{13}\text{C}$  saturation of the extracellular pool was delayed, something that would not occur in a glucose limited chemostat. This delay resulted in slowed saturation of all intracellular pools. The delay can be accounted for within the model used for computing the fluxes, therefore the method was carried forward.

## 6.2 Method

### 6.2.1 Fermentation and sampling

#### Turbidostat growth

A diploid heterozygous deletion yeast strain BY4743  $ho /HO$ , (YDL227C; MATa/ $\alpha$ MAT;  $his3 \Delta /his3 \Delta$ ;  $leu2 \Delta /leu2 \Delta$ ;  $met15 \Delta /MET15$ ;  $LYS2/lys2 \Delta$ ;  $ura3 \Delta /ura3 \Delta$ ) was grown aerobically in a fermentor (Applikon Biotechnology, Netherlands) operating under turbidostat conditions [133]. In brief, a 3L vessel was operated using a working volume of 2L. The strain of *S.cerevisiae* was grown in modified F1 Medium [134]. A growth control mechanism was established using a feedback loop in which a biomass probe measured the capacitance of the culture (Aber instruments). The feedback loop controlled the culture at 75% of the maximum biomass yield by the addition of fresh medium, and excretion of excess culture, using separate pumps, in order to maintain a constant working volume. The calibration for the biomass percentage was set by growing batch cultures to stationary phase, the equivalent of 100% biomass. The pH of the culture was maintained at 6.5 automatically using a feedback loop and automatic addition of either 1M sodium hydroxide (NaOH) or 1M sulphuric acid ( $\text{H}_2\text{SO}_4$ ). The vessel was kept at

a temperature of 30°C, with an agitation of 650 rpm. The culture was aerated by the addition of air at 1.5 mL/min. A steady state was established over a minimum of 5 volume changes, using medium containing normal glucose. Once the steady state had been established the medium was changed to one containing [U -  $^{13}\text{C}$ ] glucose, for one volume change. Samples were taken at  $t = 0, 2, 5, 8, 10, 20, 30, 45, 60$  mins and then every 30 mins thereafter until the volume of  $^{13}\text{C}$  was depleted. These measurements were repeated for a volume change back to  $^{12}\text{C}$ .

A second experiment was performed using the above conditions, but with a set point of 85%, i.e. a higher biomass density.

### **Sample collection**

For the metabolic footprinting (exometabolome) data, 5ml aliquots of culture were sampled from the steady state cultures. Separation of the footprint sample from cells was performed by syringe filtration (0.22 $\mu\text{m}$ , Sartorius, UK) of the culture and snap-freezing of the filtrate. The Intracellular data was obtained by spraying culture into a quenching solution of 60% methanol/water (MEOH/H<sub>2</sub>O) solution [22, 127, 21]. This was then centrifuged for 5 mins at -9°C at 4500 x g. An aliquot of supernatant was removed from each sample to assess the metabolite leakage for quality control. Cell extractions were performed in 80%/20% MEOH/H<sub>2</sub>O solution. Three freeze/thaw cycle in liquid nitrogen/dry ice were performed, before centrifuging (12000 x g, -9°C, 5 mins) to pellet the biomass. The supernatant was returned and a further aliquot of 80% aqueous MEOH was added to the biomass. The freeze thaw cycles were repeated and the samples were centrifuged and the supernatant was added to the first aliquot. Both footprint and intracellular samples were lyophilised using 200 $\mu\text{l}$ -footprint, 1000 $\mu\text{l}$  of sample, respectively

### **6.2.2 Mass Spectrometry (MS) analysis**

Two separate analytical experiments were performed; (a) quantitative determination of  $^{13}\text{C}$  labelling patterns and (b) absolute quantitation of metabolites. GC-ToF-MS analysis was employed in both experiments to analyse the metabolites from samples of the intra-cellular metabolome and exo-metabolome. Derivatisation is required to convert

metabolites into volatile compounds for MS analysis. All samples were chemically derivatised by applying a two-stage process of re-suspending and heating the sample in 50 $\mu$ l of 20mg/ml O-methoxylamine in pyridine solution at 40°C for 90 minutes followed by heating the resultant solution in MSTFA at 40°C for 90 minutes (full details provided in [135]). The samples were analysed using an Agilent 6890 GC and 7673 autosampler coupled to a LECO Pegasus III ToF mass spectrometer using the optimal settings previously determined for *Saccharomyces cerevisiae* [136]. The raw data were processed using LECO ChromaTof V2.12 software and its associated chromatographic deconvolution algorithm with the following settings, baseline 1.0, data point averaging of 3 and average peak width of 1.8s [135]. All samples were analysed in a random order.

For quantitative determination of  $^{13}\text{C}$  labelling a reference database was prepared comprising the mass spectrum and retention index of unique metabolite peaks detected in the analyses. For each unique metabolite two entries were present in the reference database (i) the endogenous unlabelled metabolite ( $^{12}\text{C}$  metabolite peak) and (ii) the endogenous partially labelled metabolite ( $^{13}\text{C}$  metabolite peak). Unique quantification ions were chosen for the  $^{12}\text{C}$  and  $^{13}\text{C}$  metabolite peaks to allow peak construction and peak area determination. The quantification ion for the  $^{13}\text{C}$  metabolite was chosen to represent the maximum labelling pattern for that metabolite. The quantification ions for the  $^{12}\text{C}$  and  $^{13}\text{C}$  metabolite peaks were chosen so to be similar.

For absolute quantitation of metabolites, samples were analysed with appropriate calibration standards (of a range of metabolite concentrations operating in the linear dynamic range of the analytical method) prepared from authentic chemical standards. A calibration curve was constructed with data acquired before and after sample analysis in a single analytical experiment. A reference database was prepared comprising the mass spectrum and retention index of unique metabolite peaks detected in the analyses. The peak areas for calibration standards were applied to construct a single calibration curve for each metabolite. The peak areas determined for each metabolite in each sample were compared to the calibration curves to define the concentration of each metabolite in the derivatised sample. The number of moles of each metabolite in the extracted sample was then determined.

### 6.2.3 Computational Analysis

A reduced metabolic network including glycolysis, glycerol synthesis, pentose phosphate pathway (PPP) and tricarboxylic acid cycle (TCA) pathways was assembled as in Figure 6.2. This network was representative of the  $^{12}\text{C}$  metabolite pools, and highlighted key points of flux differentiation within the network. The addition of  $^{13}\text{C}$  to the system should not disturb the macroscopic steady state of the system, so this addition should leave the pool sizes containing the same overall concentration. In order to keep this consistent the  $^{13}\text{C}$  pools were set as assignments within the model, according to Equation 6.2.1.

$$[^{13}\text{C}_i] + [^{12}\text{C}_i] = [X_i] \quad (6.2.1)$$

Where  $[X_i]$  is the metabolite pool concentration of the corresponding metabolite  $i$ . Concentrations are denoted by square brackets, in accordance with the traditional chemical representation. Both  $[^{13}\text{C}_i]$  and  $[^{12}\text{C}_i]$  represent the concentrations of each isotope within the pool at a given time point. The  $^{13}\text{C}$  enrichment of the pools was reported as a percentage.

The rate law of each reaction was set as mass action because the metabolic system is at steady state for the modelling. Therefore the modelling of the atom transfer does not need a complicated dynamic rate law, but it does need to respond to 0 concentrations in the metabolic pool to ensure stability in the final steady state computation of the fitted model.

The model was fitted to a 3 point rolling average, smoothed version of the raw experimental data. The rolling average was used to reduce the noise present in the original data set. Regression analysis with weighted least squares using the simulated annealing algorithm was carried out in COPASI [47]. All parameter constants  $k$  were allowed to vary between  $1 \times 10^{-5}$  and  $1 \times 10^{+6}$ .

To ensure that the achieved best fit was close to a system solution, the initial rates using the fitted values were taken at time 0, before isotopic perturbation. The model was then altered so a full steady state of the fitted system could be calculated. Where a good fit is achieved the initial rates should match closely the steady state rates of the system.

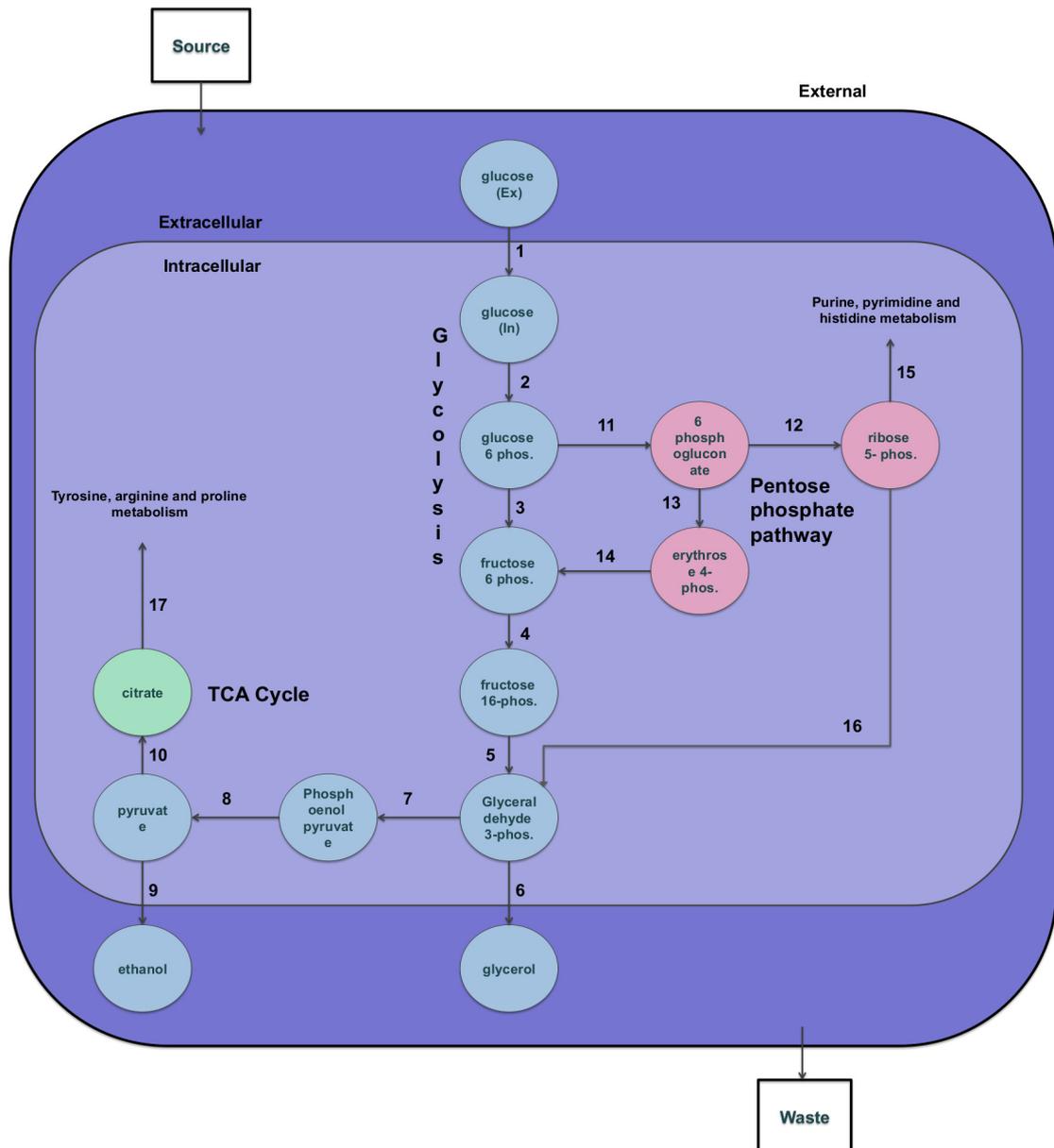


Figure 6.2: Reduced network used for modelling. It identifies only the pools that carry carbon flux through the network and focusses primarily on the ‘hub’ reactions within the network.

The concentrations at steady state were also compared to the measured concentrations of the system. Again, where a good fit is achieved these should match closely.

## 6.3 Results and Discussion

### 6.3.1 Experimental results

Figure 6.3 shows the time at which the highest percentage of  $^{13}\text{C}$  was recorded, and the percentage  $^{13}\text{C}$  that was present in the metabolic pool at that time, for a wide range of

metabolites detectable by GC-MS. The data shows that on the whole the pools respond quickly to the switch from the  $^{13}\text{C}$  media feed back to the  $^{12}\text{C}$ . This is demonstrated by the tight clustering of points around 210 minutes. Figure 6.4 shows that extracellular glucose did not achieve full saturation with  $^{13}\text{C}$ , and with the highest recorded  $^{13}\text{C}$  level of  $\approx 80\%$ . A number of pools cluster above this point showing a higher  $^{13}\text{C}$  percentage than should be feasible given the levels recorded in the extracellular glucose. It is possible that this is the result of noise within the measurements. Table 6.1 shows that metabolite pools that show an unfeasibly high percentage of  $^{13}\text{C}$  at the peak, when compared to the extracellular glucose, are more probable when the peak area is low (in this case  $10^6$ ). In these pools, small calculation errors are more likely to lead to larger errors in the final data, and could explain the data anomalies seen.

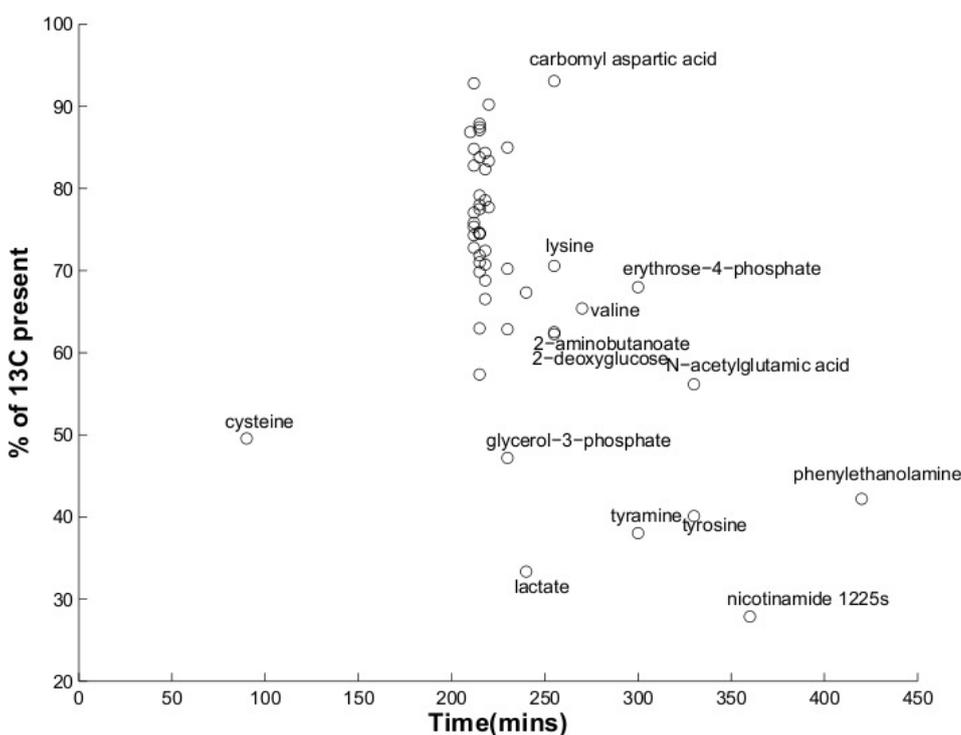


Figure 6.3: Time points where  $^{13}\text{C}$  concentration was highest in each metabolic pool

Peak area	$10^7$	$[10^7, 10^8]$	$10^8$
Expected distribution (%)	47	30	23
highest $\%^{13}\text{C}$	(%)	(%)	(%)
> 90	82	0	18
> 80	59	18	23
< 80	44	35	21
< 80 > 70	50	30	20
< 70	50	30	20

Table 6.1: Peak area associated with highest percentage  $^{13}\text{C}$  levels

A trend can also be seen in the data in Figure 6.3 from the central point of saturation towards the right, showing a slower response of some pools to the changes in glucose input. This behaviour is supported by the time-course drift present in the tyrosine metabolic pool (Figure 6.6). A drift is most likely the result of a metabolic pool being far downstream from the system input, and also carrying a low flux. The low flux would result in a significant time delay in the transfer of isotopic label.

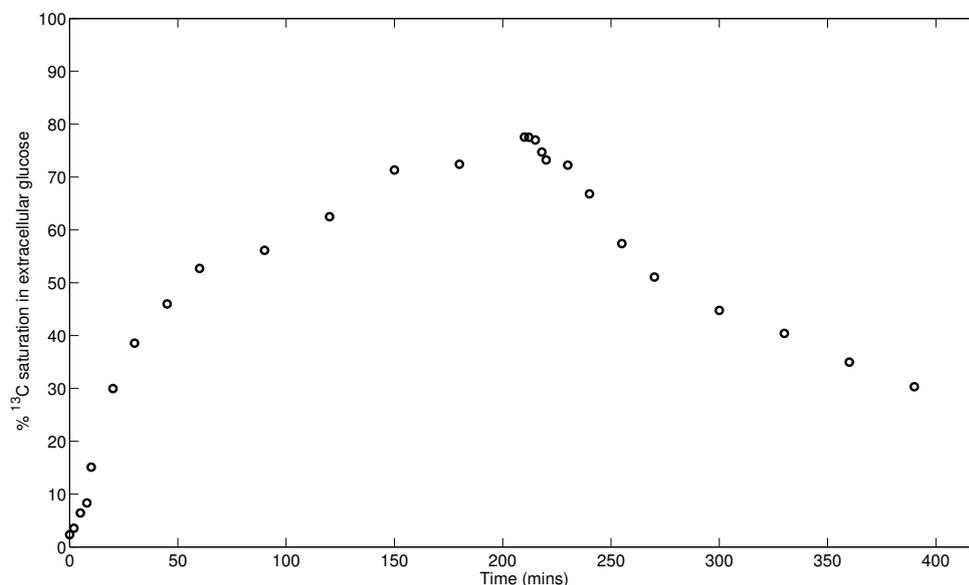
Figure 6.4:  $^{13}\text{C}$  saturation in extracellular glucose

Figure 6.5 shows the time course data obtained for the four pathways that will be modelled to investigate the metabolic flux. These are glycolysis, glycerol synthesis, PPP and TCA pathways. All four of the pathways show a saturation rate similar to that of

the extracellular glucose, showing that these pathways carry flux that is quick to respond to the isotopic perturbation. TCA does appear to saturate at  $\approx 10\%$  higher than that of the extracellular glucose pool, but the pool concentrations in the TCA cycle are relatively low, at around 0.2 mM. It is possible that the high percentage of  $^{13}\text{C}$  is attributable to small calculation deviations when computing the peak areas. For other time points across the TCA data set, the saturation trends are compatible with those in glycolysis.

The metabolite pools that were detected in glycolysis (Figure 6.5 (a)) are the first three pools of the pathway. Intracellular glucose and glucose 6-phosphate show saturation rates that are virtually identical showing their close proximity. There is a larger delay in the rate at which  $^{13}\text{C}$  saturates into the fructose 6-phosphate pool. This could be caused by PPP which also uses glucose 6-phosphate as a precursor. The PPP feeds through to purine and pyrimidine synthesis, but also feeds metabolic flux back into fructose 6-phosphate. The delay, therefore, may be a result of a longer route of production that circumnavigates the phosphoglucosomerase enzyme.

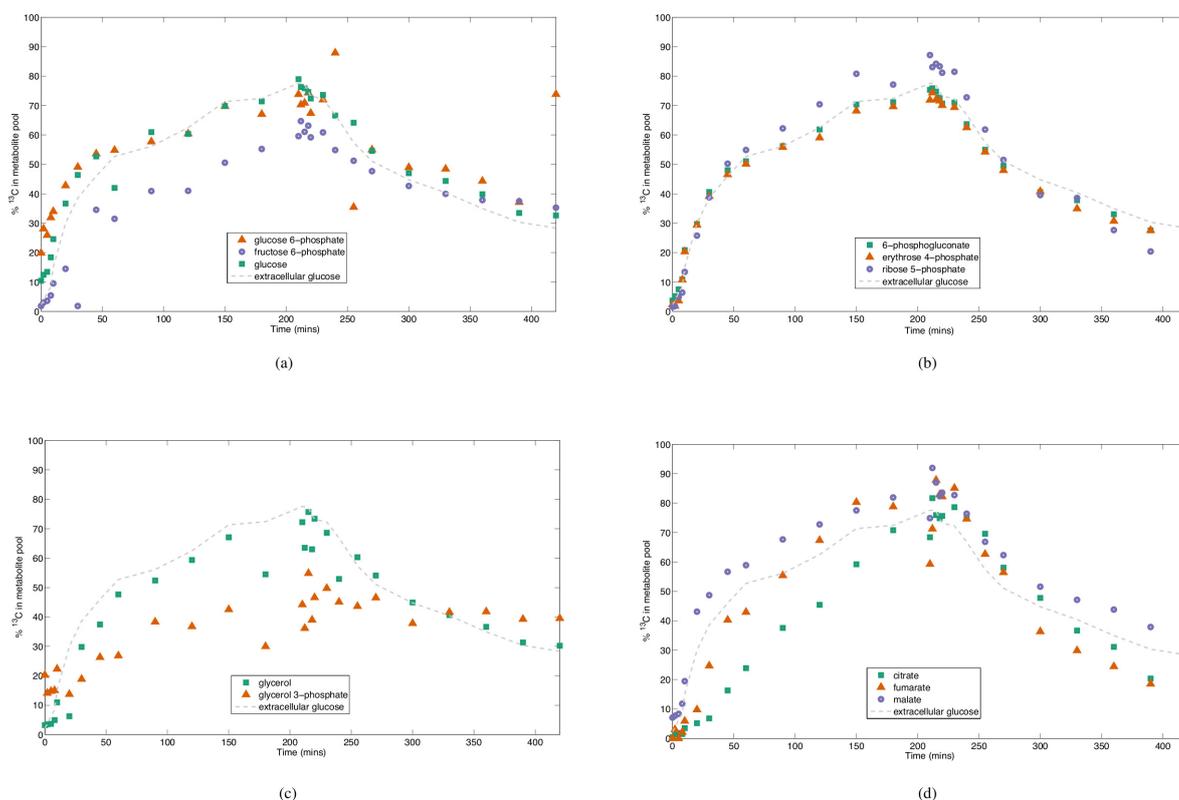


Figure 6.5: Dynamic flux data of metabolite pools within metabolic pathways. (a) glycolysis (b) pentose phosphate pathway (c) glycerol (d) citric acid cycle

The glycerol pathway shows unexpected behaviour. Its structure, in conjunction with the network structure, indicate that the primary direction of flux should be first glycerol 3-phosphate then glycerol. Figure 6.5 (c) shows quite convincingly that isotopic saturation of the glycerol pool occurs at a faster rate than that of glycerol 3-phosphate. The concentrations of the two pools are not hugely different with glycerol 3-phosphate at 0.1 mM and glycerol at 0.09 mM. In this experiment, each metabolic pool is already at steady state, and the system is well mixed. For linear branches, this means that the maximum percentage of  $^{13}\text{C}$  that can be incorporated into the downstream pool is equal to the percentage contained in the upstream pool. Therefore, when glycerol 3-phosphate has a  $^{13}\text{C}$  percentage equal to 20, the enzymatic conversion of glycerol 3-phosphate to glycerol can only result in a 20%  $^{13}\text{C}$ , or less, concentration in the glycerol pool. This behaviour was also noted in the 85% biomass culture. It is possible that the glycerol is being made through another metabolic route, but given the well studied nature of the pathway, this seems unlikely.

The Tricarboxylic Acid Cycle (TCA), when operating as a cycle, should process flux in the direction of citrate to fumarate to malate. Figure 6.5 (d) shows that for the growth conditions presented, TCA does not appear to be operating as a cycle, and also appears to be using the glyoxylate shunt. This is also something that was identified in the FBA in Chapter 5. It is a phenomenon that has been noted in some instances to be associated with pH ranges of 4-5 [118], used for optimal growth. Although findings by Schuster in 1999 [137], using elementary flux modes, suggest that the idea of the TCA operating a cycle is a misnomer, and under most biological conditions it will be acting as a range of disjoint branches.

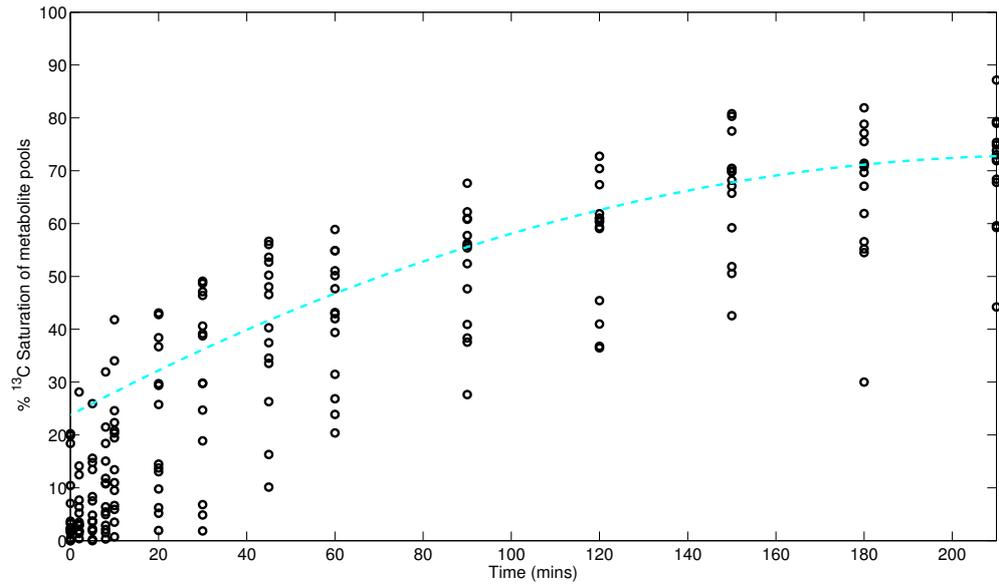


Figure 6.7: Mapping the  $^{13}\text{C}$  saturation trend of 14 different metabolic pools

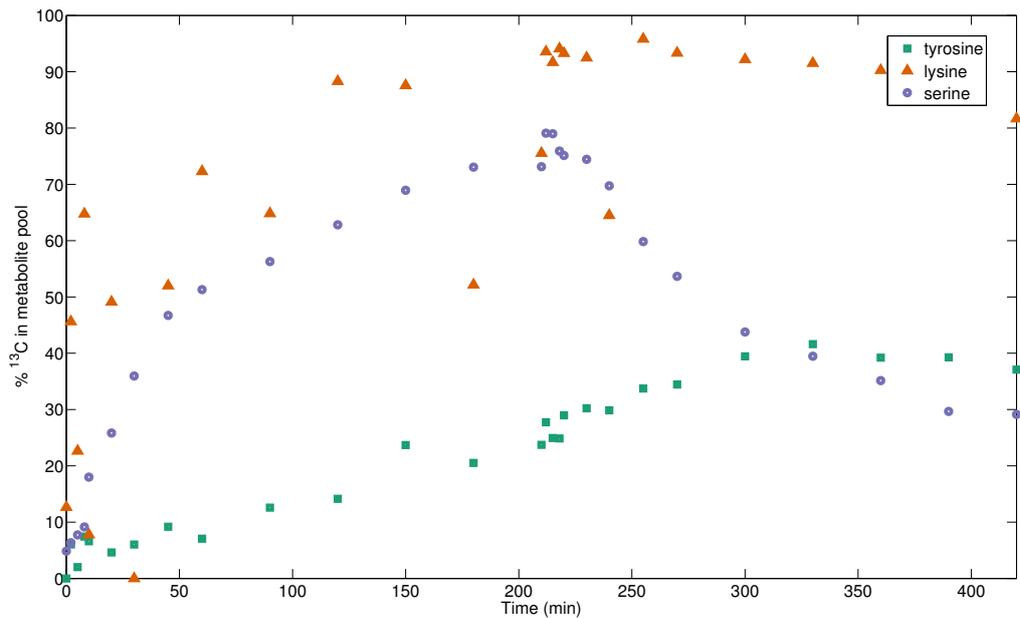


Figure 6.6: Showing the trends of  $^{13}\text{C}$  saturation in amino acid pools

All of the time course data demonstrate a distinct curve, as can be seen in Figure 6.7, which maps 14 different metabolic pools as they respond to the  $^{13}\text{C}$ . It would be difficult to replicate these data with a straight line best fit. The saturation rate of the extracellular glucose pool (Figure 6.4) also shows similar behaviour. This behaviour can be attributed to the different time scales in which the pools operate.

### 6.3.2 Computational results and discussion

The calculated flux data can be seen in Figure 6.9, and in Table 6.2. Given that the model at time 0 should be in a macroscopic steady state, a good fit would be demonstrated by steady state fluxes and concentrations closely matching the initial rate values, before isotopic perturbation, because both states should be equivalent. The data show that this is true for most reactions in the pathway, and further evidence of this can be seen in Appendix C. However, the fluxes associated with reactions 5, 6, 7, 8 and 9 show discrepancies. The concentration data shown in Table 6.3 supports this, showing that all metabolites demonstrate good correlations between the concentration at steady state, and the measured concentrations apart from fructose 1,6-bisphosphate, phosphoenol pyruvate and pyruvate. Again these are metabolites that are centred around the poorer fitted fluxes. The poorer fitting centres around reactions associated with the complex branching point, represented in the model by the combined pool, named glyceraldehyde 3-phosphate (see Diagram 6.8). There were no ratio data available for the glyceraldehyde 3-phosphate pool, hence the need for it to be presented as a combined pool. It is possible that collecting ratio data for this metabolite could help to improve the fit.

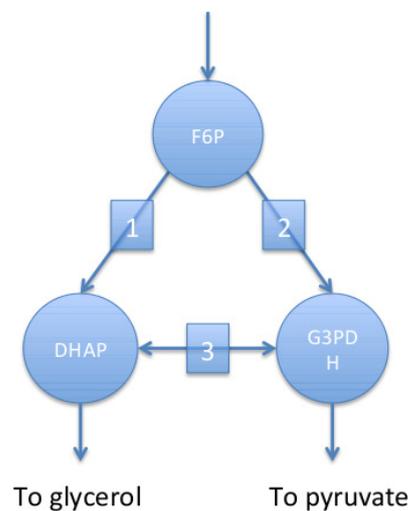


Figure 6.8: Glyceraldehyde 3-phosphate has been used as a collective pool in the reduced model (see Figure 6.2) when in glycolysis the branch would be a branch point as illustrated. Reaction 3 is a very fast equilibrium reaction.

Focusing on the flux results of the model (Figure 6.9) it can be seen that the data show that fructose 6-phosphate is not made solely by the action of the enzyme phosphoglucosomerase (represented by reaction 3), but actually bypasses this step and

is produced via the PPP. This is something that has been noted in *E.coli* [26] and yeast [118]. The magnitude of this effect, as calculated in Blank and Sauer (2004) [118], is only 4% of the total flux, meaning 96% is still carried through phosphoglucoisomerase. The fitted flux model shows that 93% of the flux is diverted through PPP. This is surprising behaviour. Some of the largest fluxes that have been reported for PPP, in recombinant strains of *Aspergillus oryzae*, show up to 47%. These data were calculated using the sMFA approaches, so it is possible that these approaches were not able to trace the magnitude of the flux going through the PPP. It could be advantageous to the cell to carry flux through the PPP because it processes flux for purine, pyrimidine and histidine metabolism. Given that purines and pyrimidines are necessary precursors for DNA and RNA, it could be that maintaining this pathway allows the cell to respond quickly to environmental changes that would increase or decrease protein demand. It should be noted, however, that the flux behaviour of the system is limited to the network that is used for the fit. In this instance, the non-targeted approach for metabolite identification led to the network presented in Figure 6.9 being the most comprehensive possible with the data. It is possible that a network that allows for more complex interactions within the PPP to be included could alter the flux distribution.

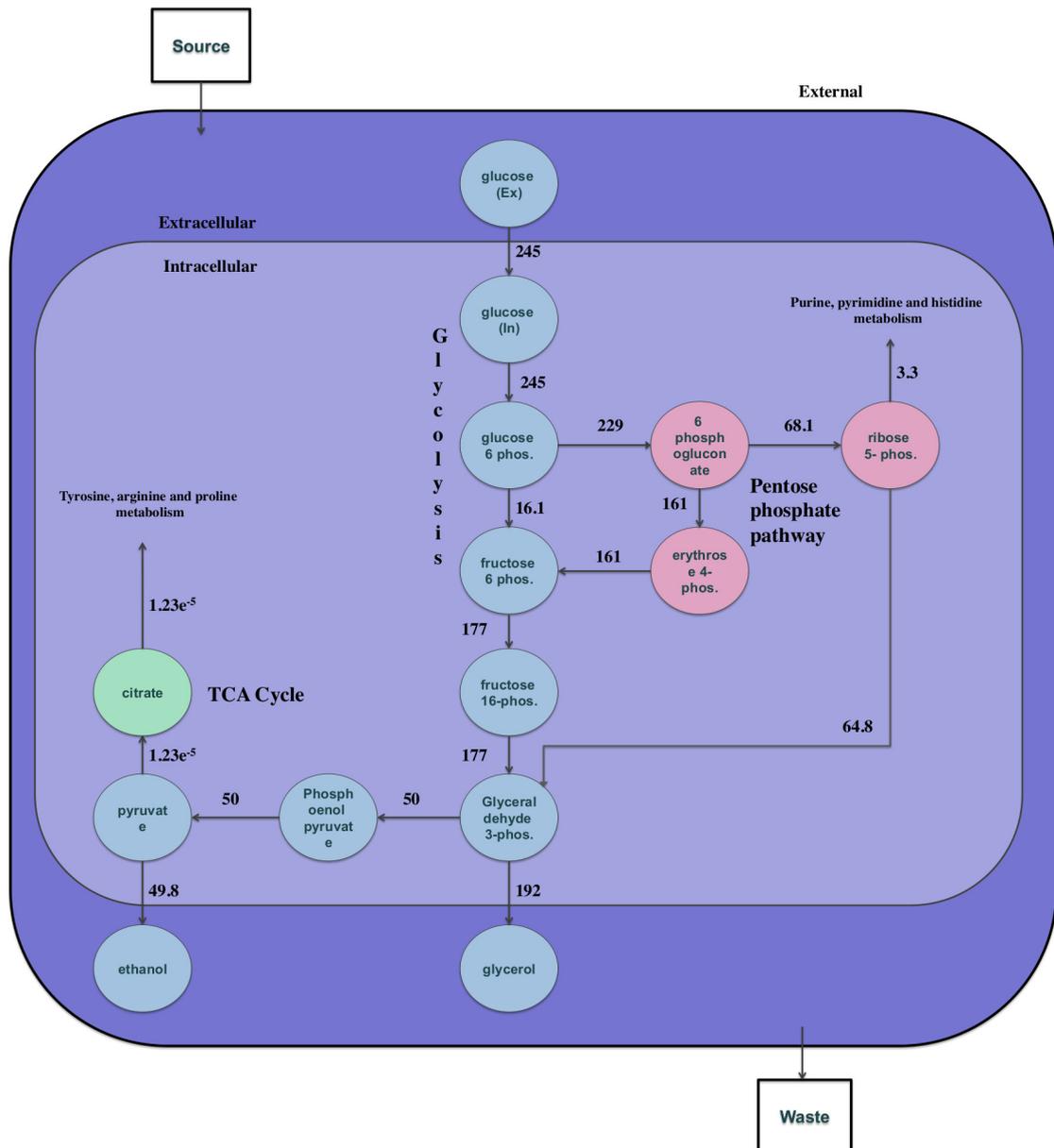


Figure 6.9: Flux results. All flux values are given in mM/s

The flux split through the pentose phosphate pathway between fructose 6-phosphate and glyceraldehyde 3-phosphate has been calculated in glucose limited cultures to be 78% and 22% respectively [138]. This was calculated using distribution data. In this model the fluxes show a route to fructose 6-phosphate to carry 60% of the flux back to glycolysis and the route to glyceraldehyde 3-phosphate to carry 40% of the flux back to glycolysis. This is a different distribution, but is possibly accountable for by the glucose excess.

The flux through the TCA is shown here to be minimal. The flux through the TCA has been noted in the literature to be inversely correlated with the glucose uptake and growth rate [118], although this was only a weak correlation. The flux calculated here supports

Table 6.2: Quantified flux results for each reaction

#	Reaction	Initial Rate (mM/L)	Steady State (mM/L)
1	glucose (ext) → glucose (int)	245	245
2	glucose(int) → glucose 6-phosphate	264	245
3	glucose 6-phosphate → fructose 6-phosphate	19.6	16.1
4	fructose 6-phosphate → fructose 1,6-bisphosphate	153	177
5	fructose 1,6-bisphosphate → glyceraldehyde 3-phosphate	743	177
6	glyceraldehyde 3-phosphate → glycerol	57.5	192
7	glyceraldehyde 3-phosphate → phosphoenolpyruvate	14.9	50
8	phosphoenolpyruvate → pyruvate(fixed)	$1.62 \times 10^{-2}$	50
9	pyruvate → ethanol	317	49.8
10	pyruvate(fixed) → citrate	$7.79 \times 10^{-5}$	$1.23 \times 10^{-5}$
11	glucose6phosphate → 6phosphogluconate	279	229
12	6phosphogluconate → ribose 5-phosphate	69.0	68.1
13	6phosphogluconate → erythrose4phosphate	163	161
14	erythrose 4-phosphate → fructose 6-phosphate	159	161
15	ribose 5-phosphate → purine, pyrimidine and histidine metabolism	3.3	3.3
16	ribose 5-phosphate → glyceraldehyde 3-phosphate	65.0	64.8
17	citrate → tyrosine, arginine and proline metabolism	$1.27 \times 10^{-5}$	$1.23 \times 10^{-5}$

this finding.

Overall the fluxes through the pathways are around 3 times higher than similar fluxes presented in already available glycolysis models [15]. The glycolysis model fluxes have typically been calculated using *in vitro* kinetic data. As such they are not directly comparable.

Table 6.3: Comparison of measured concentrations and steady state concentrations

Species	Measured Concentration (mM)	Steady State (mM)
glucose 6-phosphate	$3.28 \times 10^{-1}$	$4.00 \times 10^{-1}$
fructose 6-phosphate	$1.16 \times 10^{-1}$	$1.00 \times 10^{-1}$
fructose 1,6-bisphosphate	$4.76 \times 10^{-1}$	2.00
glyceraldehyde 3-phosphate	$6.67 \times 10^{-1}$	$2.00 \times 10^{-1}$
phosphoenolpyruvate	$9.22 \times 10^{+2}$	$3.00 \times 10^{-1}$
pyruvate	$1.57 \times 10^{-1}$	1.00
glycerol	$9.88 \times 10^{-2}$	$9.00 \times 10^{-2}$
citrate	$8.68 \times 10^{-2}$	$9.00 \times 10^{-2}$
6-phosphogluconate	$6.90 \times 10^{-2}$	$7.00 \times 10^{-2}$
ribose 5-phosphate	$2.00 \times 10^{-2}$	$2.00 \times 10^{-2}$
erythrose 4-phosphate	$2.02 \times 10^{-2}$	$2.00 \times 10^{-2}$
glucose (intracellular)	2.78	3.00

## 6.4 Conclusions

The aim of this chapter was to gain further insights into the physiology of the cell by directly measuring and quantifying the fluxes *in vivo* using dMFA. The results were surprising, showing a much larger than expected flux through the PPP and overall a higher than expected flux through each reaction. There were also some potential network discrepancies highlighted in the glycerol pathway. They highlight that the current knowledge of the pathway may be not as complete as initially thought, and that further investigation into the *in vivo* behaviour of the cell is vital.

There were some discrepancies in the model fluxes, and what would be expected from a fermenting yeast strain. These are attributable to the sparse ratio data for metabolite pools after the branching point in glycolysis. These could potentially be improved by taking a more targeted approach to the data collection for future implementations of the technique. Coupling the flux fitting with constraints based on sMFA taken under the same conditions may also provide flux reproductions that are more accurate for the cell, than the techniques would when conducted separately.

Overall, the technique shows promise for increasing the knowledge of cellular

behaviour by highlighting flux behaviour that may not have previously been considered. This process of discovery, can be followed by validation before being included in the larger-scale models, and should really drive the understanding of the cell on a more holistic level.

## **6.5 Acknowledgements**

I would like to thank Warwick Dunn for performing the Mass Spectrometry analysis for the chapter.

## SUMMARY AND FUTURE DIRECTIONS

### 7.1 Summary

The title ‘Towards a genome-scale kinetic metabolic model of yeast’, describes the content of the thesis well. It aimed to extend the number of reactions characterised in kinetic models of yeast from what was typically available ( $\approx 20$  reactions) towards wider genome coverage, where there are more than 1000 known reactions. The driving force of this project was to develop methods that could be applied when only sparse data were available for characterising the network behaviour, but that could produce realistic biological insights.

A generic methodology was developed through Chapters 2, 3 and 4. Starting with Chapter 2, which used a small example network to identify a suitable rate law, and associated parameterisation methodology that could be applied to a metabolic network when minimal information was available. This was described as a ‘first-approximation’ rate law and methodology. The convenience rate law and associated approach was selected as the most appropriate based on successful reproduction of network behaviour during control analysis and extrapolation beyond the steady state to which it was fitted. Coupled to this, the rate law also described the expected behaviour of an enzyme catalysed conversion of reactants to products, such as a zero flux when substrate concentrations tend to zero, as well as showing a saturation flux when the metabolite concentration tends to infinity.

In order to apply the convenience kinetics to a complex metabolic network, with interconnected loops, a way of integrating known equilibrium constants and assigning values to those that are unknown, in a way that does not violate thermodynamics was developed in Chapter 3. This was based on the associated link between equilibrium constants and Gibbs free energies. The Gibbs free energy quantify the energy change that occurs during the structural changes associated with conversion of reactants to products. If these values are not consistent across metabolic loops the network will be in violation of the first law of thermodynamics, and will be in perpetual motion when isolated from its surroundings. The method developed can be used to check thermodynamic feasibility of the equilibrium constants in a network, and also to ‘fit’ a thermodynamically feasible set of equilibrium constants.

Further developments were made towards the application of the first-approximation methodology in Chapter 4, where it was applied to a generic example of yeast metabolism. FBA was used as a way of generating system-wide steady state fluxes, using known flux data to constrain the solution space. A way of assigning concentrations to metabolite pools when no quantitative data was available was also developed. The network was then ‘trimmed’ so that only flux-carrying reactions were considered when the first approximation kinetics were applied. I showed that the methodology is suitable for a generating a large-scale model at steady state, and that MCA analysis could be used to highlight high control reactions within the network.

This generic methodology was then extended in Chapter 5 to a specific model, that reflects the behaviour of a yeast strain in a defined set of steady state growth conditions. The model developed contains 284 reactions. Extra modifications to the methodology were made in the form of assigning new rate laws to lipid production and biomass production, that were dependent on the extracellular glucose concentration. By assigning dynamic rate laws to these reactions, network dynamics could be computed accurately. MCA results from the model were used to identify important control reactions in the network. The high control reactions were iteratively replaced with the *in vitro* calculated kinetic rate laws (similar to the method outlined in Chapter 2) until the model could show a reasonable prediction of transient behaviour to extracellular glucose perturbations

above and below the reference concentration. Remarkably, the model showed marked improvement after the replacement of just two correct rate laws.

It was noted that the success of this methodology is heavily dependent on reasonable approximations of the fluxes through different metabolic pathways. This was an area of research where minimal experimental data was available. Chapter 6 further developed an little-used, existing technique for dynamic MFA. The approach was used to directly measure the fluxes through pathways in the central carbon metabolism. The technique allowed for further insights into the flux behaviour of the metabolic network, and produced data that could be applied as FBA constraints for generation of the large-scale network.

Overall the thesis has focussed strongly on the idea that further research should be driven by previous findings in order to build a complete picture of the issues and solutions. Each chapter is driven by the findings and questions raised in the previous chapters, and boundaries between computational and experimental work have been broken down. This follows the philosophy outlined in Chapter 1 with regards to the ‘cycle of knowledge’ where all experiments are developed and conducted based on prioritising the information that is required. Often this cycle is driven by findings from modelling, which are then validated using experimental techniques, before being used to improve the original model, then repeating the cycle. This thesis shows that following a targeted methodology increases the biological knowledge that can be gained, whilst simultaneously streamlining the lengthy experimentation that needs to be conducted, making the discovery process more efficient. Clearly showing that modellers and experimentalists need to work closely together so that iterative feedback between model evolution and improvement, and direction of desirable experiments for discovery can be balanced.

## **7.2 Future directions - the thesis**

The findings in the thesis are being used to foster a collaboration in Berlin, with the aim of further improving the thermodynamic behaviour of models. Firstly there is a method being developed to detect whether the flux values computed from the FBA are thermodynamically consistent. The method will identify flux distributions within the

network that result in a model under perpetual motion. This method is only required during computation of traditional FBA techniques; gFBA, used in the thesis, produces flux distributions that are thermodynamically feasible, so the consideration was not required for the methods presented.

Per-reaction thermodynamics is also being investigated, to allow the  $K_m$  values to be balanced in accordance with the Wegscheider conditions. This advances holistic balancing that was recently developed by that group [114]. This new method will use the equilibrium balancing outlined in this thesis, and then balance the  $K_m$  values to be consistent with the equilibrium constant. The  $K_m$  values affect the transient behaviour of the rate law, by using individual reaction balancing, small parameter alterations can be made within the network to manipulate the transient behaviour of the model. If targeted correctly this will lead to a finer control over the flux behaviour when constructing models, whilst also retaining the thermodynamic integrity of the system.

Further development of dynamic MFA in Chapter 6 is also being explored. The aim is to expand the data set so a more complex interaction network can be used for flux calculation. Static MFA data will also be analysed so a broader picture of flux behaviour can be obtained. These data could then be used to validate and further refine the strain-specific model, built in Chapter 5.

Rate law considerations could also be investigated. The same style of rate law was used for each reaction in the first-approximation methodology. Given that there is a large range of first-approximation rate laws available (see [83] for a small collection) it may be useful to investigate whether certain groups of enzymes are likely to show behaviour that correlated more with one rate law, than compared to another. This could potentially improve the reaction dynamics greatly. An area where this may be most useful is for the transport reactions that are not enzymatically controlled because of the high level of control they have been shown to have over the network behaviour.

### **7.3 Future directions - the field of study**

In the development of models it is important that allosteric regulation and signalling should be integrated with the metabolic network. This is not a trivial task. Experimental

approaches are laborious, time consuming and limited. An ability to identify whether a reaction may be affected by allosteric regulation, and the conditions that cause the regulation to occur could be a more simplistic, but effective way of introducing allosteric regulation to the model. This way, alterations of reaction rate would be attributed to known changes in system behaviour, but not necessarily the exact metabolite causing the changes. The inclusion of signalling behaviour is also vital. Changes in growth conditions or metabolite concentrations can trigger a metabolic response to limit the impact on the homeostatic behaviour of the cell. This response typically involves altering transcription and translation rates in different pathways to alter the flux. To be able to include more of this behaviour into the model, would be of huge benefit for improving the generalisation of models.

The way cells ‘grow’ in the idealisation of FBA, involves a pseudo-reaction designed to reflect the composition of biomass. There is uncertainty associated with this equation and there is certainly a need to improve its accuracy. Coupled to this, its place in kinetic models is questionable. The community needs to start looking beyond the generation of models that only include metabolism, or cell cycle, or signalling, and actually integrate these together so that full simulations on whole cells that ‘grow’ *in silico* in similar ways to how the cells grow in cultures has to be made. This leap is vital for understanding more about cellular behaviours, but also should allow *in silico* experiments that can explore a much wider range of biological behaviours. Once these methods have been developed and applied to simple organisms, the techniques should then be applied to develop tissue models, and eventually complex organ models.

## 7.4 Final thoughts

The field of systems biology shows exciting prospects for the future. Careful integration and design of modelling and experimentation demonstrates an efficient way of gaining increased biological insight. This could have huge implications on many areas of research including personalised medicine, biotechnology industries, and brewing industries. This thesis has played a small part in pushing the field towards more comprehensive understanding of complex dynamic cellular behaviour. The hope is that

this will be used as a foundation for further development.

## BIBLIOGRAPHY

- [1] H.-G. Holzhütter, “The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks.,” *European journal of biochemistry / FEBS*, vol. 271, pp. 2905–22, July 2004.
- [2] M. A. O’Malley, J. Calvert, and J. Dupré, “The study of socioethical issues in systems biology.,” *The American journal of bioethics : AJOB*, vol. 7, pp. 67–78, April 2007.
- [3] J. H. F. Bothwell, “The long past of systems biology.,” *The New phytologist*, vol. 170, pp. 6–10, Jan. 2006.
- [4] H. V. Westerhoff and B. O. Palsson, “The evolution of molecular biology into systems biology.,” *Nature biotechnology*, vol. 22, pp. 1249–52, Oct. 2004.
- [5] D. B. Kell and S. G. Oliver, “Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era,” *BioEssays*, vol. 26, pp. 99–105, 2004.
- [6] D. B. Kell and J. D. Knowles, *The role of modelling in systems biology*. System Modeling in Cellular Biology: From Concepts to Nuts and Bolts, MIT Press, Apr. 2006.
- [7] C. Li, M. Donizelli, N. Rodriguez, H. Dharuri, L. Endler, V. Chelliah, L. Li, E. He, A. Henry, M. I. Stefan, J. L. Snoep, M. Hucka, N. Le Novère, and C. Laibe, “BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models.,” *BMC systems biology*, vol. 4, p. 92, Jan. 2010.

- [8] R. Curto, E. O. Voit, A. Sorribas, and M. Cascante, “Mathematical models of purine metabolism in man,” *Mathematical Biosciences*, vol. 151, pp. 1–49, July 1998.
- [9] E. Voit, “Biochemical and genomic regulation of the trehalose cycle in yeast: review of observations and canonical model analysis,” *Journal of Theoretical Biology*, vol. 223, pp. 55–78, July 2003.
- [10] M. Bier, B. M. Bakker, and H. V. Westerhoff, “How yeast cells synchronize their glycolytic oscillations: a perturbation analytic treatment.,” *Biophysical journal*, vol. 78, pp. 1087–93, Mar. 2000.
- [11] B. Teusink, M. C. Walsh, K. van Dam, and H. V. Westerhoff, “The danger of metabolic pathways with turbo design,” *Trends in Biochemical Sciences*, vol. 23, pp. 162–169, May 1998.
- [12] J. Wolf, J. Passarge, O. J. Somsen, J. L. Snoep, R. Heinrich, and H. V. Westerhoff, “Transduction of intracellular and intercellular dynamics in yeast glycolytic oscillations.,” *Biophysical journal*, vol. 78, pp. 1145–53, Mar. 2000.
- [13] G. C. Conant and K. H. Wolfe, “Increased glycolytic flux as an outcome of whole-genome duplication in yeast.,” *Molecular systems biology*, vol. 3, p. 129, Jan. 2007.
- [14] L. Pritchard and D. B. Kell, “Schemes of flux control in a model of *Saccharomyces cerevisiae* glycolysis,” *European Journal of Biochemistry*, vol. 269, pp. 3894–3904, Aug. 2002.
- [15] B. Teusink, J. Passarge, C. A. Reijenga, E. Esgalhado, C. C. van Der Weijden, M. Schepper, M. C. Walsh, B. M. Bakker, K. van Dam, H. V. Westerhoff, and J. L. Snoep, “Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry.,” *European journal of biochemistry / FEBS*, vol. 267, pp. 5313–29, Sept. 2000.
- [16] F. Hynne, “Full-scale model of glycolysis in *Saccharomyces cerevisiae*,” *Biophysical Chemistry*, vol. 94, pp. 121–163, Dec. 2001.

- [17] K. Nielsen, “Sustained oscillations in glycolysis: an experimental and theoretical study of chaotic and complex periodic behavior and of quenching of simple oscillations,” *Biophysical Chemistry*, vol. 72, pp. 49–62, May 1998.
- [18] M. Ralser, M. M. Wamelink, A. Kowald, B. Gerisch, G. Heeren, E. A. Struys, E. Klipp, C. Jakobs, M. Breitenbach, H. Lehrach, and S. Krobitsch, “Dynamic rerouting of the carbohydrate flux is key to counteracting oxidative stress.,” *Journal of biology*, vol. 6, p. 10, Jan. 2007.
- [19] F. Lei, M. Rotbo ll, and S. B. Jorgensen, “A biochemically structured model for *saccharomyces cerevisiae*,” *Journal of Biotechnology*, vol. 88, pp. 205–221, July 2001.
- [20] G. Cronwright, J. Rohwer, and B. A. Prior, “Metabolic control analysis of glycerol synthesis in *saccharomyces cerevisiae*,” *Applied and Environmental Microbiology*, vol. 68, pp. 4448–4456, 2002.
- [21] C. L. Winder, W. B. Dunn, S. Schuler, D. Broadhurst, R. Jarvis, G. M. Stephens, and R. Goodacre, “Global metabolic profiling of *Escherichia coli* cultures: an evaluation of methods for quenching and extraction of intracellular metabolites.,” *Analytical chemistry*, vol. 80, pp. 2939–48, Apr. 2008.
- [22] R. P. Maharjan and T. Ferenci, “Global metabolite analysis: the influence of extraction methodology on metabolome profiles of *Escherichia coli*.,” *Analytical biochemistry*, vol. 313, pp. 145–54, Feb. 2003.
- [23] V. Lange, P. Picotti, B. Domon, and R. Aebersold, “Selected reaction monitoring for quantitative proteomics: a tutorial.,” *Molecular systems biology*, vol. 4, p. 222, Jan. 2008.
- [24] U. Sauer, “Metabolic networks in motion: <sup>13</sup>C-based flux analysis.,” *Mol Syst Biol*, vol. 2, p. 62, 2006.
- [25] U. Sauer, “High-throughput phenomics: experimental methods for mapping fluxomes,” *Curr Opin Biotechnol*, vol. 15, pp. 58–63, Feb. 2004.

- [26] E. Fischer and U. Sauer, "Metabolic flux profiling of *Escherichia coli* mutants in central carbon metabolism using GC-MS.," *European journal of biochemistry / FEBS*, vol. 270, pp. 880–91, Mar. 2003.
- [27] E. Fischer, N. Zamboni, and U. Sauer, "High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived  $^{13}\text{C}$  constraints.," *Analytical biochemistry*, vol. 325, pp. 308–16, Feb. 2004.
- [28] U. Sauer, D. R. Lasko, J. Fiaux, M. Hochuli, R. Glaser, T. Szyperski, K. Wüthrich, and J. E. Bailey, "Metabolic flux ratio analysis of genetic and environmental modulations of *Escherichia coli* central carbon metabolism.," *Journal of bacteriology*, vol. 181, pp. 6679–88, Nov. 1999.
- [29] B. Christensen, A. K. Gombert, and J. Nielsen, "Analysis of flux estimates based on  $^{13}\text{C}$ -labelling experiments.," *European journal of biochemistry / FEBS*, vol. 269, pp. 2795–800, June 2002.
- [30] K. Nöh, A. Wahl, and W. Wiechert, "Computational tools for isotopically instationary  $^{13}\text{C}$  labeling experiments under metabolic steady state conditions.," *Metabolic engineering*, vol. 8, pp. 554–77, Nov. 2006.
- [31] K. Nöh, K. Grönke, B. Luo, R. Takors, M. Oldiges, and W. Wiechert, "Metabolic flux analysis at ultra short time scale: isotopically non-stationary  $^{13}\text{C}$  labeling experiments.," *Journal of biotechnology*, vol. 129, pp. 249–67, Apr. 2007.
- [32] C. Wittmann, "Metabolic flux analysis using mass spectrometry.," *Adv Biochem Eng Biotechnol*, vol. 74, pp. 39–64, 2002.
- [33] D. B. Kell, "Metabolomics and systems biology: making sense of the soup.," *Current opinion in microbiology*, vol. 7, pp. 296–307, June 2004.
- [34] D. A. Fell, "Metabolic control analysis: a survey of its theoretical and experimental development.," *The Biochemical journal*, vol. 286 ( Pt 2, pp. 313–30, Sept. 1992.

- [35] M. A. Savageau, "Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions.," *Journal of theoretical biology*, vol. 25, pp. 365–9, Dec. 1969.
- [36] M. A. Savageau, "Biochemical systems analysis. II. The steady-state solutions for an n-pool system using a power-law approximation.," *Journal of theoretical biology*, vol. 25, pp. 370–9, Dec. 1969.
- [37] M. A. Savageau, "Biochemical systems analysis," *Journal of Theoretical Biology*, vol. 26, pp. 215–226, Feb. 1970.
- [38] H. Kacser and J. Burns, "The control of flux," *Symp. Soc. Exp. Biol.*, vol. 27, pp. 65 – 104, 1973.
- [39] R. Heinrich and T. A. Rapoport, "A linear steady-state treatment of enzymatic chains. General properties, control and effector strength.," *European journal of biochemistry / FEBS*, vol. 42, pp. 89–95, Feb. 1974.
- [40] H. V. Westerhoff and Y. D. Chen, "How do enzyme activities control metabolite concentrations? An additional theorem in the theory of metabolic control.," *European journal of biochemistry / FEBS*, vol. 142, pp. 425–30, July 1984.
- [41] J.-H. S. Hofmeyr, "Metabolic control analysis in a nutshell," *Proceedings of the 2nd International Conference on Systems Biology*, no. ii, pp. 291–300, 2001.
- [42] M. A. Savageau, E. O. Voit, and D. H. Irvine, "Biochemical systems theory and metabolic control theory: 1. fundamental similarities and differences," *Mathematical Biosciences*, vol. 86, no. 2, pp. 127–145, 1987.
- [43] R. Curto, A. Sorribas, and M. Cascante, "Comparative characterization of the fermentation pathway of *Saccharomyces cerevisiae* using biochemical systems theory and metabolic control analysis: model definition and nomenclature.," *Mathematical biosciences*, vol. 130, pp. 25–50, Nov. 1995.

- [44] J. H. Hofmeyr, H. Kacser, and K. J. Van Der Merwe, "Metabolic control analysis of moiety-conserved cycles.," *The Federation of European Biochemical Societies Journal*, vol. 155, no. 3, pp. 631–641, 1986.
- [45] C. Reder, "Metabolic control theory: a structural approach.," *Journal of theoretical biology*, vol. 135, pp. 175–201, Nov. 1988.
- [46] J. S. Hofmeyr and A. Cornish-Bowden, "Regulating the cellular economy of supply and demand.," *FEBS letters*, vol. 476, pp. 47–51, June 2000.
- [47] S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, and U. Kummer, "COPASI—a COMplex PATHway SIMulator.," *Bioinformatics*, vol. 22, pp. 3067–74, Dec. 2006.
- [48] K. J. Kauffman, P. Prakash, and J. S. Edwards, "Advances in flux balance analysis.," *Current opinion in biotechnology*, vol. 14, pp. 491–6, Oct. 2003.
- [49] R. Ramakrishna, J. S. Edwards, A. McCulloch, and B. O. Palsson, "Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints.," *American journal of physiology. Regulatory, integrative and comparative physiology*, vol. 280, pp. R695–704, Mar. 2001.
- [50] J. S. Edwards and B. O. Palsson, "The *Escherichia coli* MG1655 *in silico* metabolic genotype: its definition, characteristics, and capabilities.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, pp. 5528–33, May 2000.
- [51] R. Mahadevan, J. S. Edwards, and F. J. Doyle, "Dynamic flux balance analysis of diauxic growth in *Escherichia coli*.," *Biophysical journal*, vol. 83, pp. 1331–40, Sept. 2002.
- [52] A. P. Burgard and C. D. Maranas, "Probing the performance limits of the *Escherichia coli* metabolic network subject to gene additions or deletions.," *Biotechnology and bioengineering*, vol. 74, pp. 364–75, Sept. 2001.

- [53] A. Varma, B. W. Boesch, and B. O. Palsson, "Biochemical production capabilities of *Escherichia coli*," *Biotechnology and Bioengineering*, vol. 42, no. 1, pp. 59–73, 1993.
- [54] K. Smallbone and E. Simeonidis, "Flux balance analysis: a geometric perspective," *J Theor Biol*, vol. 258, pp. 311–315, May 2009.
- [55] S. A. Becker, A. M. Feist, M. L. Mo, G. Hannum, B. O. Palsson, and M. J. Herrgard, "Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox.," *Nature protocols*, vol. 2, pp. 727–38, Jan. 2007.
- [56] R. Mahadevan and C. H. Schilling, "The effects of alternate optimal solutions in constraint-based genome-scale metabolic models.," *Metabolic engineering*, vol. 5, pp. 264–76, Oct. 2003.
- [57] M. E. Bushell, S. I. Sequeira, C. Khannapho, H. Zhao, K. F. Chater, M. J. Butler, A. M. Kierzek, and C. A. Avignone-Rossa, "The use of genome-scale metabolic flux variability analysis for process feed formulation based on an investigation of the effects of the *zwf* mutation on antibiotic production in *Streptomyces coelicolor*," *Enzyme and Microbial Technology*, vol. 39, pp. 1347–1353, Oct. 2006.
- [58] J. Hay and J. Schwender, "Metabolic network reconstruction and flux variability analysis of storage synthesis in developing oilseed rape (*Brassica napus* L.) embryos.," *The Plant journal : for cell and molecular biology*, vol. 67, pp. 526–41, Apr. 2011.
- [59] J. L. Hjersted and M. A. Henson, "Steady-state and dynamic flux balance analysis of ethanol production by *Saccharomyces cerevisiae*," *IET systems biology*, vol. 3, pp. 167–79, May 2009.
- [60] F. A. Vargas, F. Pizarro, J. R. Perez-Correa, and E. Agosin, "Expanding a dynamic flux balance model of yeast fermentation to genome-scale," *BMC Systems Biology*, vol. 5, no. 1, p. 75, 2011.
- [61] C. S. Henry, L. J. Broadbelt, and V. Hatzimanikatis, "Thermodynamics-based metabolic flux analysis.," *Biophysical journal*, vol. 92, pp. 1792–805, Mar. 2007.

- [62] J. Schellenberger, N. E. Lewis, and B. O. Palsson, “Elimination of thermodynamically infeasible loops in steady-state metabolic models.,” *Biophysical journal*, vol. 100, pp. 544–53, Feb. 2011.
- [63] W. B. Dunn, “Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes.,” *Phys Biol*, vol. 5, p. 11001, 2008.
- [64] A. B. Canelas, A. ten Pierick, C. Ras, R. M. Seifar, J. C. van Dam, W. M. van Gulik, and J. J. Heijnen, “Quantitative evaluation of intracellular metabolite extraction techniques for yeast metabolomics.,” *Analytical chemistry*, vol. 81, pp. 7379–89, Sept. 2009.
- [65] C. J. Bolten, P. Kiefer, F. Letisse, J.-C. Portais, and C. Wittmann, “Sampling for metabolome analysis of microorganisms.,” *Analytical chemistry*, vol. 79, pp. 3843–9, May 2007.
- [66] M. A. H. Luttik, Z. Vuralhan, E. Suir, G. H. Braus, J. T. Pronk, and J. M. Daran, “Alleviation of feedback inhibition in *Saccharomyces cerevisiae* aromatic amino acid biosynthesis: quantification of metabolic impact.,” *Metabolic engineering*, vol. 10, no. 3-4, pp. 141–53, 2008.
- [67] E. Martínez-Force and T. Benítez, “Effects of varying media, temperature, and growth rates on the intracellular concentrations of yeast amino acids.,” *Biotechnology progress*, vol. 11, no. 4, pp. 386–92.
- [68] J. Ezaki, N. Matsumoto, M. Takeda-Ezaki, M. Komatsu, K. Takahashi, Y. Hiraoka, H. Taka, T. Fujimura, K. Takehana, M. Yoshida, J. Iwata, I. Tanida, N. Furuya, D.-M. Zheng, N. Tada, K. Tanaka, E. Kominami, and T. Ueno, “Liver autophagy contributes to the maintenance of blood glucose and amino acid levels.,” *Autophagy*, vol. 7, pp. 727–36, July 2011.
- [69] B. Gonzalez, J. François, and M. Renaud, “A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol.,” *Yeast (Chichester; England)*, vol. 13, pp. 1347–55, Nov. 1997.

- [70] H. C. Lange, M. Eman, G. van Zuijlen, D. Visser, J. C. van Dam, J. Frank, M. J. de Mattos, and J. J. Heijnen, “Improved rapid sampling for *in vivo* kinetics of intracellular metabolites in *Saccharomyces cerevisiae*.” *Biotechnology and bioengineering*, vol. 75, pp. 406–15, Nov. 2001.
- [71] W. de Koning and K. van Dam, “A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH.” *Analytical biochemistry*, vol. 204, pp. 118–23, July 1992.
- [72] S. G. Villas-Bôas, J. Hojer-Pedersen, M. Akesson, J. r. Smedsgaard, and J. Nielsen, “Global metabolite analysis of yeast: evaluation of sample preparation methods.” *Yeast (Chichester, England)*, vol. 22, pp. 1155–69, Oct. 2005.
- [73] A. Teleman, P. Richard, M. Toivari, and M. Penttilä, “Identification and quantitation of phosphorus metabolites in yeast neutral pH extracts by nuclear magnetic resonance spectroscopy.” *Analytical biochemistry*, vol. 272, pp. 71–9, July 1999.
- [74] W. B. Dunn, N. J. C. Bailey, and H. E. Johnson, “Measuring the metabolome: current analytical technologies.” *The Analyst*, vol. 130, pp. 606–25, May 2005.
- [75] R. Heinrich and S. M. Rapoport, “Metabolic Regulation and Mathematical Models,” *Proq. Biophys. Molec. Biol.*, vol. 32, pp. 1–82, 1977.
- [76] J. L. Snoep, F. Bruggeman, B. G. Olivier, and H. V. Westerhoff, “Towards building the silicon cell: a modular approach.” *Bio Systems*, vol. 83, no. 2-3, pp. 207–16, 2006.
- [77] P. Mendes, D. B. Kell, and G. R. Welch, *Enzymology in Vivo*, vol. 11 of *Advances in Molecular and Cell Biology*. Elsevier, 1995.
- [78] R. Grima, “Investigating the robustness of the classical enzyme kinetic equations in small intracellular compartments.” *BMC systems biology*, vol. 3, p. 101, Jan. 2009.

- [79] M. J. Herrgård, N. Swainston, P. Dobson, W. B. Dunn, K. Y. Arga, M. Arvas, N. Blüthgen, S. Borger, R. Costenoble, M. Heinemann, M. Hucka, N. Le Novère, P. Li, W. Liebermeister, M. L. Mo, A. P. Oliveira, D. Petranovic, S. Pettifer, E. Simeonidis, K. Smallbone, I. Spasić, D. Weichart, R. Brent, D. S. Broomhead, H. V. Westerhoff, B. Kirdar, M. Penttilä, E. Klipp, B. O. Palsson, U. Sauer, S. G. Oliver, P. Mendes, J. Nielsen, and D. B. Kell, “A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology.,” *Nature biotechnology*, vol. 26, pp. 1155–60, Oct. 2008.
- [80] P. D. Dobson, K. Smallbone, D. Jameson, E. Simeonidis, K. Lanthaler, P. Pir, C. Lu, N. Swainston, W. B. Dunn, P. Fisher, D. Hull, M. Brown, O. Oshota, N. J. Stanford, D. B. Kell, R. D. King, S. G. Oliver, R. D. Stevens, and P. Mendes, “Further developments towards a genome-scale metabolic model of yeast.,” *BMC systems biology*, vol. 4, p. 145, Oct. 2010.
- [81] M. Ederer and E. D. Gilles, “Thermodynamically feasible kinetic models of reaction networks.,” *Biophysical journal*, vol. 92, pp. 1846–57, Mar. 2007.
- [82] W. Liebermeister and E. Klipp, “Bringing metabolic networks to life: convenience rate law and thermodynamic constraints.,” *Theor Biol Med Model*, vol. 3, p. 41, 2006.
- [83] W. Liebermeister, J. Uhlenendorf, and E. Klipp, “Modular rate laws for enzymatic reactions: thermodynamics, elasticities, and implementation.,” *Bioinformatics (Oxford, England)*, Apr. 2010.
- [84] M. A. Savageau, *Biochemical Systems Analysis: A Study of Function and Design in Molecular Biology*. Addison-Wesley Pub (Sd), 1976.
- [85] J. J. Heijnen, “Approximative kinetic formats used in metabolic network modeling.,” *Biotechnology and bioengineering*, vol. 91, pp. 534–45, Sept. 2005.
- [86] W. Liebermeister, “Predicting physiological concentrations of metabolites from their molecular structure.,” *Journal of computational biology : a journal of computational molecular cell biology*, vol. 12, pp. 1307–15, Dec. 2005.

- [87] M. D. Jankowski, C. S. Henry, L. J. Broadbelt, and V. Hatzimanikatis, "Group contribution method for thermodynamic analysis of complex metabolic networks.," *Biophys J*, vol. 95, pp. 1487–1499, 2008.
- [88] F. Hadlich, S. Noack, and W. Wiechert, "Translating biochemical network models between different kinetic formats.," *Metabolic engineering*, vol. 11, pp. 87–100, Mar. 2009.
- [89] K. Smallbone, E. Simeonidis, D. S. Broomhead, and D. B. Kell, "Something from nothing: bridging the gap between constraint-based and kinetic modelling.," *The FEBS journal*, vol. 274, pp. 5576–85, Nov. 2007.
- [90] D. A. Adiamah, J. Handl, and J.-M. Schwartz, "Streamlining the construction of large-scale dynamic models using generic kinetic equations.," *Bioinformatics (Oxford, England)*, vol. 26, pp. 1324–31, May 2010.
- [91] L. J. W. M. Oehlen, M. E. Scholte, W. de Koning, and K. van Dam, "Decrease in glycolytic flux in *Saccharomyces cerevisiae* cdc35-1 cells at restrictive temperature correlates with a decrease in glucose transport," *Microbiology*, vol. 140, pp. 1891–1898, Aug. 1994.
- [92] B. M. Bakker, H. V. Westerhoff, F. R. Opperdoes, and P. A. Michels, "Metabolic control analysis of glycolysis in trypanosomes as an approach to improve selectivity and effectiveness of drugs.," *Molecular and biochemical parasitology*, vol. 106, pp. 1–10, Feb. 2000.
- [93] K. A. Reijenga, J. L. Snoep, J. A. Diderich, H. W. van Verseveld, H. V. Westerhoff, and B. Teusink, "Control of glycolytic dynamics by hexose transport in *Saccharomyces cerevisiae*." *Biophysical journal*, vol. 80, pp. 626–34, Feb. 2001.
- [94] T. Shimizu, J. C. Parker, H. Najafi, and F. M. Matschinsky, "Control of glucose metabolism in pancreatic beta-cells by glucokinase, hexokinase, and phosphofructokinase. Model study with cell lines derived from beta-cells," *Diabetes*, vol. 37, pp. 1524–1530, Nov. 1988.

- [95] R. Moreno-Sánchez, E. Saavedra, S. Rodríguez-Enríquez, and V. Olín-Sandoval, “Metabolic Control Analysis: A Tool for Designing Strategies to Manipulate Metabolic Pathways,” *Journal of Biomedicine and Biotechnology*, vol. 2008, p. 597913, 2008.
- [96] S. J. Van Dien and M. E. Lidstrom, “Stoichiometric model for evaluating the metabolic capabilities of the facultative methylotroph *Methylobacterium extorquens* AM1, with application to reconstruction of C(3) and C(4) metabolism.,” *Biotechnology and bioengineering*, vol. 78, pp. 296–312, May 2002.
- [97] J. Kennedy and R. Eberhart, *Particle swarm optimization*. IEEE.
- [98] F. C. Bing, “The history of the word ‘metabolism’.,” *Journal of the history of medicine and allied sciences*, vol. 26, pp. 158–80, Apr. 1971.
- [99] M. R. S. Vaseghi, A. Baumeister, and M. Reuss, “*In Vivo* Dynamics of The Pentose Phosphate Pathway in *Saccharomyces cerevisiae*,” *Metabolic Engineering*, vol. 1, pp. 128–140, 1999.
- [100] N. C. Duarte, S. A. Becker, N. Jamshidi, I. Thiele, M. L. Mo, T. D. Vo, R. Srivas, and B. O. Palsson, “Global reconstruction of the human metabolic network based on genomic and bibliomic data.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, pp. 1777–82, Feb. 2007.
- [101] D. LeRoith, “Dyslipidemia and glucose dysregulation in overweight and obese patients.,” *Clinical cornerstone*, vol. 8, pp. 38–52, Jan. 2007.
- [102] L. B. Ray, “Metabolism. Metabolism is not boring. Introduction.,” *Science (New York, N.Y.)*, vol. 330, p. 1337, Dec. 2010.
- [103] J. L. Snoep and B. G. Olivier, “Java Web Simulation (JWS); a web based database of kinetic models.,” *Molecular Biology Reports*, vol. 29, no. 1-2, pp. 259–263, 2002.
- [104] J. M. Lee, E. P. Gianchandani, and J. A. Papin, “Flux balance analysis in the era of metabolomics.,” *Briefings in bioinformatics*, vol. 7, pp. 140–50, June 2006.

- [105] C. Khannapho, H. Zhao, B. K. Bonde, A. M. Kierzek, C. A. Avignone-Rossa, and M. E. Bushell, "Selection of objective function in genome scale flux balance analysis for process feed development in antibiotic production.," *Metabolic engineering*, vol. 10, pp. 227–33, Sept. 2008.
- [106] K. Bettenbrock, S. Fischer, A. Kremling, K. Jahreis, T. Sauter, and E.-D. Gilles, "A quantitative approach to catabolite repression in *Escherichia coli*," *The Journal of biological chemistry*, vol. 281, pp. 2578–84, Feb. 2006.
- [107] A. Varma and B. O. Palsson, "Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110.," *Applied and environmental microbiology*, vol. 60, pp. 3724–31, Oct. 1994.
- [108] P. Mendes, D. B. Kell, and H. V. Westerhoff, "Why and when channelling can decrease pool size at constant net flux in a simple dynamic channel.," *Biochimica et biophysica acta*, vol. 1289, pp. 175–86, Mar. 1996.
- [109] K. Smallbone, E. Simeonidis, N. Swainston, and P. Mendes, "Towards a genome-scale kinetic model of cellular metabolism.," *BMC systems biology*, vol. 4, p. 6, Jan. 2010.
- [110] P. Li, J. O. Dada, D. Jameson, I. Spasic, N. Swainston, K. Carroll, W. Dunn, F. Khan, N. Malys, H. L. Messiha, E. Simeonidis, D. Weichart, C. Winder, J. Wishart, D. Broomhead, C. A. Goble, S. J. Gaskell, D. B. Kell, H. V. Westerhoff, P. Mendes, and N. W. Paton, "Systematic integration of experimental data and models in systems biology.," *BMC bioinformatics*, vol. 11, p. 582, Nov. 2010.
- [111] M. Kanehisa and S. Goto, "KEGG: kyoto encyclopedia of genes and genomes.," *Nucleic acids research*, vol. 28, pp. 27–30, Jan. 2000.
- [112] M. Kanehisa, S. Goto, M. Hattori, K. F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki, and M. Hirakawa, "From genomics to chemical genomics: new developments in KEGG.," *Nucleic acids research*, vol. 34, pp. D354–7, Jan. 2006.

- [113] M. Kanehisa, S. Goto, M. Furumichi, M. Tanabe, and M. Hirakawa, “KEGG for representation and analysis of molecular networks involving diseases and drugs.,” *Nucleic acids research*, vol. 38, pp. D355–60, Jan. 2010.
- [114] T. Lubitz, M. Schulz, E. Klipp, and W. Liebermeister, “Parameter Balancing in Kinetic Models of Cell Metabolism.,” *The journal of physical chemistry. B*, Nov. 2010.
- [115] T. Shlomi, O. Berkman, and E. Ruppin, “Regulatory on/off minimization of metabolic flux changes after genetic perturbations.,” *Proc Natl Acad Sci U S A*, vol. 102, pp. 7695–7700, May 2005.
- [116] D. B. Kell, “Systems biology, metabolic modelling and metabolomics in drug discovery and development.,” *Drug discovery today*, vol. 11, pp. 1085–92, Dec. 2006.
- [117] K. Smallbone, N. Malys, H. L. Messiha, and E. S. J A Wishart, *Building a kinetic model of trehalose biosynthesis in Saccharomyces cerevisiae*, vol. 500 of *Methods in systems biology*. Academic Press, in press.
- [118] L. M. Blank and U. Sauer, “TCA cycle activity in *Saccharomyces cerevisiae* is a function of the environmentally determined specific growth and glucose uptake rates.,” *Microbiology (Reading, England)*, vol. 150, pp. 1085–93, Apr. 2004.
- [119] J. Schaub, K. Mauch, and M. Reuss, “Metabolic flux analysis in *Escherichia coli* by integrating isotopic dynamic and isotopic stationary  $^{13}\text{C}$  labeling data,” *Biotechnol Bioeng*, vol. 99, pp. 1170–1185, Apr. 2008.
- [120] H. Dominguez, C. Rollin, A. Guyonvarch, J.-L. Guerquin-Kern, M. Coccagn-Bousquet, and N. D. Lindley, “Carbon-flux distribution in the central metabolic pathways of *Corynebacterium glutamicum* during growth on fructose,” *European Journal of Biochemistry*, vol. 254, pp. 96–102, May 1998.
- [121] L. M. Blank, L. Kuepfer, and U. Sauer, “Large-scale  $^{13}\text{C}$ -flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast.,” *Genome Biol*, vol. 6, p. R49, 2005.

- [122] T. Szyperski, “ $^{13}\text{C}$ -NMR, MS and metabolic flux balancing in biotechnology research.,” *Quarterly reviews of biophysics*, vol. 31, pp. 41–106, Feb. 1998.
- [123] S. Tännler, S. Decasper, and U. Sauer, “Maintenance metabolism and carbon fluxes in *Bacillus* species.,” *Microbial cell factories*, vol. 7, p. 19, Jan. 2008.
- [124] N. Zamboni, S.-M. Fendt, M. Rühl, and U. Sauer, “ $^{13}\text{C}$ -based metabolic flux analysis.,” *Nature protocols*, vol. 4, no. 6, pp. 878–92, 2009.
- [125] L.-E. Quek, C. Wittmann, L. K. Nielsen, and J. O. Krömer, “OpenFLUX: efficient modelling software for  $^{13}\text{C}$ -based metabolic flux analysis.,” *Microbial cell factories*, vol. 8, p. 25, Jan. 2009.
- [126] N. Zamboni, E. Fischer, and U. Sauer, “FiatFlux—a software for metabolic flux analysis from  $^{13}\text{C}$ -glucose experiments.,” *BMC bioinformatics*, vol. 6, p. 209, Jan. 2005.
- [127] M. O. Loret, L. Pedersen, and J. François, “Revised procedures for yeast metabolites extraction: application to a glucose pulse to carbon-limited yeast cultures, which reveals a transient activation of the purine salvage pathway.,” *Yeast (Chichester, England)*, vol. 24, pp. 47–60, Jan. 2007.
- [128] W. Wiechert and M. Wurzel, “Metabolic isotopomer labeling systems. Part I: global dynamic behavior,” *Math Biosci*, vol. 169, pp. 173–205, Feb. 2001.
- [129] W. Wiechert and A. A. de Graaf, “Bidirectional reaction steps in metabolic networks: I. Modeling and simulation of carbon isotope labeling experiments,” *Biotechnol Bioeng*, vol. 55, pp. 101–117, July 1997.
- [130] W. Wiechert, C. Siefke, A. A. de Graaf, and A. Marx, “Bidirectional reaction steps in metabolic networks: II. Flux estimation and statistical analysis,” *Biotechnol Bioeng*, vol. 55, pp. 118–135, July 1997.
- [131] W. Wiechert, M. Möllney, N. Isermann, M. Wurzel, and A. A. de Graaf, “Bidirectional reaction steps in metabolic networks: III. Explicit solution and

- analysis of isotopomer labeling systems,” *Biotechnol Bioeng*, vol. 66, pp. 69–85, 1999.
- [132] M. Möllney, W. Wiechert, D. Kownatzki, and A. A. de Graaf, “Bidirectional reaction steps in metabolic networks: IV. Optimal design of isotopomer labeling experiments,” *Biotechnol Bioeng*, vol. 66, pp. 86–103, 1999.
- [133] H. M. Davey, C. L. Davey, A. M. Woodward, A. N. Edmonds, A. W. Lee, and D. B. Kell, “Oscillatory, stochastic and chaotic growth rate fluctuations in permissively controlled yeast cultures,” *Bio Systems*, vol. 39, pp. 43–61, Jan. 1996.
- [134] C. Winder and K. Lanthaler, *The use of continuous culture in systems biology investigations.*, vol. 500 of *Methods in systems biology*. Academic Press, in press.
- [135] G. A. Pope, D. A. MacKenzie, M. Defernez, M. A. M. M. Aroso, L. J. Fuller, F. A. Mellon, W. B. Dunn, M. Brown, R. Goodacre, D. B. Kell, M. E. Marvin, E. J. Louis, and I. N. Roberts, “Metabolic footprinting as a tool for discriminating between brewing yeasts,” *Yeast Chichester England*, vol. 24, no. 8, pp. 667–679, 2007.
- [136] S. O’Hagan, W. B. Dunn, M. Brown, J. D. Knowles, , and D. B. Kell, “Closed-loop, multiobjective optimization of analytical instrumentation: Gas chromatography/time-of-flight mass spectrometry of the metabolomes of human serum and of yeast fermentations,” *Analytical chemistry*, vol. 77, pp. 290–303, January 2005.
- [137] S. Schuster, “Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering,” *Trends in Biotechnology*, vol. 17, pp. 53–60, Feb. 1999.
- [138] T. L. Nissen, U. Schulze, J. Nielsen, and J. Villadsen, “Flux distributions in anaerobic, glucose-limited continuous cultures of *Saccharomyces cerevisiae.*,” *Microbiology (Reading, England)*, vol. 143 ( Pt 1, pp. 203–18, Jan. 1997.

GENERATING THERMODYNAMICALLY-CONSISTENT  
AND EXPERIMENTALLY FOUNDED KINETIC  
MODELS AT THE GENOME-SCALE -  
SUPPLEMENTARY INFORMATION

## Flux and concentrations contained in the model

Table A.1: Reaction Fluxes for the reduced network.

Metabolite	Flux (mM sec <sup>-1</sup> )
1,3-beta-glucan synthase	$6.31 \times 10^{-2}$
1,4-alpha-glucan branching enzyme	$2.88 \times 10^{-2}$
1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]imidazole-4-carboxamide is.	$4.17 \times 10^{-3}$
1-acyl-sn-glycerol-3-phosphate acyltransferase	$3.48 \times 10^{-4}$
2,5-diamino-6-ribitylamino-4(3H)-pyrimidinone 5-phosphate deaminase	$5.00 \times 10^{-5}$
2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5-phosphate reductase (NADPH)	$5.00 \times 10^{-5}$
2-aceto-2-hydroxybutanoate synthase	$9.52 \times 10^{-3}$
2-aminoadipate transaminase	$1.33 \times 10^{-2}$
2-deoxy-D-arabino-heptulosonate 7-phosphate synthetase	$1.33 \times 10^{-2}$
2-isopropylmalate hydratase	$1.39 \times 10^{-2}$
2-isopropylmalate synthase	$1.39 \times 10^{-2}$
2-methylcitrate dehydratase	$1.33 \times 10^{-2}$
2-oxo-4-methyl-3-carboxypentanoate decarboxylation	$1.39 \times 10^{-2}$
3,5-bisphosphate nucleotidase	$5.16 \times 10^{-3}$
3,4-dihydroxy-2-butanone-4-phosphate synthase	$9.99 \times 10^{-5}$

Continued on Next Page...

Table A.1 – Continued

Reaction	Flux(mmol/s)
3-dehydroquininate dehydratase	$1.33 \times 10^{-2}$
3-dehydroquininate synthase	$1.33 \times 10^{-2}$
3-dehydrosphinganine reductase	$2.31 \times 10^{-5}$
3-hydroxyacyl-CoA dehydrogenase (3-oxohexacosyl-CoA)	$2.73 \times 10^{-2}$
3-hydroxyacyl-CoA dehydrogenase (3-oxohexadecanoyl-CoA)	$2.73 \times 10^{-2}$
3-hydroxyacyl-CoA dehydrogenase (3-oxooctadecanoyl-CoA)	$2.73 \times 10^{-2}$
3-hydroxyacyl-CoA dehydrogenase (3-oxotetradecanoyl-CoA)	$2.73 \times 10^{-2}$
3-isopropylmalate dehydratase	$1.39 \times 10^{-2}$
3-isopropylmalate dehydrogenase	$1.39 \times 10^{-2}$
3-phosphoshikimate 1-carboxyvinyltransferase	$1.33 \times 10^{-2}$
5,10-methylenetetrahydrofolatereductase (NADPH)	$3.63 \times 10^{-3}$
acetoxy acid isomeroeductase	$2.82 \times 10^{-2}$
acetolactate synthase	$2.82 \times 10^{-2}$
acetyl-CoA C-acetyltransferase	$2.28 \times 10^{-3}$
acetyl-Coa carboxylase	$4.17 \times 10^{-4}$
acetyl-CoA hydrolase	$2.23 \times 10^{-1}$
acetyl-CoA synthetase	$1.57 \times 10^{-1}$
acetylglutamate kinase	$7.54 \times 10^{-3}$
acetylornithine transaminase	$7.54 \times 10^{-3}$
adenosine kinase	$8.55 \times 10^{-4}$
adenosylhomocysteinase	$8.55 \times 10^{-4}$
adenylate kinase	$4.16 \times 10^{-2}$
adenylate kinase (GTP)	$5.28 \times 10^{-2}$
adenylosuccinate lyase	$6.07 \times 10^{-3}$
adenylosuccinate synthase	$7.20 \times 10^{-3}$
adenylsuccinate lyase	$7.20 \times 10^{-3}$
adenylyl-sulfate kinase	$5.16 \times 10^{-3}$
alanine glyoxylate aminotransferase	$2.23 \times 10^{-2}$
alcohol dehydrogenase, reverse rxn (acetaldehyde -> ethanol)	$8.19 \times 10^{-1}$
aldehyde dehydrogenase (acetaldehyde, NADP)	$5.73 \times 10^{-2}$
alpha, alpha-trehalose-phosphate synthase (UDP-forming)	$1.30 \times 10^{-3}$
anthranilate phosphoribosyltransferase	$1.55 \times 10^{-3}$
anthranilate synthase	$1.55 \times 10^{-3}$
argininosuccinate lyase	$7.54 \times 10^{-3}$
argininosuccinate synthase	$7.54 \times 10^{-3}$
asparagine synthase (glutamine-hydrolysing)	$9.52 \times 10^{-3}$
aspartate carbamoyltransferase	$6.83 \times 10^{-3}$
aspartate kinase	$8.11 \times 10^{-2}$
aspartate transaminase	$1.28 \times 10^{-1}$
aspartate-semialdehyde dehydrogenase	$8.11 \times 10^{-2}$
ATP phosphoribosyltransferase	$4.17 \times 10^{-3}$
ATP synthase	3.38
ATPase, cytosolic	$4.17 \times 10^{-1}$

Continued on Next Page...

Table A.1 – Continued

Reaction	Flux(mmol/s)
bicarbonate formation	$2.61 \times 10^{-1}$
C-14 sterol reductase	$3.79 \times 10^{-4}$
C-3 sterol dehydrogenase	$3.76 \times 10^{-4}$
C-3 sterol dehydrogenase (4-methylzymosterol)	$3.76 \times 10^{-4}$
C-3 sterol keto reductase (4-methylzymosterol)	$3.76 \times 10^{-4}$
C-3 sterol keto reductase (zymosterol)	$3.76 \times 10^{-4}$
C-4 methyl sterol oxidase	$3.76 \times 10^{-4}$
C-4 methyl sterol oxidase	$3.76 \times 10^{-4}$
C-4 methyl sterol oxidase	$3.76 \times 10^{-4}$
C-4 sterol methyl oxidase (4,4-dimethylzymosterol)	$3.76 \times 10^{-4}$
C-8 sterol isomerase	$5.33 \times 10^{-6}$
C-s24 sterol reductase	$3.56 \times 10^{-4}$
carbamoyl-phosphate synthase (glutamine-hydrolysing)	$1.44 \times 10^{-2}$
catalase	$3.41 \times 10^{-3}$
CDP-diacylglycerol synthase	$3.04 \times 10^{-4}$
ceramide-1 hydroxylase (24C)	$2.31 \times 10^{-5}$
ceramide-1 synthase (24C)	$2.31 \times 10^{-5}$
cholestenol delta-isomerase, lumped reaction	$3.63 \times 10^{-4}$
chorismate mutase	$1.17 \times 10^{-2}$
chorismate synthase	$1.33 \times 10^{-2}$
cis-aconitate(3-) to isocitrate	$8.01 \times 10^{-2}$
citrate synthase	$8.01 \times 10^{-2}$
citrate to cis-aconitate(3-)	$8.01 \times 10^{-2}$
CTP synthase (NH <sub>3</sub> )	$2.78 \times 10^{-3}$
cystathionine beta-synthase	$2.38 \times 10^{-3}$
cystathionine g-lyase	$5.95 \times 10^{-3}$
cystathionine gamma-synthase	$3.57 \times 10^{-3}$
cytidylate kinase (CMP)	$2.47 \times 10^{-3}$
cytochrome P450 lanosterol 14-alpha-demethylase (NAD)	$3.79 \times 10^{-4}$
D-arabinose 1-dehydrogenase (NAD)	$2.73 \times 10^{-2}$
D-arabinose 1-dehydrogenase (NADP)	$2.73 \times 10^{-2}$
dCMP deaminase	$1.35 \times 10^{-4}$
deoxyadenylate kinase	$1.99 \times 10^{-4}$
deoxyguanylate kinase (dGMP:ATP)	$1.35 \times 10^{-4}$
diacylglycerol acyltransferase	$9.99 \times 10^{-5}$
diacylglycerol pyrophosphate phosphatase	$4.34 \times 10^{-5}$
dihydroorotic acid dehydrogenase	$6.83 \times 10^{-3}$
dihydrofolate reductase	$1.99 \times 10^{-4}$
dihydroorotase	$6.83 \times 10^{-3}$
dihydroxy-acid dehydratase (2,3-dihydroxy-3-methylbutanoate)	$2.82 \times 10^{-2}$
dihydroxy-acid dehydratase (2,3-dihydroxy-3-methylpentanoate)	$9.52 \times 10^{-3}$
dihydroxyacetone kinase	$4.53 \times 10^{-2}$
dimethylallyltranstransferase	$7.61 \times 10^{-4}$

Continued on Next Page...

Table A.1 – Continued

Reaction	Flux(mmol/s)
dolichyl-phosphate D-mannosyltransferase	$4.56 \times 10^{-2}$
dolichyl-phosphate-mannose-protein mannosyltransferase	$4.56 \times 10^{-2}$
enolase	1.23
fatty acid synthase (n-C10:0)	$1.16 \times 10^{-5}$
fatty acid synthase (n-C12:0)	$1.16 \times 10^{-5}$
fatty acid synthase (n-C14:0)	$1.16 \times 10^{-5}$
fatty acid synthase (n-C16:0)	$1.16 \times 10^{-5}$
fatty acid synthase (n-C18:0)	$1.16 \times 10^{-5}$
fatty acid synthase (n-C24:0), lumped reaction	$2.31 \times 10^{-5}$
fatty acyl-CoA synthase (n-C10:0CoA)	$3.47 \times 10^{-5}$
fatty acyl-CoA synthase (n-C8:0CoA), lumped reaction	$4.63 \times 10^{-5}$
fatty-acid-CoA ligase (n-C24:0)	$2.31 \times 10^{-5}$
fatty-acid-CoA ligase (octadecanoate)	$1.16 \times 10^{-5}$
fatty-acid-CoA ligase (octanoate)	$1.16 \times 10^{-5}$
fatty-acyl-CoA synthase (n-C12:0CoA)	$3.47 \times 10^{-5}$
fatty-acyl-CoA synthase (n-C14:0CoA)	$3.47 \times 10^{-5}$
fatty-acyl-CoA synthase (n-C16:0CoA)	$3.47 \times 10^{-5}$
fatty-acyl-CoA synthase (n-C18:0CoA)	$1.16 \times 10^{-5}$
formate-tetrahydrofolate ligase	$5.29 \times 10^{-4}$
fructose-bisphosphate aldolase	$6.23 \times 10^{-1}$
fumarase	$3.79 \times 10^{-2}$
fumarate reductase	$5.87 \times 10^{-2}$
geranyltranstransferase	$7.61 \times 10^{-4}$
glucokinase	$5.99 \times 10^{-1}$
glucose-6-phosphate isomerase	$5.99 \times 10^{-1}$
glucose-6-phosphate isomerase	$8.62 \times 10^{-2}$
glutamate 5-kinase	$7.14 \times 10^{-3}$
glutamate dehydrogenase (NADP)	$1.51 \times 10^{-1}$
glutamate synthase (NADH <sub>2</sub> )	$1.51 \times 10^{-1}$
glutamate-5-semialdehyde dehydrogenase	$7.14 \times 10^{-3}$
glutamine phosphoribosylidiphosphate amidotransferase	$6.07 \times 10^{-3}$
glutamine synthetase	$2.11 \times 10^{-1}$
glyceraldehyde-3-phosphate dehydrogenase	1.23
glycerol dehydrogenase (NADP-dependent)	$4.53 \times 10^{-2}$
glycerol-3-phosphatase	$4.53 \times 10^{-2}$
glycerol-3-phosphate dehydrogenase (fad)	$5.87 \times 10^{-2}$
glycerol-3-phosphate dehydrogenase (NAD)	$1.04 \times 10^{-1}$
glycerol-3-phosphate/dihydroxyacetone phosphate acyltransferase	$3.48 \times 10^{-4}$
glycine cleavage system	$3.79 \times 10^{-2}$
glycine hydroxymethyltransferase	$1.83 \times 10^{-2}$
glycogen (starch) synthase	$2.88 \times 10^{-2}$
GMP synthase	$3.04 \times 10^{-3}$
GTP cyclohydrolase II	$5.00 \times 10^{-5}$

Continued on Next Page...

Table A.1 – Continued

Reaction	Flux(mmol/s)
guanylate kinase (GMP:ATP)	$1.09 \times 10^{-4}$
guanylate kinase (GMP:dATP)	$9.96 \times 10^{-5}$
hexokinase (D-glucose:ATP)	$1.81 \times 10^{-1}$
histidinol dehydrogenase	$4.17 \times 10^{-3}$
histidinol-phosphatase	$4.17 \times 10^{-3}$
histidinol-phosphate transaminase	$4.17 \times 10^{-3}$
homoaccontinate hydratase	$1.33 \times 10^{-2}$
homocitrate synthase	$1.33 \times 10^{-2}$
homoisocitrate dehydrogenase	$1.33 \times 10^{-2}$
homoserine dehydrogenase (NADH)	$8.11 \times 10^{-2}$
homoserine kinase	$7.24 \times 10^{-2}$
homoserine O-trans-acetylase	$8.73 \times 10^{-3}$
hydroxymethylglutaryl CoA reductase	$2.28 \times 10^{-3}$
hydroxymethylglutaryl CoA synthase	$2.28 \times 10^{-3}$
Imidazole-glycerol-3-phosphate synthase	$4.17 \times 10^{-3}$
imidazoleglycerol-phosphate dehydratase	$4.17 \times 10^{-3}$
IMP cyclohydrolase	$1.02 \times 10^{-2}$
IMP dehydrogenase	$3.04 \times 10^{-3}$
indole-3-glycerol-phosphate synthase	$1.55 \times 10^{-3}$
inorganic diphosphatase	$4.16 \times 10^{-2}$
inositolphosphotransferase	$2.31 \times 10^{-5}$
IPC synthase	$2.31 \times 10^{-5}$
isocitrate dehydrogenase (NADP)	$5.77 \times 10^{-2}$
isocitrate lyase	$2.23 \times 10^{-2}$
isoleucine transaminase	$9.52 \times 10^{-3}$
isopentenyl-diphosphate D-isomerase	$7.61 \times 10^{-4}$
ketol-acid reductoisomerase (2-aceto-2-hydroxybutanoate)	$9.52 \times 10^{-3}$
L-alanine transaminase	$4.22 \times 10^{-2}$
L-aminoadipate-semialdehyde dehydrogenase (NADH)	$1.33 \times 10^{-2}$
L-glutamate 5-semialdehyde dehydratase	$7.14 \times 10^{-3}$
L-hydroxyproline dehydrogenase (NADP)	$2.73 \times 10^{-2}$
L-hydroxyproline reductase (NAD)	$2.73 \times 10^{-2}$
L-threonine deaminase	$3.57 \times 10^{-3}$
lanosterol synthase	$3.80 \times 10^{-4}$
leucine transaminase	$1.39 \times 10^{-2}$
malate dehydrogenase	$3.79 \times 10^{-2}$
mannose-1-phosphate guanylyltransferase	$4.56 \times 10^{-2}$
mannose-6-phosphate isomerase	$4.56 \times 10^{-2}$
methenyltetrahydroficate cyclohydrolase	$1.58 \times 10^{-2}$
methionine adenosyltransferase	$8.55 \times 10^{-4}$
methionine synthase	$3.63 \times 10^{-3}$
methylenetetrahydrofolate dehydrogenase (NADP)	$1.58 \times 10^{-2}$
mevalonate kinase (ctp)	$2.28 \times 10^{-3}$

Continued on Next Page...

Table A.1 – Continued

Reaction	Flux(mmol/s)
mevalonate pyrophosphate decarboxylase	$2.28 \times 10^{-3}$
microsomal beta-keto-reductase	$2.73 \times 10^{-2}$
MIPC synthase	$2.31 \times 10^{-5}$
myo-inositol 1-phosphatase	$8.50 \times 10^{-5}$
myo-inositol-1-phosphate synthase	$1.31 \times 10^{-4}$
N-acetyl-g-glutamyl-phosphate reductase	$7.54 \times 10^{-3}$
non-enzymatic reaction	$1.33 \times 10^{-2}$
nucleoside-diphosphate kinase (ATP:CDP)	$1.89 \times 10^{-4}$
nucleoside-diphosphate kinase (ATP:UDP)	$9.59 \times 10^{-2}$
O-acetylhomoserine (thiol)-lyase	$5.16 \times 10^{-3}$
ornithine carbamoyltransferase	$7.54 \times 10^{-3}$
ornithine transacetylase	$7.54 \times 10^{-3}$
orotate phosphoribosyltransferase	$6.83 \times 10^{-3}$
orotidine-5-phosphate decarboxylase	$6.83 \times 10^{-3}$
phenylalanine transaminase	$6.35 \times 10^{-3}$
phosphatidylethanolamine methyltransferase	$1.60 \times 10^{-4}$
phosphatidylinositol synthase	$8.50 \times 10^{-5}$
phosphatidylserine decarboxylase	$1.99 \times 10^{-4}$
phosphatidylserine synthase	$2.20 \times 10^{-4}$
phosphoadenylyl-sulfate reductase (thioredoxin)	$5.16 \times 10^{-3}$
phosphofructokinase	$6.23 \times 10^{-1}$
phosphoglucomutase	$9.31 \times 10^{-2}$
phosphoglycerate kinase	1.23
phosphoglycerate mutase	1.23
phospholipid methyltransferase	$1.60 \times 10^{-4}$
phospholipid methyltransferase	$1.60 \times 10^{-4}$
phosphomannomutase	$4.56 \times 10^{-2}$
phosphomevalonate kinase	$2.28 \times 10^{-3}$
phosphoribosyl-AMP cyclohydrolase	$4.17 \times 10^{-3}$
phosphoribosyl-ATP pyrophosphatase	$4.17 \times 10^{-3}$
phosphoribosylaminoimidazole carboxylase	$6.07 \times 10^{-3}$
phosphoribosylaminoimidazole synthase	$6.07 \times 10^{-3}$
phosphoribosylaminoimidazolecarboxamide formyltransferase	$1.02 \times 10^{-2}$
phosphoribosylaminoimidazolesuccinocarboxamide synthase	$6.07 \times 10^{-3}$
phosphoribosylanthranilate isomerase	$1.55 \times 10^{-3}$
phosphoribosylformylglycinamide synthase	$6.07 \times 10^{-3}$
phosphoribosylglycinamide formyltransferase	$6.07 \times 10^{-3}$
phosphoribosylglycinamide synthase	$6.07 \times 10^{-3}$
phosphoribosylpyrophosphate synthetase	$1.86 \times 10^{-2}$
prephenate dehydratase	$6.35 \times 10^{-3}$

Continued on Next Page...

Table A.1 – Continued

Reaction	Flux(mmol/s)
prephenate dehydrogenase (NADP)	$5.36 \times 10^{-3}$
pyrimidine phosphatase	$5.00 \times 10^{-5}$
pyrroline-5-carboxylate reductase	$7.14 \times 10^{-3}$
pyruvate carboxylase	$2.46 \times 10^{-1}$
pyruvate decarboxylase	$8.18 \times 10^{-1}$
pyruvate dehydrogenase	$5.73 \times 10^{-2}$
pyruvate kinase	1.21
riboflavin synthase	$9.99 \times 10^{-5}$
riboflavin synthase	$5.00 \times 10^{-5}$
ribonucleoside-diphosphate reductase	$9.96 \times 10^{-5}$
ribonucleoside-diphosphate reductase (GDP)	$1.35 \times 10^{-4}$
ribonucleoside-diphosphate reductase (UDP)	$3.34 \times 10^{-4}$
ribonucleoside-triphosphate reductase (ATP)	$9.96 \times 10^{-5}$
ribose-5-phosphate isomerase	$1.68 \times 10^{-2}$
ribulose 5-phosphate 3-epimerase	$1.69 \times 10^{-2}$
S-adenosyl-methionine delta-24-sterol-c-methyltransferase	$1.17 \times 10^{-5}$
saccharopine dehydrogenase (NAD, L-lysine forming)	$1.33 \times 10^{-2}$
saccharopine dehydrogenase (NADP, L-glutamate forming)	$1.33 \times 10^{-2}$
serine C-palmitoyltransferase	$2.31 \times 10^{-5}$
shikimate dehydrogenase	$1.33 \times 10^{-2}$
shikimate kinase	$1.33 \times 10^{-2}$
squalene epoxidase (NAD)	$3.80 \times 10^{-4}$
squalene synthase	$3.80 \times 10^{-4}$
steryl ester hydrolase	$4.51 \times 10^{-5}$
succinate-CoA ligase (ADP-forming)	$7.96 \times 10^{-4}$
sulfate adenylyltransferase (ADP)	$5.16 \times 10^{-3}$
sulfite reductase (NADPH <sub>2</sub> )	$5.16 \times 10^{-3}$
thioredoxin reductase (NADPH)	$5.83 \times 10^{-3}$
threonine aldolase	$5.79 \times 10^{-2}$
threonine synthase	$7.24 \times 10^{-2}$
thymidylate synthase	$1.99 \times 10^{-4}$
transaldolase	$1.82 \times 10^{-3}$
transketolase	$1.82 \times 10^{-3}$
transketolase	$1.51 \times 10^{-2}$
trehalose-phosphatase	$1.30 \times 10^{-3}$
triacylglycerol lipase	$5.65 \times 10^{-5}$
triose-phosphate isomerase	$6.23 \times 10^{-1}$
tryptophan synthase (indoleglycerol phosphate)	$1.55 \times 10^{-3}$
tyrosine transaminase	$5.36 \times 10^{-3}$
UMP kinase	$3.11 \times 10^{-3}$
uridylyltransferase (dUMP)	$3.34 \times 10^{-4}$
UTP-glucose-1-phosphate uridylyltransferase	$9.31 \times 10^{-2}$
valine transaminase	$1.43 \times 10^{-2}$

Continued on Next Page...

Table A.1 – Continued

Reaction	Flux(mmol/s)
ammonia transport	$3.17 \times 10^{-1}$
CO2 transport	$7.93 \times 10^{-1}$
ethanol transport	$8.19 \times 10^{-1}$
glucose transport	$7.80 \times 10^{-1}$
O2 transport	$7.57 \times 10^{-3}$
phosphate transport	$1.16 \times 10^{-2}$
succinate transport	$8.02 \times 10^{-2}$
sulfate uniport	$6.27 \times 10^{-3}$
isa acyl-CoA	$7.96 \times 10^{-4}$
biomass production	$5.55 \times 10^{-2}$
growth	$5.55 \times 10^{-2}$
lipid production	$5.55 \times 10^{-2}$

Table A.2: Metabolite Concentrations

Metabolite	Concentration (mM)
(1->3)-beta-D-glucan [intracellular]	$1.76 \times 10^{-1}$
(2R,3R)-2,3-dihydroxy-3-methylpentanoate [intracellular]	$1.76 \times 10^{-1}$
(2R,3S)-3-isopropylmalate(2-) [intracellular]	$1.76 \times 10^{-1}$
(2S)-2-[5-amino-1-(5-phospho-beta-D-ribose)imidazole-4-carboxamido]succinic acid [intracellular]	$1.76 \times 10^{-1}$
(2S)-2-isopropyl-3-oxosuccinate(2-) [intracellular]	$1.76 \times 10^{-1}$
(6R)-5,10-methenyltetrahydrofolic acid [intracellular]	$1.76 \times 10^{-1}$
(N(omega)-L-arginino)succinic acid [intracellular]	$1.76 \times 10^{-1}$
(R)-2,3-dihydroxy-3-methylbutanoate [intracellular]	$1.76 \times 10^{-1}$
(R)-5-diphosphomevalonic acid [intracellular]	$1.76 \times 10^{-1}$
(R)-5-phosphomevalonic acid [intracellular]	$1.76 \times 10^{-1}$
(R)-mevalonate [intracellular]	$1.76 \times 10^{-1}$
(S)-2,3-epoxysqualene [intracellular]	$1.76 \times 10^{-1}$
(S)-2-acetyl-2-hydroxybutanoate [intracellular]	$1.76 \times 10^{-1}$
(S)-3-hydroxyhexacosanoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
(S)-3-hydroxypalmitoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
(S)-3-hydroxytetradecanoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
(S)-3-methyl-2-oxopentanoate [intracellular]	$1.76 \times 10^{-1}$
(S)-dihydroorotate [intracellular]	$1.76 \times 10^{-1}$
(S)-malate(2-) [intracellular]	$1.76 \times 10^{-1}$
1-(2-carboxyphenylamino)-1-deoxy-D-ribose 5-phosphate [intracellular]	$1.76 \times 10^{-1}$
1-(5-phospho-D-ribose)-5-[(5-phospho-D-riboseylamino)methylideneamino]-imidazole-4-carboxamide [intracellular]	$1.76 \times 10^{-1}$
1-(5-phosphoribosyl)-5-AMP [intracellular]	$1.76 \times 10^{-1}$
1-acyl-sn-glycerol 3-phosphate [intracellular]	$1.76 \times 10^{-1}$
1-C-(indol-3-yl)glycerol 3-phosphate [intracellular]	$1.76 \times 10^{-1}$

Continued on Next Page...

Table A.2 – Continued

Metabolite	Concentration (mM)
1-phosphatidyl-1D-myo-inositol [intracellular]	$1.76 \times 10^{-1}$
1-pyrroline-3-hydroxy-5-carboxylic acid [intracellular]	$1.76 \times 10^{-1}$
1-pyrroline-5-carboxylate [intracellular]	$1.76 \times 10^{-1}$
10-formyltetrahydrofolic acid [intracellular]	$1.76 \times 10^{-1}$
14-demethylsterol [intracellular]	$1.76 \times 10^{-1}$
1D-myo-inositol 1-phosphate [intracellular]	$1.76 \times 10^{-1}$
2,5-diamino-4-hydroxy-6-(5-phosphoribosylamino)pyrimidine [intracellular]	$1.76 \times 10^{-1}$
2,5-diamino-6-(5-phosphono)ribitylamino-4(3H)-pyrimidinone [intracellular]	$1.76 \times 10^{-1}$
2-acetamido-5-oxopentanoate [intracellular]	$1.76 \times 10^{-1}$
2-acetyllactic acid [intracellular]	$1.76 \times 10^{-1}$
2-formamido-N(1)-(5-phospho-D-ribosyl)acetamide [intracellular]	$1.76 \times 10^{-1}$
2-hydroxy-3-oxobutyl phosphate [intracellular]	$1.76 \times 10^{-1}$
2-isopropylmalate(2-) [intracellular]	$1.76 \times 10^{-1}$
2-isopropylmaleic acid [intracellular]	$1.76 \times 10^{-1}$
2-oxaloglutaric acid [intracellular]	$1.76 \times 10^{-1}$
2-oxoadipic acid [intracellular]	$1.76 \times 10^{-1}$
2-oxobutanoate [intracellular]	$1.76 \times 10^{-1}$
2-oxoglutarate [intracellular]	$1.76 \times 10^{-1}$
2-phospho-D-glyceric acid [intracellular]	$3.70 \times 10^{-2}$
2-trans,6-trans-farnesyl diphosphate [intracellular]	$1.76 \times 10^{-1}$
3-phospho-5-adenylyl sulfate [intracellular]	$1.76 \times 10^{-1}$
3-(4-hydroxyphenyl)pyruvate [intracellular]	$1.76 \times 10^{-1}$
3-(imidazol-4-yl)-2-oxopropyl dihydrogen phosphate [intracellular]	$1.76 \times 10^{-1}$
3-dehydro-4-methylzymosterol [intracellular]	$1.76 \times 10^{-1}$
3-dehydroquinone [intracellular]	$1.76 \times 10^{-1}$
3-dehydroshikimate [intracellular]	$1.76 \times 10^{-1}$
3-dehydrospinganine [intracellular]	$1.76 \times 10^{-1}$
3-hydroxy-3-methylglutaryl-CoA [intracellular]	$1.76 \times 10^{-1}$
3-hydroxyoctadecanoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
3-methyl-2-oxobutanoate [intracellular]	$1.76 \times 10^{-1}$
3-oxohexacosanoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
3-oxooctadecanoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
3-oxopalmitoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
3-oxotetradecanoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
3-phospho-D-glyceric acid [intracellular]	$2.78 \times 10^{-1}$
3-phospho-D-glyceroyl dihydrogen phosphate [intracellular]	$2.75 \times 10^{-4}$
3-phosphoshikimic acid [intracellular]	$1.76 \times 10^{-1}$
4,4-dimethyl-5alpha-cholesta-8,14,24-trien-3beta-ol [intracellular]	$1.76 \times 10^{-1}$
4-methyl-2-oxopentanoate [intracellular]	$1.76 \times 10^{-1}$
4-phospho-L-aspartate [intracellular]	$1.76 \times 10^{-1}$
4alpha-methylzymosterol [intracellular]	$1.76 \times 10^{-1}$
4beta-methylzymosterol-4alpha-carboxylic acid [intracellular]	$1.76 \times 10^{-1}$
5-adenylyl sulfate [intracellular]	$1.76 \times 10^{-1}$

Continued on Next Page...

Table A.2 – Continued

Metabolite	Concentration (mM)
5-xanthylic acid [intracellular]	$1.76 \times 10^{-1}$
5,10-methylenetetrahydrofolate(2-) [intracellular]	$1.76 \times 10^{-1}$
5,6,7,8-tetrahydrofolic acid [intracellular]	$1.76 \times 10^{-1}$
5-[(5-phospho-1-deoxy-D-ribulos-1-ylamino)methyl]- 1-(5-phospho-D-ribosyl)imidazole-4-carboxamide [intracellular]	$1.76 \times 10^{-1}$
5-amino-1-(5-phospho-D-ribosyl)imidazole [intracellular]	$1.76 \times 10^{-1}$
5-amino-1-(5-phospho-D-ribosyl)imidazole-4-carboxamide [intracellular]	$1.76 \times 10^{-1}$
5-amino-1-(5-phospho-D-ribosyl)imidazole-4-carboxylic acid [intracellular]	$1.76 \times 10^{-1}$
5-amino-6-(5-phosphoribitylamino)uracil [intracellular]	$1.76 \times 10^{-1}$
5-amino-6-(D-ribitylamino)uracil [intracellular]	$1.76 \times 10^{-1}$
5-formamido-1-(5-phospho-D-ribosyl)imidazole-4-carboxamide [intracellular]	$1.76 \times 10^{-1}$
5-methyltetrahydrofolate(2-) [intracellular]	$1.76 \times 10^{-1}$
5-O-(1-carboxyvinyl)-3-phosphoshikimic acid [intracellular]	$1.76 \times 10^{-1}$
5-O-phosphono-alpha-D-ribofuranosyl diphosphate [intracellular]	$1.76 \times 10^{-1}$
5-phospho-beta-D-ribosylamine [intracellular]	$1.76 \times 10^{-1}$
5-phosphoribosyl-ATP [intracellular]	$1.76 \times 10^{-1}$
6,7-dimethyl-8-(1-D-ribityl)lumazine [intracellular]	$1.76 \times 10^{-1}$
7-phospho-2-dehydro-3-deoxy-D-arabino-heptonic acid [intracellular]	$1.76 \times 10^{-1}$
acetaldehyde [intracellular]	$1.70 \times 10^{-1}$
acetate [intracellular]	$1.76 \times 10^{-1}$
acetoacetyl-CoA [intracellular]	$1.76 \times 10^{-1}$
acetyl-CoA [intracellular]	$1.76 \times 10^{-1}$
acyl-CoA [intracellular]	$1.76 \times 10^{-1}$
adenosine [intracellular]	$1.76 \times 10^{-1}$
adenosine 3,5-bismonophosphate [intracellular]	$1.76 \times 10^{-1}$
ADP [intracellular]	1.63
aldehydo-D-glucose 6-phosphate [intracellular]	$1.76 \times 10^{-1}$
alpha,alpha-trehalose [intracellular]	$1.76 \times 10^{-1}$
alpha,alpha-trehalose 6-phosphate [intracellular]	$1.76 \times 10^{-1}$
alpha-D-ribose 5-phosphate [intracellular]	$1.76 \times 10^{-1}$
ammonium [intracellular]	$3.80 \times 10^{+1}$
AMP [intracellular]	$7.96 \times 10^{-1}$
amylose [intracellular]	$1.76 \times 10^{-1}$
anthranilate [intracellular]	$1.76 \times 10^{-1}$
ATP [intracellular]	1.13
beta-D-glucose 6-phosphate [intracellular]	1.02
bicarbonate [intracellular]	$1.76 \times 10^{-1}$
biomass [intracellular]	$1.76 \times 10^{-1}$
but-1-ene-1,2,4-tricarboxylic acid [intracellular]	$1.76 \times 10^{-1}$
carbamoyl phosphate [intracellular]	$1.76 \times 10^{-1}$
carbon dioxide [intracellular]	1.00
CDP [intracellular]	$1.76 \times 10^{-1}$
CDP-diacylglycerol [intracellular]	$1.76 \times 10^{-1}$

Continued on Next Page...

Table A.2 – Continued

Metabolite	Concentration (mM)
chorismate(2-) [intracellular]	$1.76 \times 10^{-1}$
cis-aconitate(3-) [intracellular]	$1.76 \times 10^{-1}$
citrate(3-) [intracellular]	$1.76 \times 10^{-1}$
CMP [intracellular]	$1.76 \times 10^{-1}$
coenzyme A [intracellular]	$1.76 \times 10^{-1}$
CTP [intracellular]	$1.76 \times 10^{-1}$
D-arabinono-1,4-lactone [intracellular]	$1.76 \times 10^{-1}$
D-arabinose [intracellular]	$1.76 \times 10^{-1}$
D-erythro-1-(imidazol-4-yl)glycerol 3-phosphate [intracellular]	$1.76 \times 10^{-1}$
D-erythrose 4-phosphate(2-) [intracellular]	$1.76 \times 10^{-1}$
D-fructose 1,6-bisphosphate [intracellular]	2.82
D-fructose 6-phosphate [intracellular]	$1.12 \times 10^{-1}$
D-glucose [intracellular]	$9.06 \times 10^{-2}$
D-glucose 1-phosphate [intracellular]	$1.76 \times 10^{-1}$
D-mannose 1-phosphate [intracellular]	$1.76 \times 10^{-1}$
D-mannose 6-phosphate [intracellular]	$1.76 \times 10^{-1}$
D-ribulose 5-phosphate [intracellular]	$1.76 \times 10^{-1}$
D-xylulose 5-phosphate [intracellular]	$1.76 \times 10^{-1}$
dADP [intracellular]	$1.76 \times 10^{-1}$
dAMP [intracellular]	$1.76 \times 10^{-1}$
dATP [intracellular]	$1.76 \times 10^{-1}$
dCMP [intracellular]	$1.76 \times 10^{-1}$
decanoate [intracellular]	$1.76 \times 10^{-1}$
decanoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
dGDP [intracellular]	$1.76 \times 10^{-1}$
dGMP [intracellular]	$1.76 \times 10^{-1}$
diglyceride [intracellular]	$1.76 \times 10^{-1}$
dihydrofolic acid [intracellular]	$1.76 \times 10^{-1}$
diphosphate [intracellular]	$1.76 \times 10^{-1}$
dolichyl D-mannosyl phosphate [intracellular]	$1.76 \times 10^{-1}$
dolichyl phosphate [intracellular]	$1.76 \times 10^{-1}$
dTMP [intracellular]	$1.76 \times 10^{-1}$
dUDP [intracellular]	$1.76 \times 10^{-1}$
dUMP [intracellular]	$1.76 \times 10^{-1}$
episterol [intracellular]	$1.76 \times 10^{-1}$
ergosta-5,7,22,24(28)-tetraen-3beta-ol [intracellular]	$1.76 \times 10^{-1}$
ergosterol [intracellular]	$1.76 \times 10^{-1}$
ergosterol ester [intracellular]	$1.76 \times 10^{-1}$
ethanol [intracellular]	$5.00 \times 10^{+1}$
FAD [intracellular]	$1.76 \times 10^{-1}$
FADH2 [intracellular]	$1.76 \times 10^{-1}$
fatty acid [intracellular]	$1.76 \times 10^{-1}$
fecosterol [intracellular]	$1.76 \times 10^{-1}$

Continued on Next Page...

Table A.2 – Continued

Metabolite	Concentration (mM)
formate [intracellular]	$1.76 \times 10^{-1}$
fumarate(2-) [intracellular]	$1.76 \times 10^{-1}$
GDP [intracellular]	$1.76 \times 10^{-1}$
GDP-alpha-D-mannose [intracellular]	$1.76 \times 10^{-1}$
geranyl diphosphate [intracellular]	$1.76 \times 10^{-1}$
glyceraldehyde 3-phosphate [intracellular]	$6.90 \times 10^{-2}$
glycerol [intracellular]	2.27
glycerone [intracellular]	$1.76 \times 10^{-1}$
glycerone phosphate [intracellular]	$5.90 \times 10^{-1}$
glycine [intracellular]	$1.76 \times 10^{-1}$
glycogen [intracellular]	$1.76 \times 10^{-1}$
glyoxylate [intracellular]	$1.76 \times 10^{-1}$
GMP [intracellular]	$1.76 \times 10^{-1}$
GTP [intracellular]	$1.76 \times 10^{-1}$
homocitrate(3-) [intracellular]	$1.76 \times 10^{-1}$
homoisocitrate(3-) [intracellular]	$1.76 \times 10^{-1}$
hydrogen peroxide [intracellular]	$1.76 \times 10^{-1}$
hydrogen sulfide [intracellular]	$1.76 \times 10^{-1}$
IMP [intracellular]	$1.76 \times 10^{-1}$
inositol phosphomannosylinositol phosphoceramide [intracellular]	$1.76 \times 10^{-1}$
inositol-P-ceramide B [intracellular]	$1.76 \times 10^{-1}$
isocitrate(3-) [intracellular]	$1.76 \times 10^{-1}$
isopentenyl diphosphate [intracellular]	$1.76 \times 10^{-1}$
keto-phenylpyruvate [intracellular]	$1.76 \times 10^{-1}$
L-2-aminoadipate(2-) [intracellular]	$1.76 \times 10^{-1}$
L-alanine [intracellular]	$1.76 \times 10^{-1}$
L-allysine [intracellular]	$1.76 \times 10^{-1}$
L-arginine [intracellular]	$1.76 \times 10^{-1}$
L-asparagine [intracellular]	$1.76 \times 10^{-1}$
L-aspartate [intracellular]	$1.76 \times 10^{-1}$
L-aspartate 4-semialdehyde [intracellular]	$1.76 \times 10^{-1}$
L-citrulline [intracellular]	$1.76 \times 10^{-1}$
L-cystathionine [intracellular]	$1.76 \times 10^{-1}$
L-cysteine [intracellular]	$1.76 \times 10^{-1}$
L-gamma-glutamyl phosphate [intracellular]	$1.76 \times 10^{-1}$
L-glutamate [intracellular]	$1.76 \times 10^{-1}$
L-glutamic 5-semialdehyde [intracellular]	$1.76 \times 10^{-1}$
L-glutamine [intracellular]	$1.76 \times 10^{-1}$
L-histidine [intracellular]	$1.76 \times 10^{-1}$
L-histidinol [intracellular]	$1.76 \times 10^{-1}$
L-histidinol phosphate [intracellular]	$1.76 \times 10^{-1}$
L-homocysteine [intracellular]	$1.76 \times 10^{-1}$
L-homoserine [intracellular]	$1.76 \times 10^{-1}$

Continued on Next Page...

Table A.2 – Continued

Metabolite	Concentration (mM)
L-isoleucine [intracellular]	$1.76 \times 10^{-1}$
L-leucine [intracellular]	1.00
L-lysine [intracellular]	1.00
L-methionine [intracellular]	$1.76 \times 10^{-1}$
L-phenylalanine [intracellular]	$1.76 \times 10^{-1}$
L-proline [intracellular]	$1.76 \times 10^{-1}$
L-saccharopine [intracellular]	$1.76 \times 10^{-1}$
L-serine [intracellular]	1.00
L-threonine [intracellular]	1.00
L-tryptophan [intracellular]	1.00
L-tyrosine [intracellular]	$1.76 \times 10^{-1}$
L-valine [intracellular]	1.00
lanosterol [intracellular]	$1.76 \times 10^{-1}$
laurate [intracellular]	$1.76 \times 10^{-1}$
lauroyl-CoA [intracellular]	$1.76 \times 10^{-1}$
lignocerate [intracellular]	$1.76 \times 10^{-1}$
lipid [intracellular]	$1.76 \times 10^{-1}$
malonyl-CoA [intracellular]	$1.76 \times 10^{-1}$
mannan [intracellular]	$1.76 \times 10^{-1}$
mannosylinositol phosphorylceramide [intracellular]	$1.76 \times 10^{-1}$
myo-inositol [intracellular]	$1.76 \times 10^{-1}$
myristate [intracellular]	$1.76 \times 10^{-1}$
myristoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
N(1)-(5-phospho-D-ribosyl)glycinamide [intracellular]	$1.76 \times 10^{-1}$
N(2)-acetyl-L-ornithine [intracellular]	$1.76 \times 10^{-1}$
N(2)-formyl-N(1)-(5-phospho-D-ribosyl)glycinamide [intracellular]	$1.76 \times 10^{-1}$
N(6)-(1,2-dicarboxyethyl)-AMP [intracellular]	$1.76 \times 10^{-1}$
N-(24-hydroxytetraacosanyl)sphinganine [intracellular]	$1.76 \times 10^{-1}$
N-(5-phospho-beta-D-ribosyl)anthranilate [intracellular]	$1.76 \times 10^{-1}$
N-acetyl-L-gamma-glutamyl phosphate [intracellular]	$1.76 \times 10^{-1}$
N-acetyl-L-glutamate(2-) [intracellular]	$1.76 \times 10^{-1}$
N-carbamoyl-L-aspartate [intracellular]	$1.76 \times 10^{-1}$
N-tetraacosanyl sphinganine [intracellular]	$1.76 \times 10^{-1}$
NAD(+) [intracellular]	1.50
NADH [intracellular]	$8.61 \times 10^{-2}$
NADP(+) [intracellular]	$1.76 \times 10^{-1}$
NADPH [intracellular]	$1.76 \times 10^{-1}$
O-acetyl-L-homoserine [intracellular]	$1.76 \times 10^{-1}$
O-phospho-L-homoserine [intracellular]	$1.76 \times 10^{-1}$
octanoate [intracellular]	$1.76 \times 10^{-1}$
octanoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
ornithine [intracellular]	$1.76 \times 10^{-1}$
orotate [intracellular]	$1.76 \times 10^{-1}$

Continued on Next Page...

Table A.2 – Continued

Metabolite	Concentration (mM)
orotidine 5-(dihydrogen phosphate) [intracellular]	$1.76 \times 10^{-1}$
oxaloacetate(2-) [intracellular]	$1.76 \times 10^{-1}$
oxygen [intracellular]	$1.76 \times 10^{-1}$
palmitate [intracellular]	$1.76 \times 10^{-1}$
palmitoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
phosphate [intracellular]	$1.76 \times 10^{-1}$
phosphatidate [intracellular]	$1.76 \times 10^{-1}$
phosphatidyl-L-serine [intracellular]	$1.76 \times 10^{-1}$
phosphatidyl-N,N-dimethylethanolamine [intracellular]	$1.76 \times 10^{-1}$
phosphatidyl-N-methylethanolamine [intracellular]	$1.76 \times 10^{-1}$
phosphatidylcholine [intracellular]	$1.76 \times 10^{-1}$
phosphatidylethanolamine [intracellular]	$1.76 \times 10^{-1}$
phosphoenolpyruvate [intracellular]	$3.02 \times 10^{-2}$
prenyl diphosphate [intracellular]	$1.76 \times 10^{-1}$
prephenate(2-) [intracellular]	$1.76 \times 10^{-1}$
pyruvate [intracellular]	8.36
riboflavin [intracellular]	$1.76 \times 10^{-1}$
S-adenosyl-L-homocysteine [intracellular]	$1.76 \times 10^{-1}$
S-adenosyl-L-methionine [intracellular]	$1.76 \times 10^{-1}$
sedoheptulose 7-phosphate [intracellular]	$1.76 \times 10^{-1}$
shikimate [intracellular]	$1.76 \times 10^{-1}$
sn-glycerol 3-phosphate [intracellular]	$4.57 \times 10^{-1}$
sphinganine [intracellular]	$1.76 \times 10^{-1}$
squalene [intracellular]	$1.76 \times 10^{-1}$
stearate [intracellular]	$1.76 \times 10^{-1}$
stearoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
succinate(2-) [intracellular]	$1.76 \times 10^{-1}$
succinyl-CoA [intracellular]	$1.76 \times 10^{-1}$
sulphate [intracellular]	$1.76 \times 10^{-1}$
sulphite [intracellular]	$1.76 \times 10^{-1}$
tetracosanoyl-CoA [intracellular]	$1.76 \times 10^{-1}$
trans-4-hydroxy-L-proline [intracellular]	$1.76 \times 10^{-1}$
triglyceride [intracellular]	$1.76 \times 10^{-1}$
UDP [intracellular]	$1.76 \times 10^{-1}$
UDP-D-glucose [intracellular]	$1.76 \times 10^{-1}$
UMP [intracellular]	$1.76 \times 10^{-1}$
UTP [intracellular]	$1.76 \times 10^{-1}$
zymosterol [intracellular]	$1.76 \times 10^{-1}$
zymosterol intermediate 1a [intracellular]	$1.76 \times 10^{-1}$
zymosterol intermediate 1b [intracellular]	$1.76 \times 10^{-1}$
zymosterol intermediate 1c [intracellular]	$1.76 \times 10^{-1}$
zymosterol intermediate 2 [intracellular]	$1.76 \times 10^{-1}$
thioredoxin disulfide [intracellular]	$1.76 \times 10^{-1}$

Continued on Next Page...

Table A.2 – Continued

<b>Metabolite</b>	<b>Concentration (mM)</b>
thioredoxin dithiol [intracellular]	$1.76 \times 10^{-1}$
H <sup>+</sup> [intracellular]	$1.11 \times 10^{+1}$
water [intracellular]	$1.11 \times 10^{+1}$
ammonium [extracellular]	$1.11 \times 10^{+1}$
biomass [extracellular]	$1.11 \times 10^{+1}$
carbon dioxide [extracellular]	$1.00 \times 10^{-1}$
D-glucose [extracellular]	$1.11 \times 10^{+1}$
ethanol [extracellular]	$1.11 \times 10^{+1}$
H <sup>+</sup> [extracellular]	$1.11 \times 10^{+1}$
oxygen [extracellular]	$1.11 \times 10^{+1}$
phosphate [extracellular]	$1.11 \times 10^{+1}$
succinate(2-) [extracellular]	1.00
sulphate [extracellular]	$4.22 \times 10^{+1}$

NEW INSIGHTS FROM OLD DATA: EMERGENT  
 PROPERTIES FROM KINETIC TRANSFORMATION OF  
 THE CONSENSUS YEAST NETWORK BUILT USING  
 NON-HOMOGENOUS DATA SOURCES -  
 SUPPLEMENTARY INFORMATION

## Flux and concentrations contained in the model

Table B.1: Reaction Fluxes for the reduced network

Metabolite	Flux (mM min <sup>-1</sup> )
1,3-beta-glucan synthase	$6.09 \times 10^{-3}$
1,4-alpha-glucan branching enzyme	$2.78 \times 10^{-3}$
1-acyl-sn-glycerol-3-phosphate acyltransferase	$3.36 \times 10^{-5}$
2,5-diamino-6-ribitylamino-4(3H)-pyrimidinone 5-phosphate deaminase	$4.82 \times 10^{-6}$
2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5-phosphate reductase (NADPH)	$4.82 \times 10^{-6}$
2-aceto-2-hydroxybutanoate synthase	$9.19 \times 10^{-4}$
2-aminoadipate transaminase	$1.28 \times 10^{-3}$
2-deoxy-D-arabino-heptulosonate 7-phosphate synthetase	$1.28 \times 10^{-3}$
2-methylcitrate dehydratase	$1.28 \times 10^{-3}$
3,5-bisphosphate nucleotidase	$4.98 \times 10^{-4}$
3,4-dihydroxy-2-butanone-4-phosphate synthase	$9.65 \times 10^{-6}$
3-dehydroquinate dehydratase	$1.28 \times 10^{-3}$
3-dehydroquinate synthase	$1.28 \times 10^{-3}$

Continued on Next Page...

Table B.1 – Continued

Reaction	Flux(mmol/s)
3-dehydrosphinganine reductase	$2.23 \times 10^{-6}$
3-hydroxyacyl-CoA dehydrogenase (3-oxohexacosyl-CoA)	$3.53 \times 10^{-3}$
3-hydroxyacyl-CoA dehydrogenase (3-oxohexadecanoyl-CoA)	$3.53 \times 10^{-3}$
3-hydroxyacyl-CoA dehydrogenase (3-oxooctadecanoyl-CoA)	$3.53 \times 10^{-3}$
3-hydroxyacyl-CoA dehydrogenase (3-oxotetradecanoyl-CoA)	$3.53 \times 10^{-3}$
3-phosphoshikimate 1-carboxyvinyltransferase	$1.28 \times 10^{-3}$
5,10-methylenetetrahydrofolatereductase (NADPH)	$3.51 \times 10^{-4}$
acetoxy acid isomeroeductase	$1.38 \times 10^{-3}$
acetolactate synthase	$1.38 \times 10^{-3}$
acetyl-CoA C-acetyltransferase	$2.20 \times 10^{-4}$
acetyl-Coa carboxylase	$4.02 \times 10^{-5}$
acetyl-CoA hydrolase	$1.43 \times 10^{-2}$
acetyl-CoA synthetase	$9.75 \times 10^{-3}$
acetylglutamate kinase	$7.28 \times 10^{-4}$
acetylornithine transaminase	$7.28 \times 10^{-4}$
adenosine kinase	$8.25 \times 10^{-5}$
adenosylhomocysteinase	$8.25 \times 10^{-5}$
adenylate kinase (GTP)	$4.86 \times 10^{-3}$
adenylosuccinate lyase	$5.86 \times 10^{-4}$
adenylosuccinate synthase	$2.93 \times 10^{-4}$
adenylsuccinate lyase	$2.93 \times 10^{-4}$
adenyl-sulfate kinase	$4.98 \times 10^{-4}$
alcohol dehydrogenase (ethanol)	1.29
alcohol dehydrogenase, reverse rxn (acetaldehyde -> ethanol)	$1.33 \times 10^{-2}$
aldehyde dehydrogenase (acetaldehyde, NADP)	$3.70 \times 10^{-3}$
alpha,alpha-trehalase	$9.00 \times 10^{-6}$
alpha,alpha-trehalose-phosphate synthase (UDP-forming)	$1.34 \times 10^{-4}$
anthranilate phosphoribosyltransferase	$1.50 \times 10^{-4}$
anthranilate synthase	$1.50 \times 10^{-4}$
argininosuccinate lyase	$7.28 \times 10^{-4}$
argininosuccinate synthase	$7.28 \times 10^{-4}$
asparagine synthase (glutamine-hydrolysing)	$9.19 \times 10^{-4}$
aspartate kinase	$3.08 \times 10^{-3}$
aspartate transaminase	$6.52 \times 10^{-3}$
aspartate-semialdehyde dehydrogenase	$3.08 \times 10^{-3}$
ATP synthase	1.72
ATPase, cytosolic	2.48
bicarbonate formation	$4.20 \times 10^{-2}$
C-14 sterol reductase	$3.66 \times 10^{-5}$
C-22 sterol desaturase (NAD)	$1.74 \times 10^{-5}$
C-3 sterol dehydrogenase	$3.63 \times 10^{-5}$
C-3 sterol dehydrogenase (4-methylzymosterol)	$3.63 \times 10^{-5}$
C-3 sterol keto reductase (4-methylzymosterol)	$3.63 \times 10^{-5}$

Continued on Next Page...

Table B.1 – Continued

Reaction	Flux(mmol/s)
C-3 sterol keto reductase (zymosterol)	$3.63 \times 10^{-5}$
C-4 methyl sterol oxidase	$3.63 \times 10^{-5}$
C-4 methyl sterol oxidase	$3.63 \times 10^{-5}$
C-4 methyl sterol oxidase	$3.63 \times 10^{-5}$
C-4 sterol methyl oxidase (4,4-dimethylzymosterol)	$3.63 \times 10^{-5}$
C-5 sterol desaturase	$1.74 \times 10^{-5}$
C-8 sterol isomerase	$1.79 \times 10^{-5}$
C-s24 sterol reductase	$3.44 \times 10^{-5}$
carbamoyl-phosphate synthase (glutamine-hydrolysing)	$7.28 \times 10^{-4}$
CDP-diacylglycerol synthase	$2.94 \times 10^{-5}$
ceramide-1 hydroxylase (24C)	$2.23 \times 10^{-6}$
ceramide-1 synthase (24C)	$2.23 \times 10^{-6}$
cholesterol delta-isomerase, lumped reaction	$1.77 \times 10^{-5}$
chorismate mutase	$1.13 \times 10^{-3}$
chorismate synthase	$1.28 \times 10^{-3}$
cis-aconitate(3-) to isocitrate	$5.57 \times 10^{-3}$
citrate synthase	$5.57 \times 10^{-3}$
citrate to cis-aconitate(3-)	$5.57 \times 10^{-3}$
CTP synthase (NH <sub>3</sub> )	$2.68 \times 10^{-4}$
cystathionine beta-synthase	$2.30 \times 10^{-4}$
cystathionine g-lyase	$5.75 \times 10^{-4}$
cystathionine gamma-synthase	$3.45 \times 10^{-4}$
cytidylate kinase (CMP)	$2.39 \times 10^{-4}$
cytochrome P450 lanosterol 14-alpha-demethylase (NAD)	$3.66 \times 10^{-5}$
D-arabinose 1-dehydrogenase (NAD)	$3.53 \times 10^{-3}$
D-arabinose 1-dehydrogenase (NADP)	$3.53 \times 10^{-3}$
D-sorbitol reductase	$8.02 \times 10^{-3}$
dCMP deaminase	$1.30 \times 10^{-5}$
deoxyadenylate kinase	$1.92 \times 10^{-5}$
deoxyguanylate kinase (dGMP:ATP)	$1.30 \times 10^{-5}$
diacylglycerol acyltransferase	$9.64 \times 10^{-6}$
diacylglycerol pyrophosphate phosphatase	$4.19 \times 10^{-6}$
dihydrofolate reductase	$1.92 \times 10^{-5}$
dihydroxy-acid dehydratase (2,3-dihydroxy-3-methylbutanoate)	$1.38 \times 10^{-3}$
dihydroxy-acid dehydratase (2,3-dihydroxy-3-methylpentanoate)	$9.19 \times 10^{-4}$
dimethylallyltranstransferase	$7.35 \times 10^{-5}$
dolichyl-phosphate D-mannosyltransferase	$4.40 \times 10^{-3}$
dolichyl-phosphate-mannose-protein mannosyltransferase	$4.40 \times 10^{-3}$
enolase	1.36
fatty acid synthase (n-C10:0)	$1.12 \times 10^{-6}$
fatty acid synthase (n-C12:0)	$1.12 \times 10^{-6}$
fatty acid synthase (n-C14:0)	$1.12 \times 10^{-6}$
fatty acid synthase (n-C16:0)	$1.12 \times 10^{-6}$

Continued on Next Page...

Table B.1 – Continued

Reaction	Flux(mmol/s)
fatty acid synthase (n-C18:0)	$1.12 \times 10^{-6}$
fatty acid synthase (n-C24:0), lumped reaction	$2.23 \times 10^{-6}$
fatty acyl-CoA synthase (n-C10:0CoA)	$3.35 \times 10^{-6}$
fatty acyl-CoA synthase (n-C8:0CoA), lumped reaction	$4.47 \times 10^{-6}$
fatty-acid-CoA ligase (n-C24:0)	$2.23 \times 10^{-6}$
fatty-acid-CoA ligase (octadecanoate)	$1.12 \times 10^{-6}$
fatty-acid-CoA ligase (octanoate)	$1.12 \times 10^{-6}$
fatty-acyl-CoA synthase (n-C12:0CoA)	$3.35 \times 10^{-6}$
fatty-acyl-CoA synthase (n-C14:0CoA)	$3.35 \times 10^{-6}$
fatty-acyl-CoA synthase (n-C16:0CoA)	$3.35 \times 10^{-6}$
fatty-acyl-CoA synthase (n-C18:0CoA)	$1.12 \times 10^{-6}$
ferrocytochrome-c:oxygen oxidoreductase	$4.39 \times 10^{-5}$
ferrocytochrome-c:oxygen oxidoreductase	$8.79 \times 10^{-5}$
FMN reductase	$9.08 \times 10^{-4}$
formate-tetrahydrofolate ligase	$5.10 \times 10^{-5}$
fructose-bisphosphate aldolase	$6.82 \times 10^{-1}$
fumarase	$1.23 \times 10^{-4}$
fumarate reductase FMN	$9.08 \times 10^{-4}$
geranyltranstransferase	$7.35 \times 10^{-5}$
glucose-6-phosphate isomerase	$6.72 \times 10^{-1}$
glucose-6-phosphate isomerase	$6.72 \times 10^{-1}$
glucose-6-phosphate isomerase	$1.11 \times 10^{-1}$
glutamate 5-kinase	$6.89 \times 10^{-4}$
glutamate dehydrogenase (NADP)	$1.14 \times 10^{-2}$
glutamate synthase (NADH2)	$1.14 \times 10^{-2}$
glutamate-5-semialdehyde dehydrogenase	$6.89 \times 10^{-4}$
glutamine phosphoribosyldiphosphate amidotransferase	$5.86 \times 10^{-4}$
glutamine synthetase	$1.11 \times 10^{-1}$
glutamine-fructose-6-phosphate transaminase	$9.46 \times 10^{-2}$
glyceraldehyde-3-phosphate dehydrogenase	1.36
glycerol-3-phosphatase	$5.18 \times 10^{-4}$
glycerol-3-phosphate dehydrogenase (NAD)	$5.52 \times 10^{-4}$
glycerol-3-phosphate/dihydroxyacetone phosphate acyltransferase	$3.36 \times 10^{-5}$
glycine hydroxymethyltransferase	$1.49 \times 10^{-3}$
glycogen (starch) synthase	$2.78 \times 10^{-3}$
GMP synthase	$2.93 \times 10^{-4}$
GTP cyclohydrolase II	$4.82 \times 10^{-6}$
guanylate kinase (GMP:ATP)	$1.05 \times 10^{-5}$
guanylate kinase (GMP:dATP)	$9.61 \times 10^{-6}$
hexokinase (D-glucose:ATP)	$7.92 \times 10^{-1}$
homoaccontinate hydratase	$1.28 \times 10^{-3}$
homocitrate synthase	$1.28 \times 10^{-3}$
homoisocitrate dehydrogenase	$1.28 \times 10^{-3}$

Continued on Next Page...

Table B.1 – Continued

Reaction	Flux(mmol/s)
homoserine dehydrogenase (NADH)	$3.08 \times 10^{-3}$
homoserine kinase	$2.24 \times 10^{-3}$
homoserine O-trans-acetylase	$8.43 \times 10^{-4}$
hydroxymethylglutaryl CoA reductase	$2.20 \times 10^{-4}$
hydroxymethylglutaryl CoA synthase	$2.20 \times 10^{-4}$
IMP cyclohydrolase	$5.86 \times 10^{-4}$
IMP dehydrogenase	$2.93 \times 10^{-4}$
indole-3-glycerol-phosphate synthase	$1.50 \times 10^{-4}$
inorganic diphosphatase	$8.28 \times 10^{-3}$
inositolphosphotransferase	$2.23 \times 10^{-6}$
IPC synthase	$2.23 \times 10^{-6}$
isocitrate dehydrogenase (NADP)	$5.57 \times 10^{-3}$
isoleucine transaminase	$9.19 \times 10^{-4}$
isopentenyl-diphosphate D-isomerase	$7.35 \times 10^{-5}$
ketol-acid reductoisomerase (2-aceto-2-hydroxybutanoate)	$9.19 \times 10^{-4}$
L-alanine transaminase	$1.92 \times 10^{-3}$
L-aminoadipate-semialdehyde dehydrogenase (NADH)	$1.28 \times 10^{-3}$
L-glutamate 5-semialdehyde dehydratase	$6.89 \times 10^{-4}$
L-hydroxyproline dehydrogenase (NADP)	$3.53 \times 10^{-3}$
L-hydroxyproline reductase (NAD)	$3.53 \times 10^{-3}$
L-threonine deaminase	$3.45 \times 10^{-4}$
lanosterol synthase	$3.67 \times 10^{-5}$
malate dehydrogenase	$2.66 \times 10^{-2}$
mannose-1-phosphate guanylyltransferase	$4.40 \times 10^{-3}$
mannose-6-phosphate isomerase	$4.40 \times 10^{-3}$
methenyltetrahydrifkate cyclohydrolase	$1.12 \times 10^{-3}$
methionine adenosyltransferase	$8.25 \times 10^{-5}$
methionine synthase	$3.51 \times 10^{-4}$
methylenetetrahydrofolate dehydrogenase (NADP)	$1.12 \times 10^{-3}$
mevalonate kinase (ctp)	$2.20 \times 10^{-4}$
mevalonate pyrophosphate decarboxylase	$2.20 \times 10^{-4}$
microsomal beta-keto-reductase	$3.53 \times 10^{-3}$
MIPC synthase	$2.23 \times 10^{-6}$
myo-inositol 1-phosphatase	$8.20 \times 10^{-6}$
myo-inositol-1-phosphate synthase	$1.27 \times 10^{-5}$
N-acetyl-g-glutamyl-phosphate reductase	$7.28 \times 10^{-4}$
NADH dehydrogenase, cytosolic/mitochondrial	$9.09 \times 10^{-4}$
non-enzymatic reaction	$1.28 \times 10^{-3}$
nucleoside-diphosphate kinase (ATP:CDP)	$1.82 \times 10^{-5}$
nucleoside-diphosphate kinase (ATP:UDP)	$9.27 \times 10^{-3}$

Continued on Next Page...

Table B.1 – Continued

Reaction	Flux(mmol/s)
O-acetylhomoserine (thiol)-lyase	$4.98 \times 10^{-4}$
ornithine carbamoyltransferase	$7.28 \times 10^{-4}$
ornithine transacetylase	$7.28 \times 10^{-4}$
phenylalanine transaminase	$6.13 \times 10^{-4}$
phosphatidylethanolamine methyltransferase	$1.55 \times 10^{-5}$
phosphatidylinositol synthase	$8.20 \times 10^{-6}$
phosphatidylserine decarboxylase	$1.92 \times 10^{-5}$
phosphatidylserine synthase	$2.12 \times 10^{-5}$
phosphoadenylyl-sulfate reductase (thioredoxin)	$4.98 \times 10^{-4}$
phosphoenolpyruvate carboxykinase	$2.56 \times 10^{-3}$
phosphofructokinase	$6.82 \times 10^{-1}$
phosphoglucomutase	$9.00 \times 10^{-3}$
phosphoglycerate dehydrogenase	$3.25 \times 10^{-3}$
phosphoglycerate kinase	1.36
phosphoglycerate mutase	1.36
phospholipid methyltransferase	$1.55 \times 10^{-5}$
phospholipid methyltransferase	$1.55 \times 10^{-5}$
phosphomannomutase	$4.40 \times 10^{-3}$
phosphomevalonate kinase	$2.20 \times 10^{-4}$
phosphopentomutase	$3.30 \times 10^{-4}$
phosphoribosylaminoimidazole carboxylase	$5.86 \times 10^{-4}$
phosphoribosylaminoimidazole synthase	$5.86 \times 10^{-4}$
phosphoribosylaminoimidazolecarboxamide formyltransferase	$5.86 \times 10^{-4}$
phosphoribosylaminoimidazolesuccinocarboxamide synthase	$5.86 \times 10^{-4}$
phosphoribosylanthranilate isomerase	$1.50 \times 10^{-4}$
phosphoribosylformylglycinamide synthase	$5.86 \times 10^{-4}$
phosphoribosylglycinamide formyltransferase	$5.86 \times 10^{-4}$
phosphoribosylglycinamide synthase	$5.86 \times 10^{-4}$
phosphoribosylpyrophosphate synthetase	$1.07 \times 10^{-3}$
phosphoserine phosphatase (L-serine)	$3.25 \times 10^{-3}$
phosphoserine transaminase	$3.25 \times 10^{-3}$
prephenate dehydratase	$6.13 \times 10^{-4}$
prephenate dehydrogenase (NADP)	$5.17 \times 10^{-4}$
pyrimidine phosphatase	$4.82 \times 10^{-6}$
pyrimidine-nucleoside phosphorylase (uracil)	$3.30 \times 10^{-4}$
pyroline-5-carboxylate reductase	$6.89 \times 10^{-4}$
pyruvate carboxylase	$4.12 \times 10^{-2}$
pyruvate decarboxylase	1.31
pyruvate dehydrogenase	$3.86 \times 10^{-3}$
pyruvate kinase	1.36
riboflavin synthase	$9.65 \times 10^{-6}$
riboflavin synthase	$4.82 \times 10^{-6}$
ribonucleoside-diphosphate reductase	$9.61 \times 10^{-6}$

Continued on Next Page...

Table B.1 – Continued

Reaction	Flux(mmol/s)
ribonucleoside-diphosphate reductase (GDP)	$1.30 \times 10^{-5}$
ribonucleoside-diphosphate reductase (UDP)	$3.23 \times 10^{-5}$
ribonucleoside-triphosphate reductase (ATP)	$9.61 \times 10^{-6}$
ribose-5-phosphate isomerase	$1.35 \times 10^{-3}$
ribulose 5-phosphate 3-epimerase	$1.36 \times 10^{-3}$
S-adenosyl-methionine delta-24-sterol-c-methyltransferase	$1.85 \times 10^{-5}$
saccharopine dehydrogenase (NAD, L-lysine forming)	$1.28 \times 10^{-3}$
saccharopine dehydrogenase (NADP, L-glutamate forming)	$1.28 \times 10^{-3}$
serine C-palmitoyltransferase	$2.23 \times 10^{-6}$
shikimate dehydrogenase	$1.28 \times 10^{-3}$
shikimate kinase	$1.28 \times 10^{-3}$
squalene epoxidase (NAD)	$3.67 \times 10^{-5}$
squalene synthase	$3.67 \times 10^{-5}$
steryl ester hydrolase	$4.35 \times 10^{-6}$
succinate dehydrogenase (ubiquinone-6)	$8.21 \times 10^{-4}$
succinate-CoA ligase (ADP-forming)	$7.68 \times 10^{-5}$
sulfate adenylyltransferase (ADP)	$4.98 \times 10^{-4}$
sulfite reductase (NADPH <sub>2</sub> )	$4.98 \times 10^{-4}$
thioredoxin reductase (NADPH)	$5.62 \times 10^{-4}$
threonine aldolase	$8.38 \times 10^{-4}$
threonine synthase	$2.24 \times 10^{-3}$
thymidylate synthase	$1.92 \times 10^{-5}$
transaldolase	$4.17 \times 10^{-5}$
transketolase	$4.17 \times 10^{-5}$
transketolase	$1.32 \times 10^{-3}$
trehalose-phosphatase	$1.34 \times 10^{-4}$
triacylglycerol lipase	$5.46 \times 10^{-6}$
triose-phosphate isomerase	$6.82 \times 10^{-1}$
tryptophan synthase (indoleglycerol phosphate)	$1.50 \times 10^{-4}$
tyrosine transaminase	$5.17 \times 10^{-4}$
UMP kinase	$3.00 \times 10^{-4}$
uracil phosphoribosyltransferase	$3.30 \times 10^{-4}$
uridine kinase (ATP:uridine)	$1.65 \times 10^{-4}$
uridine kinase (GTP:uridine)	$1.65 \times 10^{-4}$
uridylylate kinase (dUMP)	$3.23 \times 10^{-5}$
UTP-glucose-1-phosphate uridylyltransferase	$9.00 \times 10^{-3}$
valine transaminase	$1.38 \times 10^{-3}$
ammonia transport	$1.21 \times 10^{-1}$
CO <sub>2</sub> transport	1.28
D-glucosamine 6-phosphate uniport	$9.46 \times 10^{-2}$
D-sorbitol transport	$8.02 \times 10^{-3}$
ethanol transport	1.31
glucose transport	$8.00 \times 10^{-1}$

Continued on Next Page...

Table B.1 – Continued

Reaction	Flux(mmol/s)
glycerol transport via channel	$5.18 \times 10^{-4}$
L-histidine transport	$4.02 \times 10^{-4}$
L-leucine transport	$1.34 \times 10^{-3}$
L-malate transport	$2.65 \times 10^{-2}$
O <sub>2</sub> transport	$4.63 \times 10^{-4}$
phosphate transport	$9.57 \times 10^{-2}$
succinate transport	$1.65 \times 10^{-3}$
sulfate uniport	$6.05 \times 10^{-4}$
uracil transport	$6.59 \times 10^{-4}$
isa acyl-CoA	$7.68 \times 10^{-5}$
biomass production	$5.36 \times 10^{-3}$
growth	$5.36 \times 10^{-3}$
lipid production	$5.36 \times 10^{-3}$

Table B.2: Metabolite Concentrations

Metabolite	Concentration (mM)
(1->3)-beta-D-glucan [intracellular]	$7.96 \times 10^{-1}$
(2R,3R)-2,3-dihydroxy-3-methylpentanoate [intracellular]	$7.96 \times 10^{-1}$
(2S)-2-[5-amino-1-(5-phospho-beta-D-riboseyl)imidazole-4-carboxamido]succinic acid [intracellular]	$7.96 \times 10^{-1}$
(6R)-5,10-methenyltetrahydrofolic acid [intracellular]	$7.96 \times 10^{-1}$
(N(omega)-L-arginino)succinic acid [intracellular]	$7.96 \times 10^{-1}$
(R)-2,3-dihydroxy-3-methylbutanoate [intracellular]	$7.96 \times 10^{-1}$
(R)-5-diphosphomevalonic acid [intracellular]	$7.96 \times 10^{-1}$
(R)-5-phosphomevalonic acid [intracellular]	$7.96 \times 10^{-1}$
(R)-mevalonate [intracellular]	$7.96 \times 10^{-1}$
(S)-2,3-epoxysqualene [intracellular]	$7.96 \times 10^{-1}$
(S)-2-acetyl-2-hydroxybutanoate [intracellular]	$7.96 \times 10^{-1}$
(S)-3-hydroxyhexacosanoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
(S)-3-hydroxypalmitoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
(S)-3-hydroxytetradecanoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
(S)-3-methyl-2-oxopentanoate [intracellular]	$7.96 \times 10^{-1}$
(S)-malate(2-) [intracellular]	$7.96 \times 10^{-1}$
1-(2-carboxyphenylamino)-1-deoxy-D-ribose 5-phosphate [intracellular]	$7.96 \times 10^{-1}$
1-acyl-sn-glycerol 3-phosphate [intracellular]	$7.96 \times 10^{-1}$
1-C-(indol-3-yl)glycerol 3-phosphate [intracellular]	$7.96 \times 10^{-1}$
1-phosphatidyl-1D-myo-inositol [intracellular]	$7.96 \times 10^{-1}$
1-pyrroline-3-hydroxy-5-carboxylic acid [intracellular]	$7.96 \times 10^{-1}$
1-pyrroline-5-carboxylate [intracellular]	$7.96 \times 10^{-1}$
10-formyltetrahydrofolic acid [intracellular]	$7.96 \times 10^{-1}$
14-demethylsterol [intracellular]	$7.96 \times 10^{-1}$

Continued on Next Page...

Table B.2 – Continued

Metabolite	Concentration (mM)
1D-myo-inositol 1-phosphate [intracellular]	$7.96 \times 10^{-1}$
2,5-diamino-4-hydroxy-6-(5-phosphoribosylamino)pyrimidine [intracellular]	$7.96 \times 10^{-1}$
2,5-diamino-6-(5-phosphono)ribitylamino-4(3H)-pyrimidinone [intracellular]	$7.96 \times 10^{-1}$
2-acetamido-5-oxopentanoate [intracellular]	$7.96 \times 10^{-1}$
2-acetyllactic acid [intracellular]	$7.96 \times 10^{-1}$
2-formamido-N(1)-(5-phospho-D-ribosyl)acetamide [intracellular]	$7.96 \times 10^{-1}$
2-hydroxy-3-oxobutyl phosphate [intracellular]	$7.96 \times 10^{-1}$
2-oxaloglutaric acid [intracellular]	$7.96 \times 10^{-1}$
2-oxoadipic acid [intracellular]	$7.96 \times 10^{-1}$
2-oxobutanoate [intracellular]	$7.96 \times 10^{-1}$
2-oxoglutarate [intracellular]	$7.96 \times 10^{-1}$
2-phospho-D-glyceric acid [intracellular]	$2.01 \times 10^{-2}$
2-trans,6-trans-farnesyl diphosphate [intracellular]	$7.96 \times 10^{-1}$
3-phospho-5-adenylyl sulfate [intracellular]	$7.96 \times 10^{-1}$
3-(4-hydroxyphenyl)pyruvate [intracellular]	$7.96 \times 10^{-1}$
3-dehydro-4-methylzymosterol [intracellular]	$7.96 \times 10^{-1}$
3-dehydroquininate [intracellular]	$7.96 \times 10^{-1}$
3-dehydroshikimate [intracellular]	$7.96 \times 10^{-1}$
3-dehydrosphinganine [intracellular]	$7.96 \times 10^{-1}$
3-hydroxy-3-methylglutaryl-CoA [intracellular]	$7.96 \times 10^{-1}$
3-hydroxyoctadecanoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
3-methyl-2-oxobutanoate [intracellular]	$7.96 \times 10^{-1}$
3-oxohexacosanoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
3-oxooctadecanoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
3-oxopalmitoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
3-oxotetradecanoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
3-phospho-D-glyceric acid [intracellular]	$1.22 \times 10^{-1}$
3-phospho-D-glyceroyl dihydrogen phosphate [intracellular]	$1.19 \times 10^{-3}$
3-phosphonatooxypyruvate(3-) [intracellular]	$7.96 \times 10^{-1}$
3-phosphoshikimic acid [intracellular]	$7.96 \times 10^{-1}$
4,4-dimethyl-5 $\alpha$ -cholesta-8,14,24-trien-3 $\beta$ -ol [intracellular]	$7.96 \times 10^{-1}$
4-phospho-L-aspartate [intracellular]	$7.96 \times 10^{-1}$
4 $\alpha$ -methylzymosterol [intracellular]	$7.96 \times 10^{-1}$
4 $\beta$ -methylzymosterol-4 $\alpha$ -carboxylic acid [intracellular]	$7.96 \times 10^{-1}$
5-adenylyl sulfate [intracellular]	$7.96 \times 10^{-1}$
5-xanthylic acid [intracellular]	$7.96 \times 10^{-1}$
5,10-methylenetetrahydrofolate(2-) [intracellular]	$7.96 \times 10^{-1}$
5,6,7,8-tetrahydrofolic acid [intracellular]	$7.96 \times 10^{-1}$
5-amino-1-(5-phospho-D-ribosyl)imidazole [intracellular]	$7.96 \times 10^{-1}$
5-amino-1-(5-phospho-D-ribosyl)imidazole-4-carboxamide [intracellular]	$7.96 \times 10^{-1}$
5-amino-1-(5-phospho-D-ribosyl)imidazole-4-carboxylic acid [intracellular]	$7.96 \times 10^{-1}$
5-amino-6-(5-phosphoribitylamino)uracil [intracellular]	$7.96 \times 10^{-1}$
5-amino-6-(D-ribitylamino)uracil [intracellular]	$7.96 \times 10^{-1}$

Continued on Next Page...

Table B.2 – Continued

Metabolite	Concentration (mM)
5-formamido-1-(5-phospho-D-ribosyl)imidazole-4-carboxamide [intracellular]	$7.96 \times 10^{-1}$
5-methyltetrahydrofolate(2-) [intracellular]	$7.96 \times 10^{-1}$
5-O-(1-carboxyvinyl)-3-phosphoshikimic acid [intracellular]	$7.96 \times 10^{-1}$
5-O-phosphono-alpha-D-ribofuranosyl diphosphate [intracellular]	$7.96 \times 10^{-1}$
5-phospho-beta-D-ribosylamine [intracellular]	$7.96 \times 10^{-1}$
6,7-dimethyl-8-(1-D-ribityl)lumazine [intracellular]	$7.96 \times 10^{-1}$
7-phospho-2-dehydro-3-deoxy-D-arabino-heptonic acid [intracellular]	$7.96 \times 10^{-1}$
acetaldehyde [intracellular]	1.16
acetate [intracellular]	$7.96 \times 10^{-1}$
acetoacetyl-CoA [intracellular]	$7.96 \times 10^{-1}$
acetyl-CoA [intracellular]	$7.96 \times 10^{-1}$
acyl-CoA [intracellular]	$7.96 \times 10^{-1}$
adenosine [intracellular]	$7.96 \times 10^{-1}$
adenosine 3,5-bismonophosphate [intracellular]	$7.96 \times 10^{-1}$
ADP [intracellular]	1.10
aldehyde-D-glucose 6-phosphate [intracellular]	$7.96 \times 10^{-1}$
alpha,alpha-trehalose [intracellular]	$5.00 \times 10^{-2}$
alpha,alpha-trehalose 6-phosphate [intracellular]	$2.00 \times 10^{-2}$
alpha-D-glucosamine 6-phosphate [intracellular]	$7.96 \times 10^{-1}$
alpha-D-ribose 1-phosphate [intracellular]	$7.96 \times 10^{-1}$
alpha-D-ribose 5-phosphate [intracellular]	$7.96 \times 10^{-1}$
ammonium [intracellular]	$7.96 \times 10^{-1}$
AMP [intracellular]	$1.19 \times 10^{-1}$
amylose [intracellular]	$7.96 \times 10^{-1}$
anthranilate [intracellular]	$7.96 \times 10^{-1}$
ATP [intracellular]	4.58
beta-D-glucose 6-phosphate [intracellular]	3.33
bicarbonate [intracellular]	$7.96 \times 10^{-1}$
biomass [intracellular]	$7.96 \times 10^{-1}$
but-1-ene-1,2,4-tricarboxylic acid [intracellular]	$7.96 \times 10^{-1}$
carbamoyl phosphate [intracellular]	$7.96 \times 10^{-1}$
carbon dioxide [intracellular]	$7.96 \times 10^{-1}$
CDP [intracellular]	$7.96 \times 10^{-1}$
CDP-diacylglycerol [intracellular]	$7.96 \times 10^{-1}$
chorismate(2-) [intracellular]	$7.96 \times 10^{-1}$
cis-aconitate(3-) [intracellular]	$7.96 \times 10^{-1}$
citrate(3-) [intracellular]	$7.96 \times 10^{-1}$
CMP [intracellular]	$7.96 \times 10^{-1}$
coenzyme A [intracellular]	$7.96 \times 10^{-1}$
CTP [intracellular]	$7.96 \times 10^{-1}$
D-arabinono-1,4-lactone [intracellular]	$7.96 \times 10^{-1}$
D-arabinose [intracellular]	$7.96 \times 10^{-1}$
D-erythrose 4-phosphate(2-) [intracellular]	$7.96 \times 10^{-1}$

Continued on Next Page...

Table B.2 – Continued

Metabolite	Concentration (mM)
D-fructose 1,6-bisphosphate [intracellular]	1.68
D-fructose 6-phosphate [intracellular]	$9.08 \times 10^{-1}$
D-glucitol [intracellular]	$7.96 \times 10^{-1}$
D-glucose [intracellular]	$7.69 \times 10^{-1}$
D-glucose 1-phosphate [intracellular]	$1.00 \times 10^{-1}$
D-mannose 1-phosphate [intracellular]	$7.96 \times 10^{-1}$
D-mannose 6-phosphate [intracellular]	$7.96 \times 10^{-1}$
D-ribulose 5-phosphate [intracellular]	$7.96 \times 10^{-1}$
D-xylulose 5-phosphate [intracellular]	$7.96 \times 10^{-1}$
dADP [intracellular]	$7.96 \times 10^{-1}$
dAMP [intracellular]	$7.96 \times 10^{-1}$
dATP [intracellular]	$7.96 \times 10^{-1}$
dCMP [intracellular]	$7.96 \times 10^{-1}$
decanoate [intracellular]	$7.96 \times 10^{-1}$
decanoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
dGDP [intracellular]	$7.96 \times 10^{-1}$
dGMP [intracellular]	$7.96 \times 10^{-1}$
diglyceride [intracellular]	$7.96 \times 10^{-1}$
dihydrofolic acid [intracellular]	$7.96 \times 10^{-1}$
diphosphate [intracellular]	$7.96 \times 10^{-1}$
dolichyl D-mannosyl phosphate [intracellular]	$7.96 \times 10^{-1}$
dolichyl phosphate [intracellular]	$7.96 \times 10^{-1}$
dTMP [intracellular]	$7.96 \times 10^{-1}$
dUDP [intracellular]	$7.96 \times 10^{-1}$
dUMP [intracellular]	$7.96 \times 10^{-1}$
episterol [intracellular]	$7.96 \times 10^{-1}$
ergosta-5,7,22,24(28)-tetraen-3beta-ol [intracellular]	$7.96 \times 10^{-1}$
ergosta-5,7,24(28)-trien-3beta-ol [intracellular]	$7.96 \times 10^{-1}$
ergosterol [intracellular]	$7.96 \times 10^{-1}$
ergosterol ester [intracellular]	$7.96 \times 10^{-1}$
ethanol [intracellular]	$2.08 \times 10^{+2}$
fatty acid [intracellular]	$7.96 \times 10^{-1}$
fecosterol [intracellular]	$7.96 \times 10^{-1}$
ferricytochrome c [intracellular]	$7.96 \times 10^{-1}$
ferrocyclochrome c [intracellular]	$7.96 \times 10^{-1}$
FMN [intracellular]	$7.96 \times 10^{-1}$
FMNH <sub>2</sub> [intracellular]	$7.96 \times 10^{-1}$
formate [intracellular]	$7.96 \times 10^{-1}$
fumarate(2-) [intracellular]	$7.96 \times 10^{-1}$
GDP [intracellular]	$7.96 \times 10^{-1}$
GDP-alpha-D-mannose [intracellular]	$7.96 \times 10^{-1}$
geranyl diphosphate [intracellular]	$7.96 \times 10^{-1}$
glyceraldehyde 3-phosphate [intracellular]	$5.21 \times 10^{-2}$

Continued on Next Page...

Table B.2 – Continued

Metabolite	Concentration (mM)
glycerol [intracellular]	$2.05 \times 10^{+1}$
glycerone phosphate [intracellular]	1.24
glycine [intracellular]	$7.96 \times 10^{-1}$
glycogen [intracellular]	$7.96 \times 10^{-1}$
GMP [intracellular]	$7.96 \times 10^{-1}$
GTP [intracellular]	$7.96 \times 10^{-1}$
homocitrate(3-) [intracellular]	$7.96 \times 10^{-1}$
homoisocitrate(3-) [intracellular]	$7.96 \times 10^{-1}$
hydrogen sulfide [intracellular]	$7.96 \times 10^{-1}$
IMP [intracellular]	$7.96 \times 10^{-1}$
inositol phosphomannosylinositol phosphoceramide [intracellular]	$7.96 \times 10^{-1}$
inositol-P-ceramide B [intracellular]	$7.96 \times 10^{-1}$
isocitrate(3-) [intracellular]	$7.96 \times 10^{-1}$
isopentenyl diphosphate [intracellular]	$7.96 \times 10^{-1}$
keto-phenylpyruvate [intracellular]	$7.96 \times 10^{-1}$
L-2-aminoadipate(2-) [intracellular]	$7.96 \times 10^{-1}$
L-alanine [intracellular]	$7.96 \times 10^{-1}$
L-allysine [intracellular]	$7.96 \times 10^{-1}$
L-arginine [intracellular]	$7.96 \times 10^{-1}$
L-asparagine [intracellular]	$7.96 \times 10^{-1}$
L-aspartate [intracellular]	$7.96 \times 10^{-1}$
L-aspartate 4-semialdehyde [intracellular]	$7.96 \times 10^{-1}$
L-citrulline [intracellular]	$7.96 \times 10^{-1}$
L-cystathionine [intracellular]	$7.96 \times 10^{-1}$
L-cysteine [intracellular]	$7.96 \times 10^{-1}$
L-gamma-glutamyl phosphate [intracellular]	$7.96 \times 10^{-1}$
L-glutamate [intracellular]	$7.96 \times 10^{-1}$
L-glutamic 5-semialdehyde [intracellular]	$7.96 \times 10^{-1}$
L-glutamine [intracellular]	$7.96 \times 10^{-1}$
L-histidine [intracellular]	$9.90 \times 10^{-2}$
L-homocysteine [intracellular]	$7.96 \times 10^{-1}$
L-homoserine [intracellular]	$7.96 \times 10^{-1}$
L-isoleucine [intracellular]	$7.96 \times 10^{-1}$
L-leucine [intracellular]	$1.35 \times 10^{-1}$
L-lysine [intracellular]	$7.96 \times 10^{-1}$
L-methionine [intracellular]	$7.96 \times 10^{-1}$
L-phenylalanine [intracellular]	$7.96 \times 10^{-1}$
L-proline [intracellular]	$7.96 \times 10^{-1}$
L-saccharopine [intracellular]	$7.96 \times 10^{-1}$
L-serine [intracellular]	$7.96 \times 10^{-1}$
L-threonine [intracellular]	$7.96 \times 10^{-1}$
L-tryptophan [intracellular]	$7.96 \times 10^{-1}$
L-tyrosine [intracellular]	$7.96 \times 10^{-1}$

Continued on Next Page...

Table B.2 – Continued

Metabolite	Concentration (mM)
L-valine [intracellular]	$7.96 \times 10^{-1}$
lanosterol [intracellular]	$7.96 \times 10^{-1}$
laurate [intracellular]	$7.96 \times 10^{-1}$
lauroyl-CoA [intracellular]	$7.96 \times 10^{-1}$
lignocerate [intracellular]	$7.96 \times 10^{-1}$
lipid [intracellular]	$7.96 \times 10^{-1}$
malonyl-CoA [intracellular]	$7.96 \times 10^{-1}$
mannan [intracellular]	$7.96 \times 10^{-1}$
mannosylinositol phosphorylceramide [intracellular]	$7.96 \times 10^{-1}$
myo-inositol [intracellular]	$7.96 \times 10^{-1}$
myristate [intracellular]	$7.96 \times 10^{-1}$
myristoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
N(1)-(5-phospho-D-ribosyl)glycinamide [intracellular]	$7.96 \times 10^{-1}$
N(2)-acetyl-L-ornithine [intracellular]	$7.96 \times 10^{-1}$
N(2)-formyl-N(1)-(5-phospho-D-ribosyl)glycinamide [intracellular]	$7.96 \times 10^{-1}$
N(6)-(1,2-dicarboxyethyl)-AMP [intracellular]	$7.96 \times 10^{-1}$
N-(24-hydroxytetracosanyl)sphinganine [intracellular]	$7.96 \times 10^{-1}$
N-(5-phospho-beta-D-ribosyl)anthranilate [intracellular]	$7.96 \times 10^{-1}$
N-acetyl-L-gamma-glutamyl phosphate [intracellular]	$7.96 \times 10^{-1}$
N-acetyl-L-glutamate(2-) [intracellular]	$7.96 \times 10^{-1}$
N-tetracosanylsphinganine [intracellular]	$7.96 \times 10^{-1}$
NAD(+) [intracellular]	1.42
NADH [intracellular]	$8.67 \times 10^{-2}$
NADP(+) [intracellular]	$7.96 \times 10^{-1}$
NADPH [intracellular]	$7.96 \times 10^{-1}$
O-acetyl-L-homoserine [intracellular]	$7.96 \times 10^{-1}$
O-phospho-L-homoserine [intracellular]	$7.96 \times 10^{-1}$
O-phospho-L-serine [intracellular]	$7.96 \times 10^{-1}$
octanoate [intracellular]	$7.96 \times 10^{-1}$
octanoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
ornithine [intracellular]	$7.96 \times 10^{-1}$
oxaloacetate(2-) [intracellular]	$7.96 \times 10^{-1}$
oxygen [intracellular]	$7.96 \times 10^{-1}$
palmitate [intracellular]	$7.96 \times 10^{-1}$
palmitoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
phosphate [intracellular]	$7.96 \times 10^{-1}$
phosphatidate [intracellular]	$7.96 \times 10^{-1}$
phosphatidyl-L-serine [intracellular]	$7.96 \times 10^{-1}$
phosphatidyl-N,N-dimethylethanolamine [intracellular]	$7.96 \times 10^{-1}$
phosphatidyl-N-methylethanolamine [intracellular]	$7.96 \times 10^{-1}$
phosphatidylcholine [intracellular]	$7.96 \times 10^{-1}$
phosphatidylethanolamine [intracellular]	$7.96 \times 10^{-1}$
phosphoenolpyruvate [intracellular]	$6.46 \times 10^{-2}$

Continued on Next Page...

Table B.2 – Continued

Metabolite	Concentration (mM)
prenyl diphosphate [intracellular]	$7.96 \times 10^{-1}$
prephenate(2-) [intracellular]	$7.96 \times 10^{-1}$
pyruvate [intracellular]	2.65
riboflavin [intracellular]	$7.96 \times 10^{-1}$
S-adenosyl-L-homocysteine [intracellular]	$7.96 \times 10^{-1}$
S-adenosyl-L-methionine [intracellular]	$7.96 \times 10^{-1}$
sedoheptulose 7-phosphate [intracellular]	$7.96 \times 10^{-1}$
shikimate [intracellular]	$7.96 \times 10^{-1}$
sn-glycerol 3-phosphate [intracellular]	$1.43 \times 10^{-2}$
sphinganine [intracellular]	$7.96 \times 10^{-1}$
squalene [intracellular]	$7.96 \times 10^{-1}$
stearate [intracellular]	$7.96 \times 10^{-1}$
stearoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
succinate(2-) [intracellular]	$7.96 \times 10^{-1}$
succinyl-CoA [intracellular]	$7.96 \times 10^{-1}$
sulphate [intracellular]	$7.96 \times 10^{-1}$
sulphite [intracellular]	$7.96 \times 10^{-1}$
tetracosanoyl-CoA [intracellular]	$7.96 \times 10^{-1}$
trans-4-hydroxy-L-proline [intracellular]	$7.96 \times 10^{-1}$
triglyceride [intracellular]	$7.96 \times 10^{-1}$
ubiquinol-6 [intracellular]	$7.96 \times 10^{-1}$
ubiquinone-6 [intracellular]	$7.96 \times 10^{-1}$
UDP [intracellular]	$4.91 \times 10^{-1}$
UDP-D-glucose [intracellular]	$7.00 \times 10^{-1}$
UMP [intracellular]	$7.96 \times 10^{-1}$
uracil [intracellular]	$9.00 \times 10^{-2}$
uridine [intracellular]	$7.96 \times 10^{-1}$
UTP [intracellular]	1.13
zymosterol [intracellular]	$7.96 \times 10^{-1}$
zymosterol intermediate 1a [intracellular]	$7.96 \times 10^{-1}$
zymosterol intermediate 1b [intracellular]	$7.96 \times 10^{-1}$
zymosterol intermediate 1c [intracellular]	$7.96 \times 10^{-1}$
zymosterol intermediate 2 [intracellular]	$7.96 \times 10^{-1}$
thioredoxin disulfide [intracellular]	$7.96 \times 10^{-1}$
thioredoxin dithiol [intracellular]	$7.96 \times 10^{-1}$
(S)-malate(2-) [extracellular]	$7.00 \times 10^{-1}$
alpha-D-glucosamine 6-phosphate [extracellular]	0.00
ammonium [extracellular]	8.82
biomass [extracellular]	8.82
carbon dioxide [extracellular]	0.00
D-glucitol [extracellular]	$7.00 \times 10^{-1}$
D-glucose [extracellular]	$7.40 \times 10^{+1}$
ethanol [extracellular]	$5.20 \times 10^{+1}$

Continued on Next Page...

Table B.2 – Continued

<b>Metabolite</b>	<b>Concentration (mM)</b>
glycerol [extracellular]	$1.75 \times 10^{+1}$
H+ [intracellular]	8.82
H+ [extracellular]	8.82
L-histidine [extracellular]	$1.10 \times 10^{-1}$
L-leucine [extracellular]	$1.50 \times 10^{-1}$
oxygen [extracellular]	8.82
phosphate [extracellular]	8.82
succinate(2-) [extracellular]	$7.00 \times 10^{-1}$
sulphate [extracellular]	8.82
uracil [extracellular]	$1.00 \times 10^{-1}$
water [intracellular]	8.82

DYNAMIC FLUX EXPERIMENTS OFFER FURTHER  
INSIGHT INTO CELLULAR BEHAVIOUR -  
SUPPLEMENTARY INFORMATION

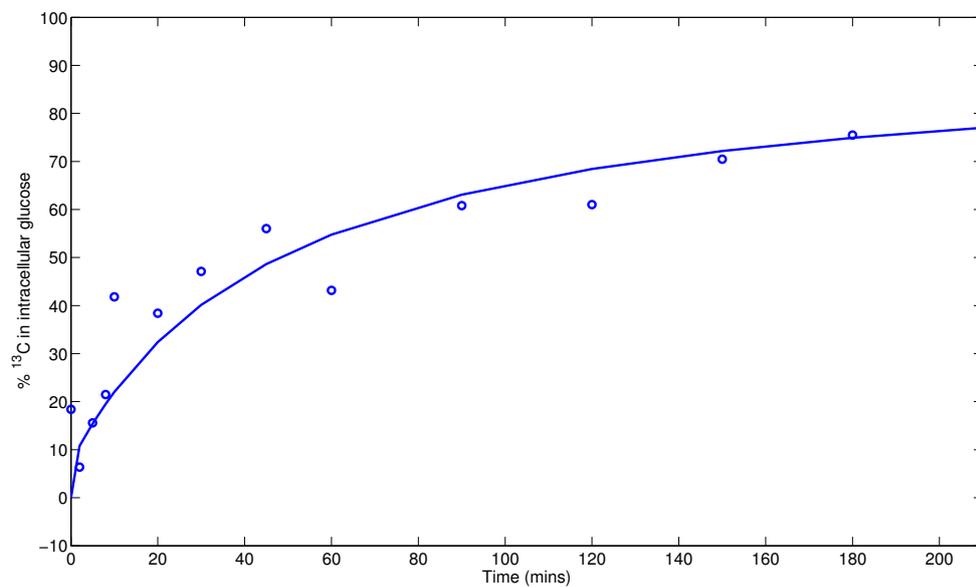


Figure C.1: Extracellular glucose. This was not fitted using a least squares approach, but was set as time based Michaelis Menten rate law

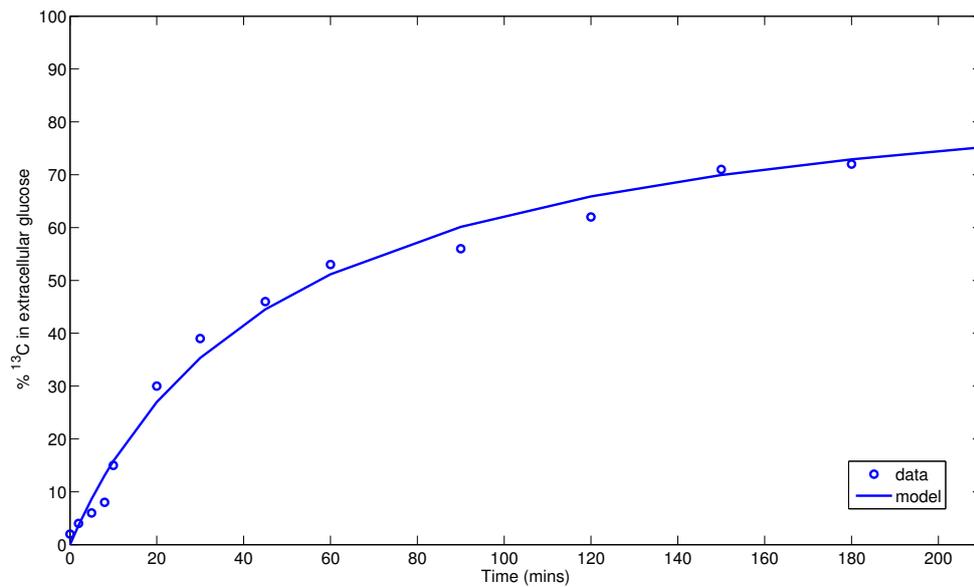


Figure C.2: Data fit for the changing  $^{13}\text{C}$  percentage in Intracellular glucose

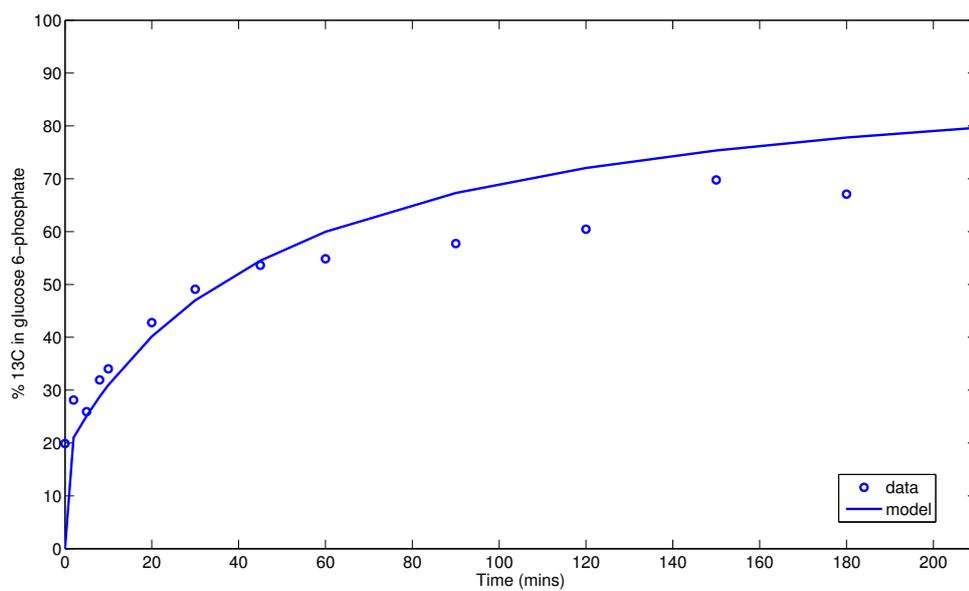


Figure C.3: Data fit for the changing  $^{13}\text{C}$  percentage in glucose 6-phosphate

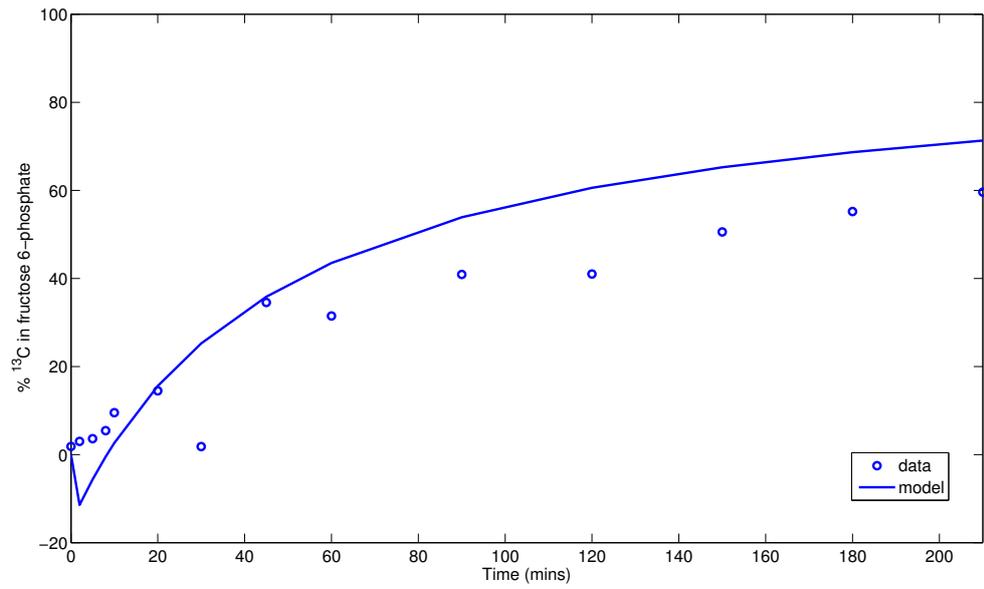


Figure C.4: Data fit for the changing  $^{13}\text{C}$  percentage in fructose 6-phosphate

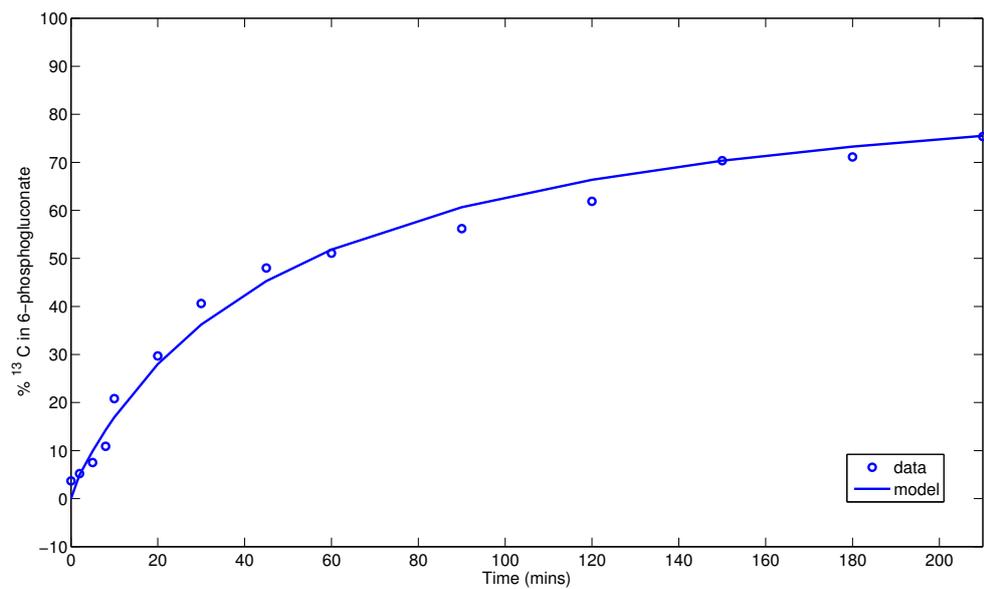


Figure C.5: Data fit for the changing  $^{13}\text{C}$  percentage in 6-phosphogluconate

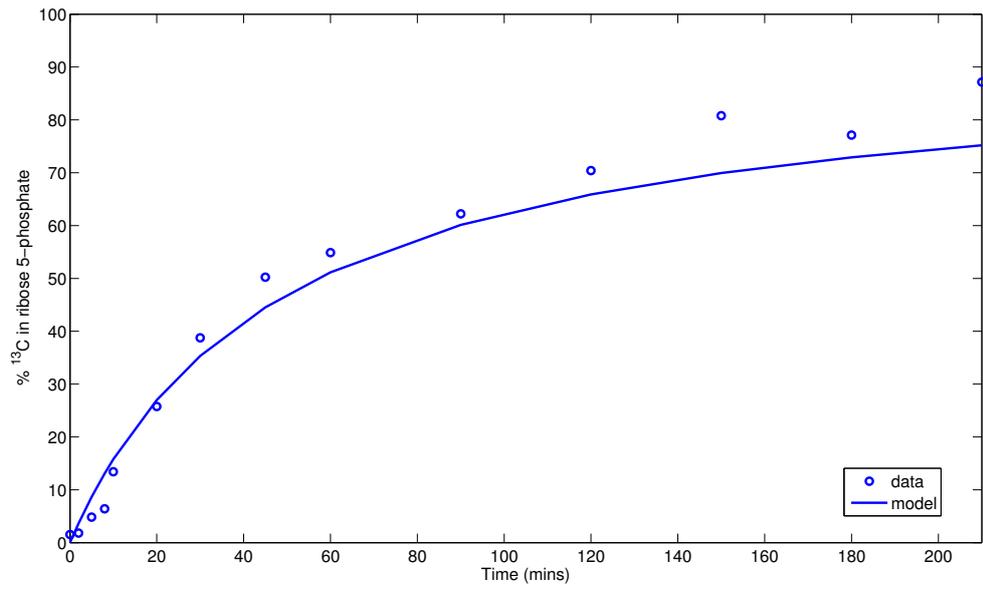


Figure C.6: Data fit for the changing  $^{13}\text{C}$  percentage in ribose 5-phosphate

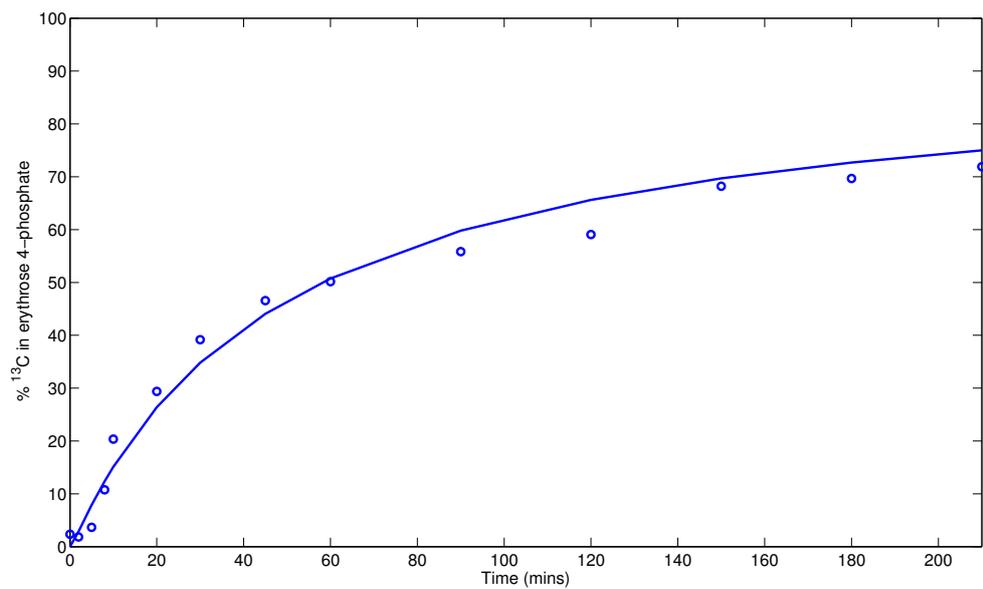


Figure C.7: Data fit for the changing  $^{13}\text{C}$  percentage in erythrose 4-phosphate

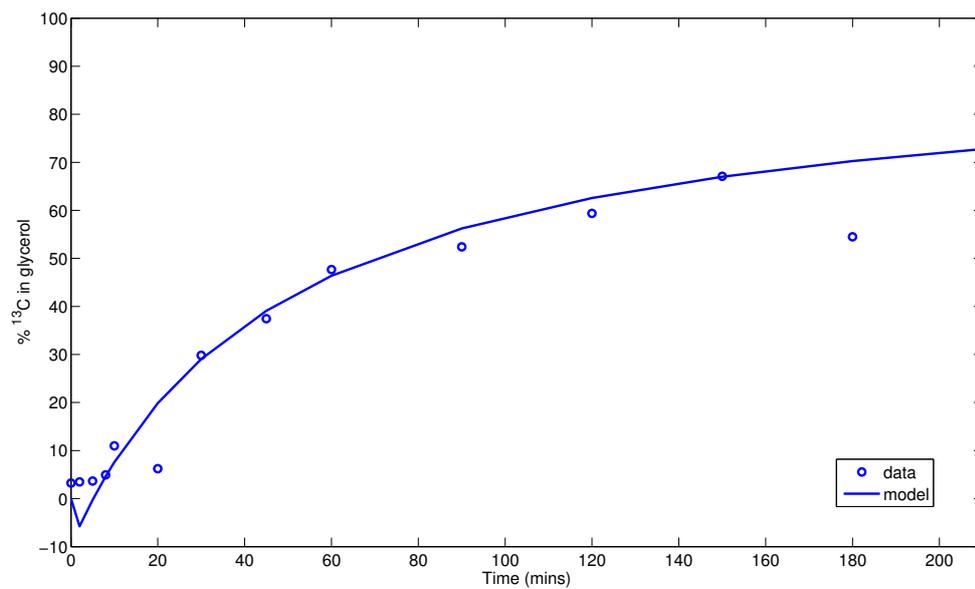


Figure C.8: Data fit for the changing  $^{13}\text{C}$  percentage in glycerol

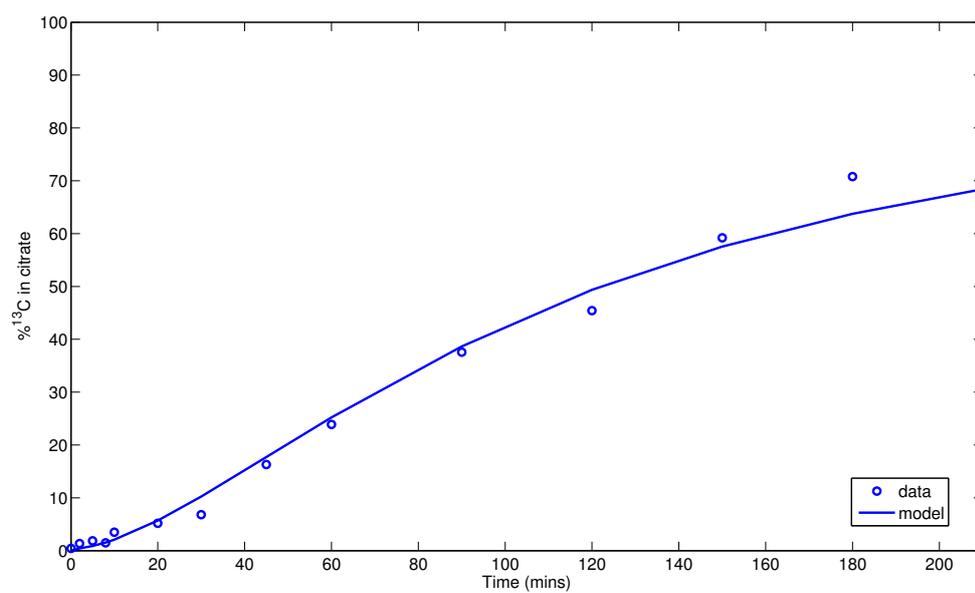


Figure C.9: Data fit for the changing  $^{13}\text{C}$  percentage in citrate