Nonstationary Metabolic Flux Analysis (NMFA) for the Elucidation of Cellular Physiology

by

Jason L. Walther

B.S., Stanford University (2004)
M.S., Stanford University (2004)

Submitted to the Department of Chemical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

April 2010
[June 2010]

Author .....
Department of Chemical Engineering
April 20, 2010

Certified by .....
Gregory Stephanopoulos
W. H. Dow Professor of Chemical Engineering
Thesis Supervisor

Accepted by .....
William M. Deen
C. P. Dubbs Professor of Chemical Engineering
Chairman, Committee for Graduate Students

Nonstationary Metabolic Flux Analysis (NMFA) for the Elucidation of Cellular Physiology

by

Jason L. Walther

Submitted to the Department of Chemical Engineering
on April 20, 2010, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Chemical Engineering

Abstract

Many current and future applications of biological engineering hinge on our ability to measure, understand, and manipulate metabolism. Many diseases for which we seek cures are metabolic in nature. Small-molecule biomanufacturing almost always involves metabolic engineering. Biofuels, a current topic of great interest, is essentially a metabolic problem. Even bioprocesses that involve complex products, such as enzyme or antibody manufacturing, still rely on a healthy and optimal metabolism and can benefit from a greater understanding therein.

A cell's metabolic flux distribution has been proposed to be one of the most solid and meaningful indicators and descriptors of metabolism. Metabolic fluxes represent integrative information and are a function of gene expression, translation, posttranslational modifications, and protein-metabolite interactions. Metabolic flux analysis (MFA) is a powerful method for determining these flux distribution through a cellular reaction network. However, MFA has experimental limitations (most notably, a requirement for isotopic steady state) that restrict the scope of biological contexts in which it can be applied.

Nonstationary metabolic flux analysis (NMFA) has recently emerged as a combined computational and experimental method that improves upon MFA with the capacity to estimate fluxes even during periods of isotopic transience in metabolism, allowing flux analysis to be applied in a broader range of experimental settings. In this thesis, we have developed and applied robust and efficient NMFA tools and techniques and applied them to understand various cellular physiologies.

We built a software package (MetranCL) that combines the elementary metabolite unit (EMU) framework, a new network decomposition strategy termed block decoupling, and a customized differential equation solver. MetranCL performs flux estimations as much as 5000 times faster than the previous state-of-the-art NMFA methods, opening entirely new types of biological systems to the possibility of flux analysis.

We applied MetranCL to a simulated large network representing *E. coli* metabolism and were able to successfully estimate reaction fluxes and metabolite concentrations

and their 95% confidence intervals. We investigated a number of different experimental arrangements of measurement time points, and found that in general, measurements earlier in isotopic transience were more sensitive to network parameters and yielded more precise confidence intervals. We also observed that the addition of concentration measurements significantly increased estimate quality.

We next used NMFA to compute fluxes from actual experimental measurements taken from brown adipocytes. We designed an appropriate network and successfully fitted simulated measurements to actual measurements (taken at 2, 4, and 6 hours after introducing tracer). A flux distribution was obtained that indicated a high level of pyruvate cycling, a low flux through the TCA cycle, and high lactate production.

We developed computational and experimental tools to assist with the design of flux analysis experiments. We built a simulator that calculates the effect of different tracers on flux estimate precision and used it to study a range of different glucose and glutamine tracers in carcinoma metabolism. Of all the stand-alone tracers we tested, we found that [1,2-$^{13}C_2$]glucose estimated flux distributions with the greatest precision. We built upon this work by constructing an evolutionary algorithm to generate optimal tracer mixtures for different organisms and their respective metabolisms. We applied this algorithm to the same cancer network and found optimal tracer mixtures for the system. We ran experiments with an optimized tracer mixture and compared it to results from typical tracers and saw significant improvements in flux precision.

Finally, we applied these methods and tools to evaluate and understand the flux distribution and metabolism of a lipid-overproducing strain of the yeast *Yarrowia lipolytica*. Since NMFA of this organism required metabolite extracts taken at very precise and proximate time points, we built a rapid sampling apparatus to draw and quench samples of *Yarrowia* cell culture with a one-second time step. After conducting NMFA under different environmental conditions and at different stages of growth, we found that lipid synthesis fluxes increased when aeration of the cell culture was increased, and observed several corresponding changes in the intracellular flux distribution explaining the overall change in metabolism that occurs with this shift in environmental conditions. In particular, we found that *Yarrowia* primarily powers lipid production by regulating flux through the pentose phosphate pathway.

Thesis Supervisor: Gregory Stephanopoulos
Title: W. H. Dow Professor of Chemical Engineering

Acknowledgments

The more work I have put into my research, especially as this thesis has taken shape and my time at MIT has drawn to a close, the more the I have realized how diffusely the credit must be spread for anything I might claim to have accomplished here. I've been enormously blessed by the people in my life, and behind every success or strength that I might initially think of as my own stands one (or more) of them.

I couldn't have asked for a better advisor than Greg Stephanopoulos. He is a good scientist and a good man. His enthusiasm, optimism, and ideas have kept me solidly on track. He has always pushed me to do the best research I am capable of. I am also grateful for the members of my thesis committee, Joanne Kelleher, Dave Kraynie, and Charlie Cooney and for their consistent support from my initial meeting with them back in 2005 all the way through to the present.

I have had the good fortune to work closely with some brilliant and engaging people. Jamey Young was the mastermind behind MetranCL, and without his guidance I would still be grappling with the basics of flux analysis. Chapters 2 through 4 were written in close collaboration with Jamey. Maciek Antoniewicz, the inventor of EMU decomposition and one of the premier experts in flux analysis today, also played an influential role in my educational upbringing during our time in the lab together. Christian Metallo asked the original questions that led to our research in tracer evaluation and optimization, and just as importantly, he was the wizard of mammalian cell culture behind all of the experimental carcinoma research. Chapters 6 and 7 were written in partnership with Christian. The brown adipocyte flux analysis in Chapter 4 is based upon experimental results obtained by Hyuntae Yoo. I owe Hussain Abidi and Mitchell Tai for their freely shared advice and experience regarding *Yarrowia lipolytica* upon which I heavily relied for the experiments conducted in Chapter 8. Early in my graduate career, I was also able to work with Joel Moxley, Mark Styczynski, and Lily Tong on some metabolomic research. I learned a lot from them and I enjoyed every minute of it.

I have had a wonderful time in the Stephanopoulos lab. Part of that is due, of

course, to the interesting research, but just as much is due to the amazing people I've been around every day. Besides those I've already mentioned, I'm especially grateful for Keith Tyo, Christine Santos, and Adel Ghaderi, our esteemed line of safety coordinators that always made sure I wore my coat and safety glasses in the lab. Ben Wang and Hang Zhou forever have my respect for going above and beyond the call of duty to keep our lab running smoothly. I've enjoyed my time and friendship with Curt Fisher, Simon Carlsen, Vikram Yadav, Deepak Dugar, Hal Alper, and so many others. To all of you: Thanks!

And last and most comes family, that amazing class of people who know me so well and still, for some reason, love me anyways. I have learned so much from watching and listening to my parents. They taught me to value work, set and pursue goals, and to have high expectations for myself. I'm a lucky man to have in my life my three fantastic boys, Matthew, Ben, and Joseph. I'm already proud of the choices they make and the people they are. And finally, I'm so thankful for my beautiful and wonderful wife, Anna. She truly is "one of the great ones". She's also the reason that I enjoy every single day of my life. It is an understatement when I say that this thesis is more hers than mine. Anna, I love you!

For Matthew, Ben, and Joseph[1]

And most of all for Anna

[1]And whomever else may follow!

Contents

List of Figures

List of Tables

Chapter 1

Introduction to Metabolic Flux Analysis

1.1 Assaying Cell Metabolism

Our ability to genetically mold and manipulate cellular systems has grown dramatically in the past few decades. Many different techniques have been developed, including targeted gene knockouts [5, 6], overexpressions [4], transcription factor engineering [8, 9], RNAi [54], foreign gene insertion [36], and protein engineering [153, 163]. General strategies such as rational design [15] and directed evolution [100] have also played important roles. But as we create new cellular genotypes in pursuit of various goals, it is essential that we are able to accurately measure and understand the corresponding new phenotypes. In a wide range of applications, the most important of these phenotypic aspects is metabolism, the set of chemical reactions that occur within a living organism [135]. Metabolism is in many ways a cumulative, culminating signal of the many upstream bioprocesses (genetic, transcriptional, translational, and kinetic) leading toward it [145].

Many current and future applications of cellular engineering hinge on metabolism. Many diseases for which we seek cures are metabolic in nature [58, 85]. Small-molecule biomanufacturing almost always relies on metabolic engineering [7]. Biofuels, a current topic of great interest, is a metabolic problem [86, 137]. Even bioprocesses that

involve biologically complex products, such as enzyme or antibody manufacturing, still rely on a healthy and optimal metabolism and can benefit from a greater understanding there [14, 59]. A thorough knowledge of metabolism will aid on both "sides" of the process to adapt and alter cellular function. Understanding metabolism before intervention reveals how and where to implement change, and understanding metabolism afterwards allows us to see if and to what degree we have succeeded.

Though important, measuring metabolism is difficult. And measuring metabolism without perturbing it in the process is especially difficult. Various strategies have been created over the years, each with strengths and weaknesses. Measuring enzyme kinetics is difficult to carry out *in vivo* and on a global scale across many different enzymes. Studies into transcriptional profiling continue to show that, while useful, mRNA abundance simply does not equal protein function [49]. Proteomics methods could potentially bypass this concern and yield direct measurements of intracellular enzymes, but concentrations sensitivities are currently limiting, and cost per analysis is still relatively high [13, 45]. Metabolomics, the measurement of intracellular metabolite pool sizes, takes a broad snapshot of metabolism at a given moment; however, metabolite pool sizes are often both difficult to measure accurately and also are not always directly linked to biological insight [51, 52, 53].

Metabolic flux distributions have been proposed to be one of the most solid and meaningful indicators and descriptors of cellular metabolism [136]. Metabolic fluxes represent integrative information and are a function of gene expression, translation, posttranslational protein modifications, and protein-metabolite interactions [104]. Elucidation of fluxes within a cell is a difficult task. Extracellular analysis of metabolic byproducts and rates allows us to gauge some cellular fluxes; unfortunately, the vast majority of metabolic happenings are still left hidden and unknown within the cell. Flux balance analysis allows for estimation of an entire metabolic network of fluxes simultaneously; however, it usually must rely on assumptions in order to solve underdetermined systems of equations [123, 146, 149]. In the late 1990s, a novel, powerful method emerged that combined measurement and modeling to estimate intracellular flux distributions. This method was termed metabolic flux

analysis.

1.2 Metabolic Flux Analysis via Isotopic Labeling

Metabolic flux analysis (MFA) is a relatively new method that uses measurements of isotopic labeling to measure intracellular fluxes for the reactions in a metabolic network. Figure 1-1 explains the structure of a flux analysis experiment. First, a tracer study is conducted in which substrate labeled with stable isotopes (usually ^{13}C) is introduced to the cells of interest while they are in a metabolic steady state. The labeled atoms spread throughout the intracellular metabolites in patterns which are sensitive to the relative fluxes of the various network reactions. Metabolic byproducts are extracted from the culture and labeling measurements are obtained.

Second, we construct a mathematical model of the metabolism. The model requires for each reaction (1) a flux value and (2) the atom transitions that occur from reactant to product. With this information, the same labeling patterns measured in the actual experiment can now be simulated *in silico*. This measurement simulation is known as the "forward problem" of metabolic flux analysis.

To estimate fluxes, we repeatedly solve the forward problem, adjusting the flux distribution at each iteration in order to minimize the lack of fit observed between the experimental and simulated labeling measurements. This strategy of using measurements to get to fluxes is known as the "inverse problem", and the set of fluxes that successfully minimizes this lack of fit represents the solution to that problem.

1.3 Brief History of Metabolic Flux Analysis

The origins of MFA can be traced to early studies that used rudimentary equations to calculate a handful of key fluxes from ^{13}C NMR [93] and GC/MS measurements [77]. Network modeling and a general mathematical approach were later introduced when fluxes generated a stoichiometric analysis were shown to correctly simulate ^{1}H NMR data [166]. Although most early flux analysis relied upon NMR measurements,

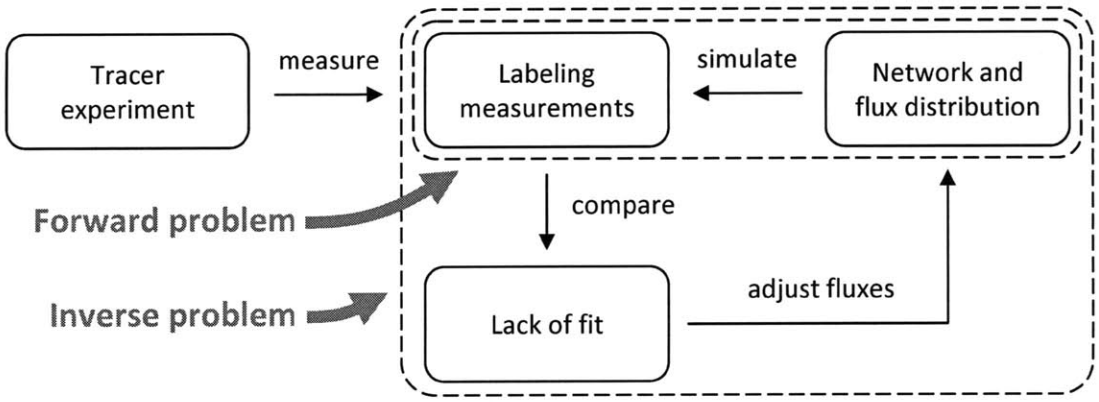

Figure 1-1: Diagram of the structure of a typical MFA experiment. First, an experiment is conducted in which isotopic tracer is introduced to cells and labeling of metabolic byproducts is measured (usually by MS or NMR). A model of cellular reactions and the atom transitions comprising those reactions can be constructed and used to simulate the labeling of those same byproducts measured experimentally. This is the forward problem. By repeatedly solving the forward problem, each time varying the model's flux distribution, we can eventually minimize the lack of fit between measurement and simulation, arriving at a set of estimated intracellular and extracellular fluxes. This iterative procedure that extracts fluxes from measurements is the inverse problem.

the majority of experiments gradually transitioned over to GC/MS [34, 37, 38, 63] and eventually LC/MS [80, 106, 148].

In the late 1990s, the simulation of labeling measurements was incorporated into parameter-fitting optimization schemes to estimate fluxes from NMR data [94, 128]. An important advancement occured when Wiechert *et al* showed that the nonlinear system of equations representing isotopic labeling fractions could be decomposed into a cascaded system of linear equations [99, 157]. Antoniewicz *et al* introduced a major breakthrough with the elementary metabolite unit (EMU) framework, an even more efficient and elegant method for decomposing the simulation problem into minimally sized systems of equations [11].

MFA was first applied extensively to microbial systems. Flux analysis has been used to great effect in showing the impact of genetic manipulations on metabolism. For instance, the rerouting of *E. coli's* carbon through PEP carboxylase and malic enzyme in response to a pyruvate kinase has been demonstrated by metabolic flux analysis of [U-$^{13}C_6$]glucose experiments [47]. Analogous experiments were also later conducted for phosphoglucose isomerase and glucose-6-dehydrogenase knockouts in *E. coli* [74].

MFA has also been used to study and improve microbial substrate utilization for industrial production; examples include the growth of *A. nidulans* on glucose versus xylose and lysine production of *C. glutamicum* using glucose versus fructose [39, 79]. Other studies have focused on *E. coli* production of compounds such as 1,3-propanediol or amorphadiene [12, 139].

Metabolic flux analysis has also been used to study relatively unknown metabolic phenomena. MFA has produced information and understanding in relatively unknown pathways in well-known organisms (such as the glucose oxidation cycle of *E. coli*) and it also has served as a useful probe of metabolism in novel organisms such as *Shewanella oneidensis*, *Geobacter metallireducens*, *Desulfovibrio vulgaris*, and *Actinobacillus succinogenes* [97, 120, 140, 141]. Flux analysis is an ideal tool for these investigations since genetic understanding is lacking but reaction network information can be inferred from similar organisms.

Metabolic flux analysis has been increasingly applied to plant and mammalian cell metabolism. Studies of *C. roseus* roots and soybeans have estimated flux values for reactions occurring in multiple-compartment networks [3, 133]. Industrial CHO cells have been analyzed in perfusion culture [65]. Several medical applications have been found as well; studies of breast cancer cells have suggested potential pathway targets for cancer therapy [57], while flux profiling of mammalian cells infected by HCMV have suggested potential targets for antiviral therapy [102].

1.4 Nonstationary Metabolic Flux Analysis

It is clear that metabolic flux analysis has found a broad array of applications, and its deployment has become more sophisticated over time. However, there are still many biological instances in which standard MFA is too limited to be useful. All of the previously mentioned MFA experiments rest upon two major assumptions. First, the system under study must be in a metabolic steady state (i.e., fluxes and metabolite concentrations are constant with respect to time). Second, the system must be allowed to come to an isotopic steady state after the initial introduction of labeled substrate (i.e., all intracellular metabolite labeling patterns are constant with respect to time). These requirements constrain experiments and sometimes limit the scope and usefulness of metabolic flux analysis.

Nonstationary metabolic flux analysis (NMFA), the topic of this thesis, is similar to MFA with the provision that metabolite labeling can be sampled and measured during the transient period before the system comes to an isotopic steady state. This gives researchers more freedom and flexibility in experimental design. NMFA offers significant advantages as compared to MFA:

1. NMFA experiments are much less costly (in terms of both time and money) since one does not need to wait for isotopic steady state to be established [106].

2. NMFA is particularly suited for systems that cannot be held at a metabolic steady state indefinitely (e.g., primary cells or animal studies) because experi-

mental durations are greatly reduced.

3. In some cases, NMFA identifies metabolic fluxes with greater precision because some isotopically transient measurements have greater sensitivities to fluxes [107].

4. In some cases, NMFA measurement data can be used to estimate metabolite concentrations in addition to fluxes.

5. ^{13}C NMFA can successfully estimate fluxes in systems that rely solely on single-carbon substrates (e.g., photoautotrophs and methylotrophs) whereas at isotopic steady state, metabolites are uniformly labeled and no new information is generated by a ^{13}C tracer [131].

However, NMFA also introduces several complexities both in computation and in experimentation. (For example, differential equations must be solved and samples must be taken rapidly within the duration of isotopic transience.) The resolution of these challenges and the successful implementation of NMFA is the goal of this thesis.

1.5 Thesis Outline

This thesis covers several aspects of NMFA, starting with the general theory and algorithms behind nonstationary simulation and estimation (Chapter 2). We move on and report some initial applications of NMFA (Chapters 3 and 4) after which we discuss topics in NMFA experimental design (Chapters 5, 6, and 7). We conclude with a rigorous NMFA experiment leading to biological insight in an important experimental system (Chapter 8).

- **NMFA theory and computation (Chapter 2):** We applied elementary metabolite unit (EMU) theory to nonstationary flux analysis, dramatically reducing computational difficulty. We also introduced block decoupling, a new method that systematically and comprehensively divides EMU systems of equations into smaller subproblems to further reduce computational difficulty. These

improvements led to a 5000-fold reduction in simulation times, enabling an entirely new and more complicated set of problems to be analyzed with NMFA. We capped our theoretical work by developing a software package (MetranCL) that uses these new methods for measurement simulation and flux estimation.

- **NMFA of a simulated large *E. coli* network (Chapter 3):** We constructed a large, biologically realistic network representing *E. coli* metabolism. This network's size would normally render NMFA infeasible, but using our new EMU-based methods, the problem was tractable. We simulated a series of nonstationary and stationary GC/MS measurements for the network that was then used to estimate parameters and their associated confidence intervals. We found that fluxes could be successfully estimated using only nonstationary labeling data and external flux measurements. The addition of concentration measurements increased the precision of most parameters.

- **NMFA of brown adipocytes (Chapter 4):** We also applied EMU-based NMFA to experimental nonstationary measurements taken from brown adipocytes and successfully estimated fluxes and some metabolite concentrations. Adipocyte metabolism provides an important window into many metabolic diseases, especially those that are related to obesity, such as diabetes. By using NMFA instead of traditional MFA, the experiment required only 6 hours instead of 50 (the time necessary for most metabolite labeling to reach 99% of isotopic steady state). In our results, we observed a large pyruvate recycle flux, a small TCA cycle flux, and a large lactate flux.

- **Rapid sampling (Chapter 5):** Organisms with highly active metabolisms have short periods of isotopic transience. To accurately measure metabolite labeling in this period, rapid sampling must be employed. We developed a vacuum-powered rapid sampler capable of taking samples on a time scale of seconds. We built the sampler so that it would be compatible with a variety of bioreactors and flasks, while maintaining a low threshold of construction by keeping the apparatus as inexpensive and as simple as possible. We showed that

the rapid sampler measurements could generate isotopically consistent measurements (of a mixture of glucose isotopomers) and isotopically dynamic measurements (of intracellular pyruvate in *Yarrowia lipolytica*).

- **Isotopic tracer evaluation for metabolic flux analysis (Chapter 6):** Tracer selection is an important and easily adjustable parameter in NMFA experiments. As such, tracer choice is a prime candidate for manipulation in experimental design. Using our NMFA software, we computationally evaluated specifically labeled ^{13}C glucose and glutamine tracers for their ability to precisely and accurately estimate fluxes in the central carbon metabolism of carcinoma cells. These methods enabled us to identify the optimal tracer for analyzing individual fluxes, specific pathways, and central carbon metabolism as a whole. These results provide valuable, quantitative information on the performance of ^{13}C-labeled substrates and can aid in the design of more informative MFA experiments in mammalian cell culture. In particular, we found that [U-$^{13}C_6$]glutamine was the best tracer for ascertaining TCA cycle fluxes while [1,2-$^{13}C_2$]glucose was the ideal tracer for glycolysis, the pentose phosphate pathway, and the overall network.

- **Optimization of isotopic tracer mixtures for metabolic flux analysis (Chapter 7):** Tracers need not be used in isolation in flux analysis; they can be combined in different proportions to improve estimate precision. To our knowledge, no systematic approach exists for searching the space of tracer mixtures for experimental design. To that end, we created a strategy that finds an optimal mixture of tracers for a given metabolic network and a given set of potential tracers. We use a genetic algorithm to search the space of possible tracer mixtures, and select for and recombine those mixtures that maximize the precision with which the flux distribution is estimated. We applied this algorithm to carcinoma metabolism and found two optimal tracer mixtures. We then experimentally applied one of these tracer mixtures to a culture of carcinoma cells and saw a corresponding improvement in flux precision, validating

our genetic algorithm and tracer evaluation strategy.

- **NMFA of *Yarrowia lipolytica* (Chapter 8):** *Y. lipolytica* is an oleaginous (lipid-producing) yeast with great promise in biofuels applications. We studied a lipid-overproducing strain using nonstationary flux analysis. We measured extracellular fluxes and conducted NMFA under different experimental conditions using a combination of rapid and manual sampling. We obtained flux distributions during the late growth phase and the stationary phase in bioreactor conditions and in conditions with high aeration and oxygenation in order to better understand the effect of oxygen availability on *Y. lipolytica's* production of fatty acids. We determined that lipid production increases with greater aeration, and that the cells manage the energy burden of the high production by increasing flux through the pentose phosphate pathway.

- **Recommendations for Future Research (Chapter 9):** Dynamic metabolic flux analysis is the next step in the progression of flux analysis and deserves study in future research. The ability to measure fluxes in metabolically unstable systems. MetranCL also can use further development to make it more accessible and more powerful. Some of these specific areas of improvement include the creation of a graphical user interface and improved parallelization. Network sensitivity analysis is a third direction that could yield valuable fruit as we seek to better measure fluxes.

Chapter 2

NMFA Theory and Computation

2.1 Introduction

MFA and NMFA are concerned with solving an "inverse problem" in which fluxes (and in the case of NMFA, concentrations) are estimated from metabolite labeling distributions by means of an iterative least-squares fitting procedure. At each iteration, a "forward problem" must be solved in which metabolite labeling distributions are simulated for a given metabolic network and a given set of parameter estimates. The mismatch between the simulated and experimental measurements is assessed and the parameter estimates are updated to achieve an improving fit.

In the context of MFA, the forward problem can be represented by systems of linear algebraic equations. NMFA, on the other hand, requires the solution of systems of ordinary differential equations, a significantly more difficult task. This additional complexity means that the algorithms for NMFA must be carefully designed so that the computational expense for large metabolic networks does not become prohibitive. Currently, state-of-the-art algorithms (using cumomer fractions as state variables) require more than an hour to simulate isotopic labeling of a realistic network model [107, 157].

In this chapter, we propose a new approach based upon the Elementary Metabolite Unit (EMU) framework [11] that efficiently and robustly handles the inverse problem of NMFA by solving the forward problem thousands of times faster than currently

available methods. Because of these improvements in the NMFA model, we are able to show for the first time that fluxes and concentrations can be estimated from non-stationary data for realistically sized metabolic networks in short amounts of time.

2.2 EMU Network Decomposition

The nonstationary treatment presented here is built using the mass isotopomer distributions (MIDs) of elementary metabolite units (EMUs) as state variables [11]. An EMU is defined as a distinct subset of a metabolite's atoms. EMUs can exist in a variety of mass states depending on their isotopic compositions. An EMU in its lowest mass state is referred to as M+0, while an EMU that contains one additional atomic mass unit (e.g., due to the presence of a ^{13}C atom in place of a ^{12}C atom) is referred to as M+1, with higher mass states described accordingly. An MID is a vector that contains the fractional abundance of each mass state of an EMU.

The goal of an NMFA simulation is the calculation of metabolite labeling patterns that are measurable by mass spectroscopy; i.e., the MIDs of a certain subset of EMUs in the system. While the total number of all possible EMUs in a network is equal to the number of isotopomers or cumomers, in most cases only a small fraction of EMUs is required to simulate measurable MIDs.

EMUs of metabolites in a common reaction network can be assembled into an analogous EMU network composed of EMU reactions where the MIDs of upstream EMUs affect the MIDs of downstream EMUs. Often, EMU networks can be decoupled into separate and smaller subnetworks. Decoupling of EMU reactions based on (1) EMU size and (2) network connectivity has been discussed previously [11]. (EMU size is defined as the number of atoms comprising a particular EMU.) Because MIDs of EMUs depend only upon MIDs of equally sized or smaller EMUs, the EMU network can be partitioned into size-based networks, each containing equally sized EMUs and depending on inputs only from smaller-sized EMUs. If smaller, completely independent EMU subnetworks can be identified within these size-based networks, further decoupling can occur. Computational costs can therefore be decreased in two ways:

first, the total size of the system can be reduced, and second, the system can be divided into smaller subsystems that cumulatively can be solved more quickly.

2.3 Block Decoupling

We propose a systematic and comprehensive method of EMU reaction network decoupling in which metabolite units are grouped into blocks. A block is defined as a set of EMUs whose MIDs are mutually dependent within the context of the EMU reaction network. Thus, by definition all EMUs within a particular block (1) are of the same size, (2) mutually approach an isotopic steady state, and (3) must be solved for simultaneously and not sequentially. Blocks can be arranged such that each is a self-contained subproblem depending only upon the outputs of previously solved blocks. This lets us work with smaller and more tractable matrices, greatly increasing computational efficiency.

To arrange EMUs into blocks, we first regard the EMU reaction network as a directed graph in which nodes represent EMUs and edges represent EMU reactions. An N-by-N adjacency matrix is then constructed for the directed graph, where N is the total number of EMUs. In short, a nonzero entry $a(i,j)$ of the adjacency matrix indicates the dependence of the ith EMU's MID on the jth EMU's MID. We then perform a Dulmage-Mendelsohn decomposition on the adjacency matrix, returning an upper block triangular matrix from which the diagonal blocks are extracted [44, 114].

2.4 Simple Network

A simple metabolic network appears in Figure 2-1A as an example. Figure 2-1B delineates the atom transitions for the network. Hypothetical metabolite C is assumed to be measurable by GC/MS. After EMU decomposition, the nonstationary system can be described in terms of 16 EMUs. This represents a 44% reduction in state variables from the 29 cumomer fractions required to simulate the system with the cumomer method. After decoupling based on EMU size and connectivity, these 16

state variables can be separated into four smaller subproblems (see Figure 2-2A).

By applying Dulmage-Mendelsohn decomposition and block decoupling to the simple network, we can achieve even further system reduction. Figure 2-3 shows this decomposition and the resulting blocks in matrix form. Block decoupling improves upon previous methods, enabling the 16 essential EMUs to be divided among eight subproblems instead of four (see Figure 2-2B). Table 2.1 provides a detailed comparison of the model reductions achieved by cumomer and EMU decompositions both with and without block decoupling.

2.5 Simulation of Metabolite Labeling

Decomposition of a network into blocks of EMUs generates a cascaded system of ordinary differential equations, where level n of the cascade represents the network of EMUs within the nth block. Each system has the following form:

$$\mathbf{C}_n \cdot \frac{d\mathbf{X}_n}{dt} = \mathbf{A}_n \cdot \mathbf{X}_n + \mathbf{B}_n \cdot \mathbf{Y}_n \tag{2.1}$$

The rows of the state matrix $\mathbf{X}_n$ correspond to MIDs of EMUs within the nth block. The input matrix $\mathbf{Y}_n$ is analogous but with rows that are MIDs of EMUs that are previously calculated inputs to the nth block. The concentration matrix $\mathbf{C}_n$ is a diagonal matrix whose elements are concentrations corresponding to EMUs in $\mathbf{X}_n$. Finally, the system matrices $\mathbf{A}_n$ and $\mathbf{B}_n$ describe the network as follows:

$$\mathbf{A}_n(i,j) = \begin{cases} -\text{sum of fluxes consuming } i\text{th EMU in } \mathbf{X}_n & i = j \\ \text{flux to } i\text{th EMU in } \mathbf{X}_n \text{ from } j\text{th EMU in } \mathbf{X}_n & i \neq j \end{cases} \tag{2.2}$$

$$\mathbf{B}_n(i,j) = \text{flux to } i\text{th EMU in } \mathbf{X}_n \text{ from } j\text{th EMU in } \mathbf{Y}_n \tag{2.3}$$

Fully written matrices $\mathbf{A}_n$, $\mathbf{B}_n$, $\mathbf{X}_n$, $\mathbf{Y}_n$, and $\mathbf{C}_n$ for all eight blocks of the simple example problem (described in Figures 2-1 and 2-2) are listed in Appendix A.

The least-squares fitting algorithm employed to solve the inverse problem requires

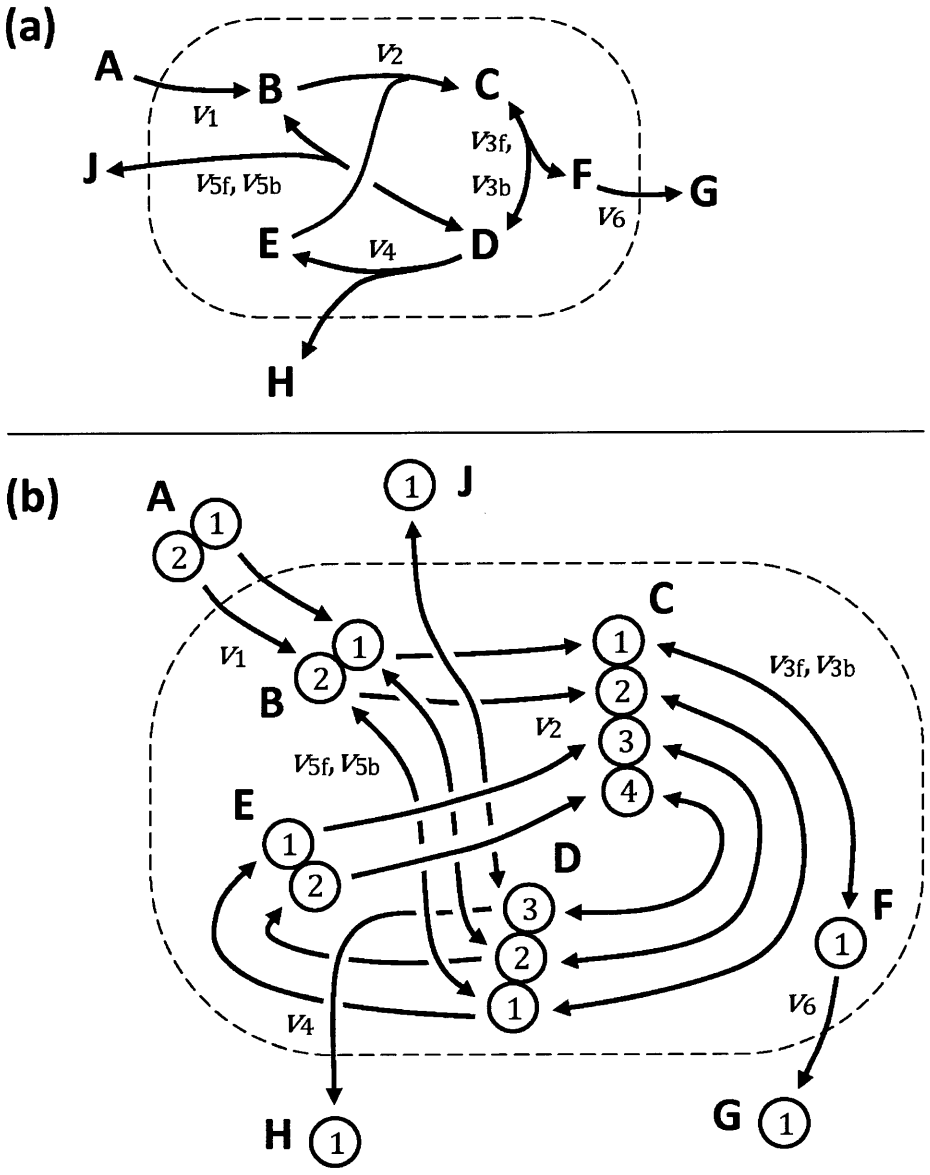

Figure 2-1: (A) A simple example network used to illustrate EMU network decomposition. The network fluxes are assumed to be constant since the system is at a metabolic steady state. Extracellular metabolites A, G, H, and J are assumed to be at a fixed state of isotopic labeling to which intracellular metabolites B, C, D, E, and F adapt over time. (B) Atom transitions for the simple example network.

Model	Cumomer	EMU	EMU
Decoupling method	Size	Size/connectivity	Blocks
(Size) # of vars	(1) 12	(1) 9	(1) 3,3,2,1
	(2) 11	(2) 4	(2) 3,1
	(3) 5	(3) 2	(3) 2
	(4) 1	(4) 1	(4) 1
Total variables	29	16	16

Table 2.1: A comparison of modeling approaches to simulate the dynamic labeling of a simple example network. EMU network decomposition followed by block decoupling minimizes the number of state variables both in the overall system and within any one subproblem. The subproblems are listed by EMU size (or in the case of cumomers, by weight). The EMU sizes (or cumomer weights) are indicated within parentheses and the number of variables within each subproblem follow. For instance, the entry "(2) 3,1" indicates that there are two subproblems involving EMUs of size 2. One subproblem contains three variables and the other only one.

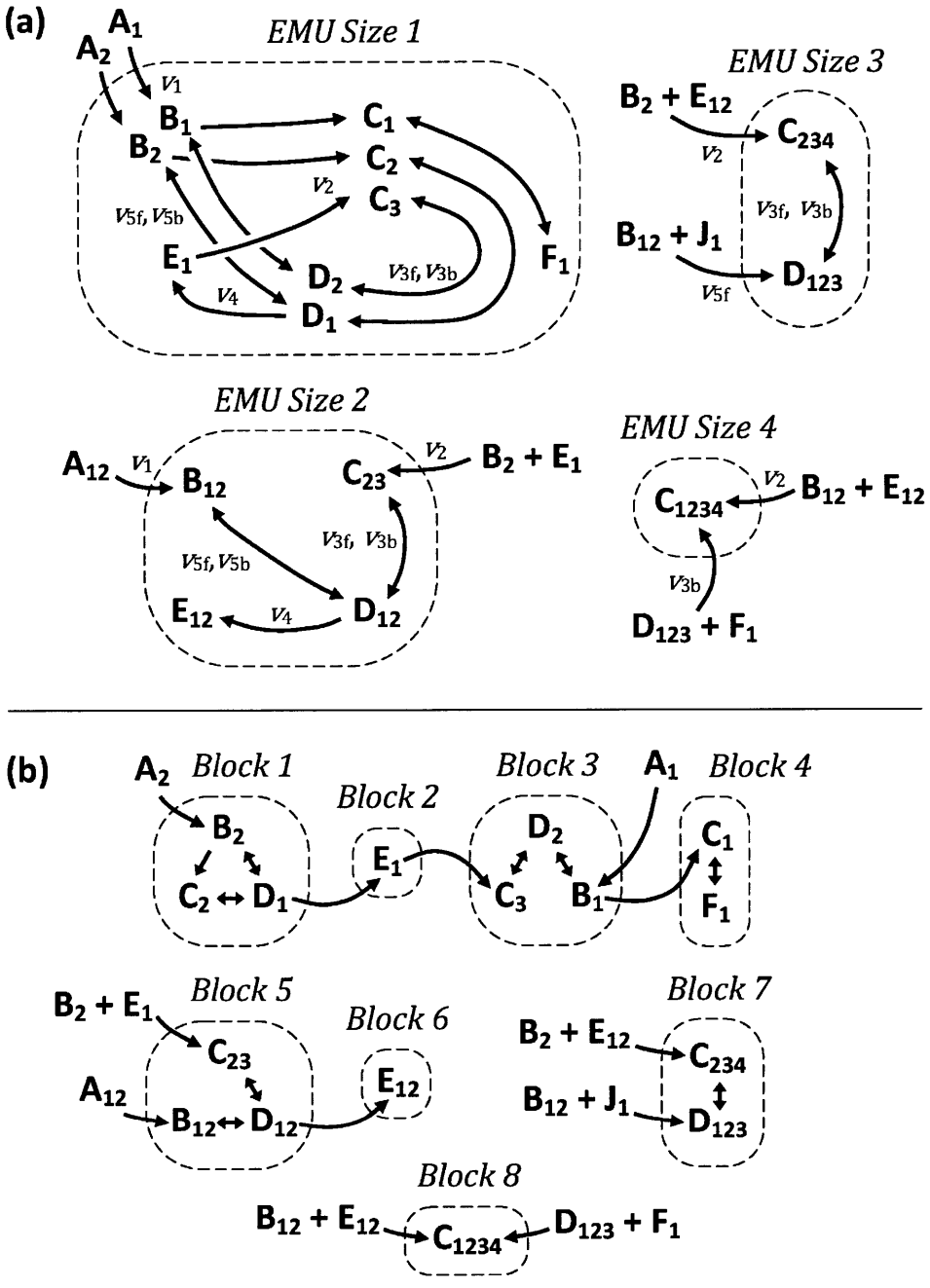

Figure 2-2: (A) EMU network decomposition for the simple example network (Figure 2-1) generated to simulate the labeling of metabolite C. The EMU reaction network was decoupled based on EMU size and network connectivity. (B) EMU network decomposition for the same network using block decoupling. Subscripts refer to the atoms of a compound that are contained within the EMU. The state and system matrices corresponding to this decomposition can be found in Appendix A.

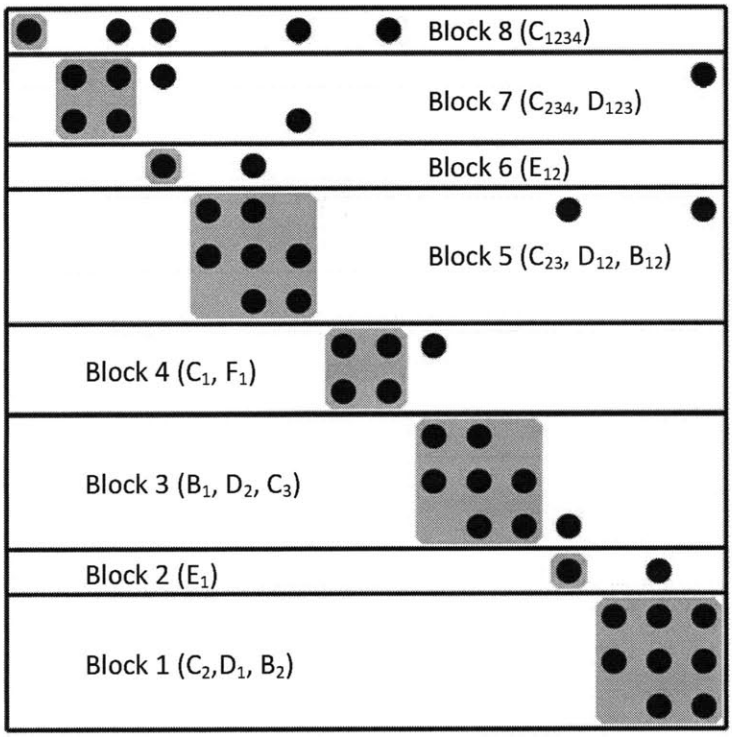

Figure 2-3: Dulmage-Mendelsohn decomposition of an adjacency matrix representing the EMU reaction network for the simple example network described in Figure 2-1. A non-zero entry (denoted by a black circle) at the ith row and jth column of the matrix represents the dependence of the ith EMU's MID on the jth EMU's MID. The upper triangular matrix resulting from the Dulmage-Mendelsohn decomposition can be separated into blocks as indicated by bold lines. Blocks can be solved in a sequential order, beginning at the lower right-hand corner and working upwards.

repeated calculation of first order derivatives, i.e., sensitivities of simulated measurements with respect to fluxes and concentrations. To this end, implicit differentiation of Equation (2.1) yields

$$\frac{d}{dt}\frac{\partial \mathbf{X}_n}{\partial \mathbf{p}} = \mathbf{C}_n^{-1} \cdot \mathbf{A}_n \cdot \frac{\partial \mathbf{X}_n}{\partial \mathbf{p}} + \frac{\partial (\mathbf{C}_n^{-1} \cdot \mathbf{A}_n)}{\partial \mathbf{p}} \cdot \mathbf{X}_n + \mathbf{C}_n^{-1} \cdot \mathbf{B}_n \cdot \frac{\partial \mathbf{Y}_n}{\partial \mathbf{p}} + \frac{\partial (\mathbf{C}_n^{-1} \cdot \mathbf{B}_n)}{\partial \mathbf{p}} \cdot \mathbf{Y}_n \quad (2.4)$$

where $\mathbf{p}$ is a vector of metabolic fluxes and concentrations.

2.6 Customized Differential Equation Solver

We integrate the system with a customized ordinary differential equation solver that discretizes Equations (2.1) and (2.4) by applying a first-order hold equivalent with adaptive step size control [115]. This method is A-stable, simple to code, and enables large time steps by making use of partial analytical solutions to the system equations.

EMU labeling states and sensitivities are described by Equations (2.1) and (2.4) and can be simplified as shown below:

$$\frac{d\mathbf{X}_n}{dt} = \mathbf{F}_n \cdot \mathbf{X}_n + \mathbf{G}_n \quad (2.5)$$

$$\frac{d}{dt}\frac{\partial \mathbf{X}_n}{\partial \mathbf{p}} = \mathbf{F}_n \cdot \frac{\partial \mathbf{X}_n}{\partial \mathbf{p}} + \mathbf{H}_n \quad (2.6)$$

by making use of the following substitutions:

$$\mathbf{F}_n = \mathbf{C}_n^{-1} \cdot \mathbf{A}_n \quad (2.7)$$

$$\mathbf{G}_n = \mathbf{C}_n^{-1} \cdot \mathbf{B}_n \cdot \mathbf{Y}_n \quad (2.8)$$

$$\mathbf{H}_n = \frac{\partial \mathbf{F}_n}{\partial \mathbf{p}} \cdot \mathbf{X}_n + \frac{\partial \mathbf{G}_n}{\partial \mathbf{p}} \quad (2.9)$$

The functions $\mathbf{G}_n$ and $\mathbf{H}_n$ potentially comprise convolutions of MIDs belonging to EMUs of previously solved blocks (a result of EMU condensation reactions) and as such Equations (2.5) and (2.6) lack analytical solutions. Partial analytical solutions, however, can be written:

$$\mathbf{X}_n(t_1) = \mathbf{e}^{\mathbf{F}_n \cdot \Delta t} \cdot \mathbf{X}_n(t_0) + \int_0^{\Delta t} \mathbf{e}^{\mathbf{F}_n \cdot (\Delta t - \tau)} \cdot \mathbf{G}_n(\tau + t_0) \cdot d\tau \tag{2.10}$$

$$\left.\frac{\partial \mathbf{X}_n}{\partial \mathbf{p}}\right|_{t_1} = \mathbf{e}^{\mathbf{F}_n \cdot \Delta t} \cdot \left.\frac{\partial \mathbf{X}_n}{\partial \mathbf{p}}\right|_{t_0} + \int_0^{\Delta t} \mathbf{e}^{\mathbf{F}_n \cdot (\Delta t - \tau)} \cdot \mathbf{H}_n(\tau + t_0) \cdot d\tau \tag{2.11}$$

where the initial state of the system at time t_0 is assumed to be known and Δt is defined as $t_1 - t_0$. Again, the integrals in Equations (2.10) and (2.11) lack analytical solutions. Instead, we evaluate at discrete points by applying a non-causal first-order-hold equivalent with adaptive step size control to numerically integrate and solve the problem [115]. This discretized approximation can be expressed as follows:

$$(\mathbf{X}_n)_{k+1} = \mathbf{\Phi}_n \cdot (\mathbf{X}_n)_k + \mathbf{\Gamma}_n \cdot (\mathbf{G}_n)_k + \mathbf{\Omega}_n \cdot [(\mathbf{G}_n)_{k+1} - (\mathbf{G}_n)_k] \tag{2.12}$$

$$\left(\frac{\partial \mathbf{X}_n}{\partial \mathbf{p}}\right)_{k+1} = \mathbf{\Phi}_n \cdot \left(\frac{\partial \mathbf{X}_n}{\partial \mathbf{p}}\right)_k + \mathbf{\Gamma}_n \cdot (\mathbf{H}_n)_k + \mathbf{\Omega}_n \cdot [(\mathbf{H}_n)_{k+1} - (\mathbf{H}_n)_k] \tag{2.13}$$

where the transition matrices $\mathbf{\Phi}_n$, $\mathbf{\Gamma}_n$, and $\mathbf{\Omega}_n$ are functions of fluxes, concentrations, and the time step magnitude according to the following relationship:

$$\begin{bmatrix} \mathbf{\Phi}_n & \mathbf{\Gamma}_n & \mathbf{\Omega}_n \\ 0 & \mathbf{I} & \mathbf{I} \\ 0 & 0 & \mathbf{I} \end{bmatrix} = \exp\left(\begin{bmatrix} \mathbf{F}_n \cdot \Delta t & \mathbf{I} \cdot \Delta t & 0 \\ 0 & 0 & \mathbf{I} \\ 0 & 0 & 0 \end{bmatrix}\right) \tag{2.14}$$

where the exponential function refers to the matrix exponential [62]. At each time point, $\mathbf{Y}_n$s and $\mathbf{X}_n$s are calculated in ascending order until the EMUs representing all desired measurements are obtained.

2.7 Flux and Concentration Estimation

Fluxes and concentrations are estimated by minimizing the difference between measured and simulated data according to the following equation [10, 107]:

$$\begin{aligned} &\min_{\mathbf{u},\mathbf{c}} \Phi = [\mathbf{m}(\mathbf{u},\mathbf{c},t) - \hat{\mathbf{m}}(t)]^{\mathrm{T}} \cdot \boldsymbol{\Sigma}_{\mathbf{m}}^{-1} \cdot [\mathbf{m}(\mathbf{u},\mathbf{c}) - \hat{\mathbf{m}}] \\ &\text{s.t. } \mathbf{N} \cdot \mathbf{u} \geq 0, \ \mathbf{c} \geq 0 \end{aligned} \tag{2.15}$$

where Φ is the objective function to be minimized, $\mathbf{u}$ is a vector of free fluxes, $\mathbf{c}$ is a vector of metabolite concentrations, t is time, $\mathbf{m}(\mathbf{u},\mathbf{c},t)$ is a vector of simulated measurements, $\hat{\mathbf{m}}(t)$ is a vector of observed measurements, $\boldsymbol{\Sigma}_{\mathbf{m}}$ is the measurement covariance matrix, and $\mathbf{N}$ is the nullspace of the stoichiometric matrix. We have implemented a reduced gradient method to handle the linear constraints of this problem within a Levenberg-Marquardt nonlinear least-squares solver [60, 90].

Calculation of parameter standard errors requires the inverse Hessian of Φ, which becomes ill-conditioned when some parameters are poorly identifiable. Because the Hessian is obtained by numerically integrating the measurement sensitivities in Equation (2.4), it is contaminated by numberical errors in the nonstationary case. Upon matrix inversion, even small errors can greatly distort standard error estimates, rendering them nearly meaningless. As such, we compute nonlinear flux confidence intervals (using parameter continuation around the optimal solutions) instead of relying upon local standard errors [10]. These confidence intervals, though more computationally expensive to obtain than local standard errors, yield a significantly more reliable and realistic description of the true parameter identity.

2.8 MetranCL Software

These algorithms were incorporated into a software package written in Matlab and named MetranCL (an abbreviation for "Command Line Metabolic Tracer Analysis"). Users can create network models, simulate labeling data, estimate flux and concentra-

tion data, generate parameter confidence intervals, and visualize results using different tools within MetranCL. Operational details can be found in Appendix B.

To compare the performance of our approach to prior methods, we reconstructed the simplified *E. coli* model described by Nöh consisting of 28 free fluxes and 16 metabolite pools [107]. Application of the EMU-based algorithm to this system using MetranCL leads to a 5000-fold reduction in the computational time required for simulation of the forward problem (from 83 minutes on an AMD Opteron 2000+ down to one second on a 2.0 GHz T2500 dual core processor). Whereas computational time for parameter estimation via cumomers was conjectured to be 24-48 hours, we estimated fluxes and concentrations in less than one minute, beginning from a randomized set of initial parameters.

2.9 Discussion

The application of the EMU framework to NMFA results in dramatic improvements in network decomposition and parameter estimation. These advances make entirely new realm of problems in nonstationary flux analysis feasible. For instance, previous analysis of systems with complicated reaction networks, multiple isotopic tracers, or large molecules were impractical targets for NMFA. By shifting to an EMU framework, these kinds of problems are now tractable. The EMU framework also makes possible the calculation of accurate confidence intervals for parameters estimated by NMFA, a computationally intensive exercise that otherwise would be infeasible.

Chapter 3

NMFA of a Large Simulated *E. coli* Network

3.1 Measurement Timing and Estimation Quality

We are interested in the effect of measurement timing (during the isotopically non-stationary period of a flux analysis experiment) on flux estimation quality. Labeling measurements at the early time points of isotopic transience have been shown to be more sensitive to (and hence better estimators of) certain fluxes [107], while late measurements' lack of sensitivity to metabolite concentrations means that less parameters must be included in the optimimzation, resulting in generally narrower confidence intervals overall. Because of these two conflicting principles and because of the nonlinearity and complexity of most metabolic networks, it is a nontrivial task to find a general set of measurement time points that optimizes flux estimation quality (by minimizing confidence intervals) for any given tracer experiment. We used a simulated metabolic network to explore the relationship between measurements and estimation quality for a single experiment and from our results extracted some potential underlying points.

3.2 Experimental Design

Because of the increased efficiency of EMU-based NMFA, we were able to apply our method to a larger and more realistic *E. coli* network. Specifically, we modeled the central metabolism of a strain capable of producing high levels of 1,3-propanediol (PDO) using a network that includes 35 free fluxes and 46 metabolite pools [12]. A complete list of reactions and atom transitions is available in Tables 3.1 and 3.2. The size of this problem can be reduced by over 90% via EMU decomposition (relative to isotopomer or cumomer decomposition) and can be further parsed into 47 subproblems with block decoupling (compared to only 14 with decoupling by size and network connectivity). Block decoupling led to a 27% decrease in computational time relative to decoupling based only upon size and connectivity. Table 3.3 provides a detailed comparison of the model reductions achieved by cumomer and EMU decompositions both with and without block decoupling. Further details can be found in Tables A.1 through A.3 of Appendix A, where we list all EMUs participating in the decomposed network, and break them into their respective decoupled blocks.

To investigate the relationship between sampling times and parameter identifiability, we generated a series of simulated data sets. We drew flux values from a previously published stationary MFA experiment involving the aforementioned PDO-producing strain [12] and metabolite concentration values from various literature sources on both *E. coli* and *S. cerevisiae* [27, 63]. Five different sets of measurements were simulated:

1. **Stationary experiment:** Measurements were conducted at a time sufficiently large such that all metabolite labeling was assumed constant. Thirty replicate sets of measurements were made such that the total number of labeling measurements in all experiments was equal.

2. **Long nonstationary experiment:** One set of measurements was taken every second for 15 seconds following the introduction of tracer. For the next 75 seconds, measurements were taken every 5 seconds, giving a total of 30 sets of measurements. By the end of this period, all measured metabolite fragments were within 99% of isotopic steady state.

Glycolysis			
v_1	G6P (abcdef)	$\leftrightarrow$	F6P (abcdef)
v_2	F6P (abcdef)	$\rightarrow$	FBP (abcdef)
v_3	FBP (abcdef)	$\leftrightarrow$	DHAP (cba) + GAP (def)
v_4	DHAP (abc)	$\leftrightarrow$	GAP (abc)
v_5	GAP (abc)	$\leftrightarrow$	3PG (abc)
v_6	3PG (abc)	$\leftrightarrow$	PEP (abc)
v_7	PEP (abc)	$\rightarrow$	Pyr (abc)
Pentose Phosphate Pathway			
v_8	G6P (abcdef)	$\rightarrow$	6PG (abcdef)
v_9	6PG (abcdef)	$\rightarrow$	Ru5P (bcdef) + CO_2 (a)
v_{10}	Ru5P (abcde)	$\leftrightarrow$	X5P (abcde)
v_{11}	Ru5P (abcde)	$\leftrightarrow$	R5P (abcde)
v_{12}	X5P (abcde)	$\leftrightarrow$	GAP (cde) + EC_2 (ab)
v_{13}	F6P (abcdef)	$\leftrightarrow$	E4P (cdef) + EC_2 (ab)
v_{14}	S7P (abcdefg)	$\leftrightarrow$	R5P (cdefg) + EC_2 (ab)
v_{15}	F6P (abcdef)	$\leftrightarrow$	GAP (def) + EC_3 (abc)
v_{16}	S7P (abcdefg)	$\leftrightarrow$	E4P (defg) + EC_3 (abc)
Entner-Doudoroff Pathway			
v_{17}	6PG (abcdef)	$\rightarrow$	KDPG (abcdef)
v_{18}	KDPG (abcdef)	$\rightarrow$	Pyr (abc) + GAP (def)
Citric Acid Cycle			
v_{19}	Pyr (abc)	$\rightarrow$	AcCoA (bc) + CO_2 (a)
v_{20}	OAA (abcd) + AcCoA (ef)	$\rightarrow$	Cit (dcbfea)
v_{21}	Cit (abcdef)	$\leftrightarrow$	ICit (abcdef)
v_{22}	ICit (abcdef)	$\leftrightarrow$	AKG (abcde) + CO_2 (f)
v_{23}	AKG (abcde)	$\rightarrow$	SucCoA (bcde) + CO_2 (a)
v_{24}	SucCoA (abcd)	$\leftrightarrow$	1/2 Suc (abcd) + 1/2 Suc (dcba)
v_{25}	Suc (abcd)	$\leftrightarrow$	1/2 Fum (abcd) + 1/2 Fum (dcba)
v_{26}	Fum (abcd)	$\leftrightarrow$	1/2 Mal (abcd) + 1/2 Mal (dcba)
v_{27}	Mal (abcd)	$\leftrightarrow$	OAA (abcd)
Anaplerotic Reactions			
v_{28}	Mal (abcd)	$\rightarrow$	Pyr (abc) + CO_2 (d)
v_{29}	PEP (abc) + CO_2 (d)	$\leftrightarrow$	OAA (abcd)
Acetic Acid Formation			
v_{30}	AcCoA (ab)	$\leftrightarrow$	Ac (ab)
PDO Biosynthesis			
v_{31}	DHAP (abc)	$\leftrightarrow$	Glyc3P (abc)
v_{32}	Glyc3P (abc)	$\rightarrow$	Glyc (abc)
v_{33}	Glyc (abc)	$\rightarrow$	HPA (abc)
v_{34}	HPA (abc)	$\rightarrow$	PDO (abc)
One Carbon Metabolism			
v_{35}	MEETHF (a)	$\rightarrow$	METHF (a)
v_{36}	MEETHF (a)	$\rightarrow$	FTHF (a)

Table 3.1: A list of reactions and atom transitions within the *E. coli* network for glycolysis, the pentose phosphate pathway, the Entner-Doudoroff pathway, the citric acid cycle, amphibolic reactions, acetic acid formation, PDO biosynthesis, and one-carbon metabolism. Carbon atom transitions are indicated within parentheses. Irreversible and reversible reactions are indicated by the symbols $\rightarrow$ and $\leftrightarrow$, respectively.

Transport			
v_{37}	$Gluc_{pre}$ (abcdef)	→	G6P (abcdef)
v_{38}	$Gluc_{ext}$ (abcdef)	→	G6P (abcdef)
v_{39}	Cit_{ext} (abcdef)	→	Cit (abcdef)
v_{40}	Glyc (abc)	↔	$Glyc_{ext}$ (abc)
v_{41}	PDO (abc)	→	PDO_{ext} (abc)
v_{42}	Ac (ab)	→	Ac_{ext} (ab)
v_{43}	CO_2 (a)	→	$CO_{2,ext}$ (a)
Amino Acid Biosynthesis			
v_{44}	AKG (abcde)	→	Glu (abcde)
v_{45}	Glu (abcde)	→	Gln (abcde)
v_{46}	Glu (abcde)	→	Pro (abcde)
v_{47}	Glu (abcde) + CO_2 (f) + Gln (ghijk) + Asp (lmno) + AcCoA (pq)	→	Arg (abcdef) + AKG (ghijk) + Fum (lmno) + Ac (pq)
v_{48}	OAA (abcd) + Glu (efghi)	→	Asp (abcd) + AKG (efghi)
v_{49}	Asp (abcd)	→	Asn (abcd)
v_{50}	Pyr (abc) + Glu (defgh)	→	Ala (abc) + AKG (defgh)
v_{51}	3PG (abc) + Glu (defgh)	→	Ser (abc) + AKG (defgh)
v_{52}	Ser (abc)	↔	Gly (ab) + MEETHF (c)
v_{53}	Gly (ab)	↔	CO_2 (a) + MEETHF (b)
v_{54}	Thr (abcd)	→	Gly (ab) + AcCoA (cd)
v_{55}	Ser (abc) + AcCoA (de)	→	Cys (abc) + Ac (de)
v_{56}	Asp (abcd) + Pyr (efg) + Glu (hijkl) + SucCoA (mnop)	→	LL-DAP (abcdgfe) + AKG (hijkl) + Suc (mnop)
v_{57}	LL-DAP (abcdefg)	→	Lys (abcdef) + CO_2 (g)
v_{58}	Asp (abcd)	→	Thr (abcd)
v_{59}	Asp (abcd) + METHF (e) + Cys (fgh) + SucCoA (ijkl)	→	Met (abcde) + Pyr (fgh) + Suc (ijkl)
v_{60}	Pyr (abc) + Pyr (def) + Glu (ghijk)	→	Val (abcef) + CO_2 (d) + AKG (ghijk)
v_{61}	AcCoA (ab) + Pyr (cde) + Pyr (fgh) + Glu (ijklm)	→	Leu (abdghe) + CO_2 (c) + CO_2 (f) + AKG (ijklm)
v_{62}	Thr (abcd) + Pyr (efg) + Glu (hijkl)	→	Ile (abfcdg) + CO_2 (e) + AKG (hijkl)
v_{63}	PEP (abc) + PEP (def) + E4P (ghij) + Glu (klmno)	→	Phe (abcefghij) + CO_2 (d) + AKG (klmno)
v_{64}	PEP (abc) + PEP (def) + E4P (ghij) + Glu (klmno)	→	Tyr (abcefghij) + CO_2 (d) + AKG (klmno)
v_{65}	Ser (abc) + R5P (defgh) + PEP (ijk) + E4P (lmno) + PEP (pqr) + Gln (stuvw)	→	Trp (abcedklmnoj) + CO_2 (i) + GAP (fgh) + Pyr (pqr) + Glu (stuvw)
v_{66}	R5P (abcde) + FTHF (f) + Gln (ghijk) + Asp (lmno)	→	His (edcbaf) + AKG (ghijk) + Fum (lmno)
Biomass Formation			
v_{67}	0.488 Ala + 0.281 Arg + 0.229 Asn + 0.229 Asp + 0.087 Cys + 0.250 Glu + 0.250 Gln + 0.582 Gly + 0.090 His + 0.276 Ile + 0.428 Leu + 0.326 Lys + 0.146 Met + 0.176 Phe + 0.210 Pro + 0.205 Ser + 0.241 Thr + 0.054 Trp + 0.131 Tyr + 0.402 Val + 0.205 G6P + 0.071 F6P + 0.754 R5P + 0.129 GAP + 0.619 3PG + 0.051 PEP + 0.083 Pyr + 2.510 AcCoA + 0.087 AKG + 0.340 OAA + 0.443 MEETHF	→	39.68 Biomass

Table 3.2: A list of reactions and atom transitions within the *E. coli* network for extracellular transport, amino acid biosynthesis, and biomass formation. Carbon atom transitions are indicated within parentheses. Irreversible and reversible reactions are indicated by the symbols → and ↔, respectively.

Model	Cumomer	EMU	EMU
Decoupling method	Size	Size/connectivity	Blocks
(Size) # of vars	(1) 54	(1) 146	(1) 117,24,1×5
	(2) 241	(2) 90	(2) 34,22,9,5×2,4,2,1×9
	(3) 527	(3) 47	(3) 26,5×2,4,1×7
	(4) 771	(4) 12,8,1×2	(4) 8,6,5,1×3
	(5) 876	(5) 5,1×4	(5) 5,1×4
	(6) 832	(6) 2	(6) 2
	(7) 655	(7) none	(7) none
	(8) 404	(8) 1	(8) 1
	(9) 183	(9) 1	(9) 1
	(10) 57		
	(11) 11		
	(12) 1		
Total variables	4612	318	318
Simulation time	Not available	22 seconds	16 seconds

Table 3.3: A comparison of modeling approaches to simulate the dynamic labeling of 33 GC/MS fragments in the large *E. coli* metabolic network. Subproblems are specified as in Table 2.1. Multiple occurrences of a particular size of subproblem are indicated with the multiplication symbol; that is, "5×2" indicates that two subproblems of size 5 exist within the system. EMU network decomposition dramatically reduces the number of state variables within the overall system. Block decoupling further simplifies the system by minimizing the number of state variables within any one subproblem, reducing computational time by an additional 27%.

3. **Short nonstationary experiment:** Two sets of replicate measurements were taken every second for 15 seconds following the introduction of tracer to the culture (for a total of 30 sets of measurements). All measured metabolite fragments remained isotopically transient during this regime.

4. **Long nonstationary experiment with concentrations:** The labeling measurements of the long nonstationary experiment were combined with concentration measurements. Concentrations were assumed to be available for all metabolites whose labeling was measured.

5. **Short nonstationary experiment with concentrations:** The labeling measurements of the short nonstationary experiment were combined with concentration measurements. Concentrations were assumed to be available for all metabolites whose labeling was measured.

The timelines of these five experiments and their corresponding labeling measurements are illustrated in Figure 3-1. Tables 3.4 and 3.5 list the available measurements, which include 33 mass spectroscopy fragments and seven external fluxes. Standard errors of 5% for external fluxes, 0.3-1 mol% for GC/MS MIDs, and 10% for concentrations were assumed and introduced randomly and normally.

3.3 Flux and Concentration Estimates

Fluxes and concentrations were estimated for each experiment. One forward nonstationary simulation of metabolite labeling required 16 seconds of computational time. Parameter estimation, beginning with a randomly distributed guess of concentrations and fluxes, ran in under 15 minutes. Nonlinear confidence intervals were also calculated for each estimated parameter. The results have been grouped as net fluxes, exchange fluxes, and metabolite concentrations and displayed in Figures 3-2, 3-3 and 3-4. Results are also listed in Tables 3.6, 3.7 and 3.8. The wide majority of parameter values were recovered within their respective 95% confidence intervals; those values

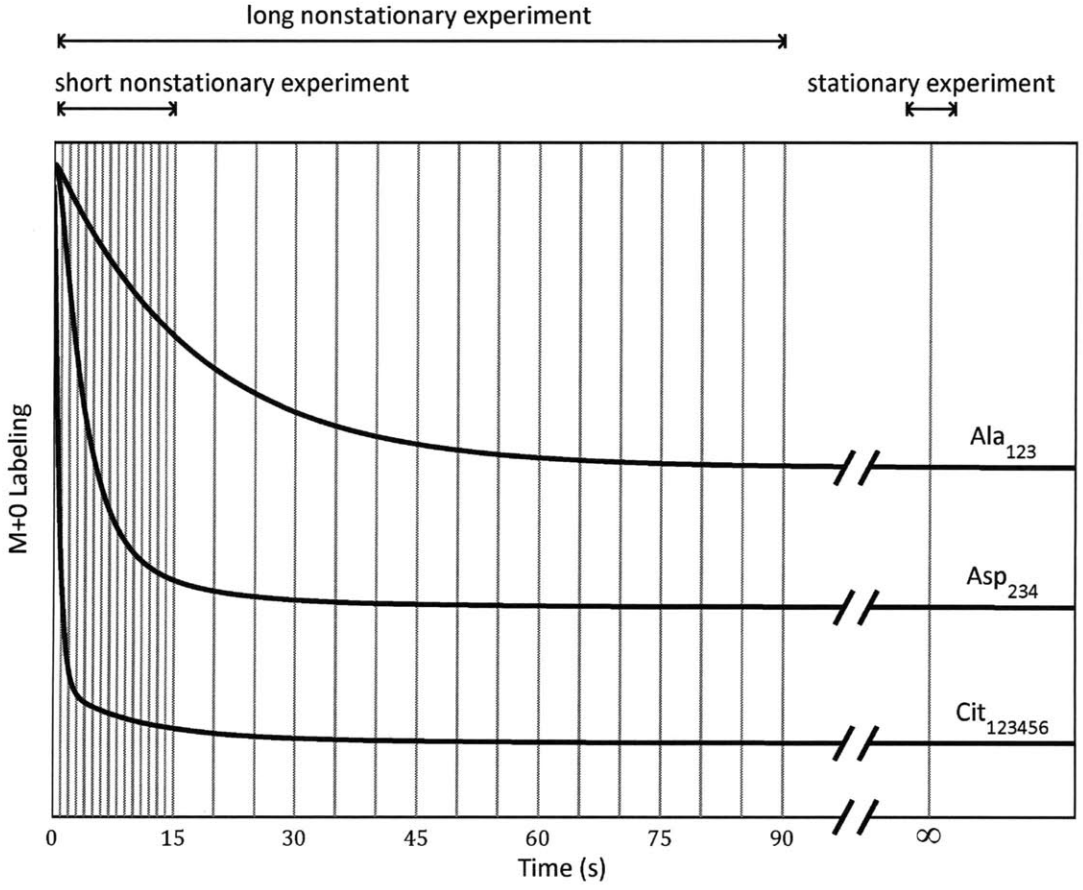

Figure 3-1: Measurement time points for the simulated experiments involving the large *E. coli* model. Time points are indicated by gray vertical lines. M+0 labeling profiles for three representative measured metabolite fragments are overlaid to convey the time scale of the isotopic transience of the system.

Metabolite	Flux
Gluc	$Gluc_{ext} \rightarrow G6P$
Cit	$Cit \rightarrow Cit_{ext}$
CO_2	$CO_2 \rightarrow CO_{2,ext}$
Glyc	$Glyc_{ext} \rightarrow Glyc$
Ac	$Ac \rightarrow Ac_{ext}$
PDO	$PDO \rightarrow PDO_{ext}$
Biomass	Various metabolites $\rightarrow$ Biomass

Table 3.4: External fluxes measured in the simulated study of the large *E. coli* network. We assumed that all fluxes measured in the previous *E. coli* study [12] would be available.

Metabolite	Mass	Carbons	Formula
AKG	346	12345	$C_{14}H_{28}O_5NSi_2$
Ala	232	23	$C_{10}H_{26}ONSi_2$
Ala	260	123	$C_{11}H_{26}O_2NSi_2$
Asp	302	12	$C_{14}H_{32}O_2NSi_2$
Asp	376	12	$C_{16}H_{38}O_3NSi_3$
Asp	390	234	$C_{17}H_{40}O_3NSi_3$
Asp	418	1234	$C_{18}H_{40}O_4NSi_3$
Cit	459	123456	$C_{20}H_{39}O_6Si_3$
Glu	330	2345	$C_{16}H_{36}O_2NSi_2$
Glu	432	12345	$C_{19}H_{42}O_4NSi_3$
Gly	218	2	$C_9H_{24}ONSi_2$
Gly	246	12	$C_{10}H_{24}O_2NSi_2$
Ile	200	23456	$C_{11}H_{26}NSi$
Ile	274	23456	$C_{13}H_{32}ONSi_2$
Leu	274	23456	$C_{13}H_{32}ONSi_2$
Mal	419	1234	$C_{18}H_{39}O_5Si_3$
Met	218	2345	$C_{10}H_{24}NSiS$
Met	292	2345	$C_{12}H_{30}NOSi_2S$
Met	320	12345	$C_{13}H_{30}NO_2Si_2S$
Phe	234	23456789	$C_{14}H_{24}NSi$
Phe	302	12	$C_{14}H_{32}O_2NSi_2$
Phe	308	23456789	$C_{16}H_{30}ONSi_2$
Phe	336	123456789	$C_{17}H_{30}O_2NSi_2$
Pyr	174	123	$C_6H_{12}O_3NSi$
Ser	288	23	$C_{14}H_{34}NOSi_2$
Ser	302	12	$C_{14}H_{32}O_2NSi_2$
Ser	362	23	$C_{16}H_{40}O_2NSi_3$
Ser	390	123	$C_{17}H_{40}O_3NSi_3$
Suc	289	1234	$C_{12}H_{25}O_4Si_2$
Thr	376	234	$C_{17}H_{42}O_2NSi_3$
Thr	404	1234	$C_{18}H_{42}O_3NSi_3$
Tyr	302	12	$C_{14}H_{32}O_2NSi_2$
Val	260	2345	$C_{12}H_{30}ONSi_2$
Val	288	12345	$C_{13}H_{30}O_2NSi_2$

Table 3.5: Metabolite MIDs measured by GC/MS in the simulated study of the large *E. coli* network. We assumed that all proteinogenic amino acid MIDs measured in the previous *E. coli* study [12] would be available in free intracellular form, as well as several organic acid fragments we observed experimentally in typical metabolite extracts. MTBSTFA is the derivatizing agent.

that were not recovered still fell reasonably near their estimated intervals. Exchange fluxes were scaled according to:

$$v_{\text{xch}}^{[0,1]} = \frac{v_{\text{xch}}}{v_{\text{xch}} + v_{\text{ref}}} \tag{3.1}$$

where v_{xch} is the unscaled exchange flux, v_{ref} is the reference flux (in this case, glucose uptake), and $v_{\text{xch}}^{[0,1]}$ is the scaled exchange flux [156].

Net fluxes were estimated most accurately and precisely. Most estimated net flux values fell within 10% of the actual values and possessed confidence intervals ranging between ±5 and ±20%. While the stationary measurements generated estimates significantly closer to actual values, confidence intervals across all five experiments were comparable in width.

Exchange flux estimation was considerably more difficult. Out of the 24 total exchange fluxes, 10 were unidentifiable (or nearly unidentifiable). Even when confidence intervals were obtained, they tended to be extremely broad. For the majority of these intervals, only an upper or a lower bound could be found. Overall, none of the five experiments could clearly claim significantly more precise confidence intervals. On a parameter-by-parameter basis, however, precision varied greatly between the different experiments.

Both upper and lower bounds were successfully found for 14 metabolite concentrations. Upper bounds were identified for the remaining concentrations. Confidence intervals ranged between ±5 and ±25% except for pyruvate (±75%). The long non-stationary experiment consistenly produced narrower confidence intervals than the short experiment. (Obviously, no metabolite concentrations could be obtained in the stationary experiment.)

To simplify comparisons between the different experimental designs, the confidence intervals of different parameters were averaged to create lumped precision scores for each parameter type (net fluxes, exchange fluxes, and concentrations) within each experiment (see Figure 3-5). The details behind precision scoring are explained in further detail in Chapter 6. In short, larger scores correlate with greater precision, where

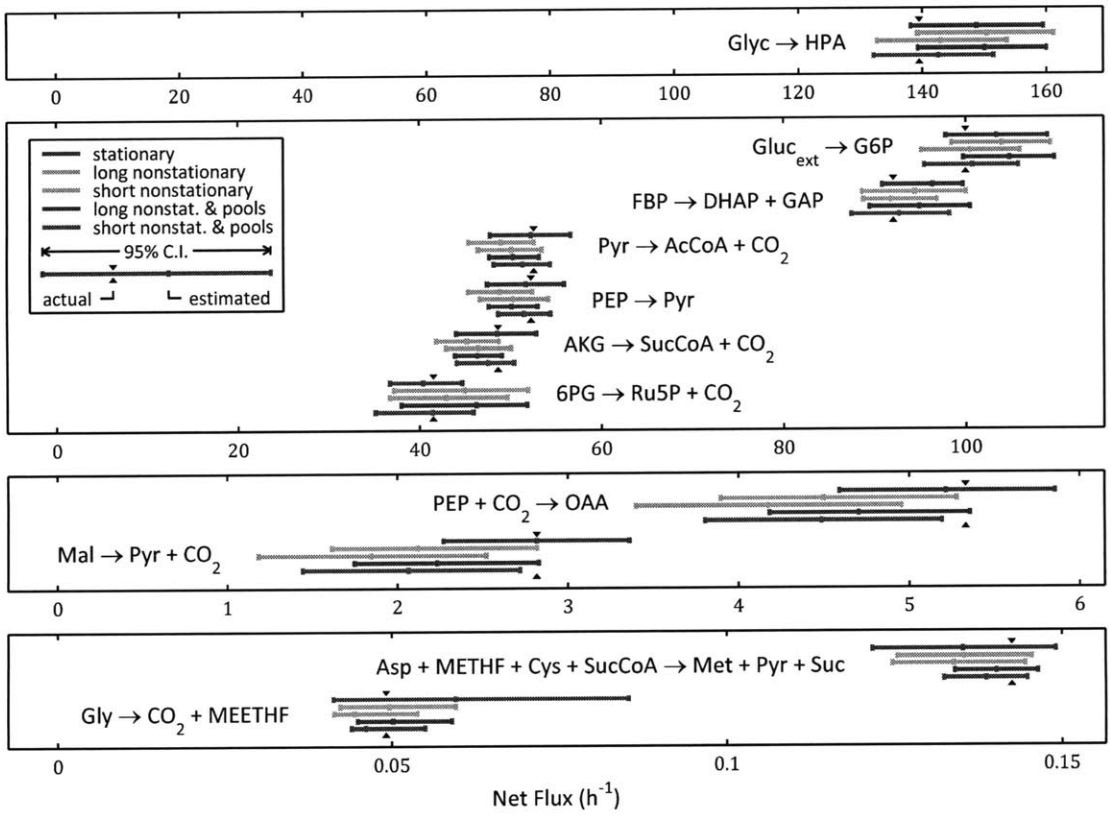

Figure 3-2: A comparison of estimated independent net fluxes in the large *E. coli* network using stationary measurements, long nonstationary measurements (with and without concentrations), and short nonstationary measurements (with and without concentrations). Actual flux values are indicated by black arrowheads. Accurate 95% confidence intervals are indicated by horizontal error bars while hash marks within these bars indicate estimated flux values. Values are in normalized units of h^{-1} and are scaled such that the $Gluc_{ext} \rightarrow G6P$ flux is 100 h^{-1}.

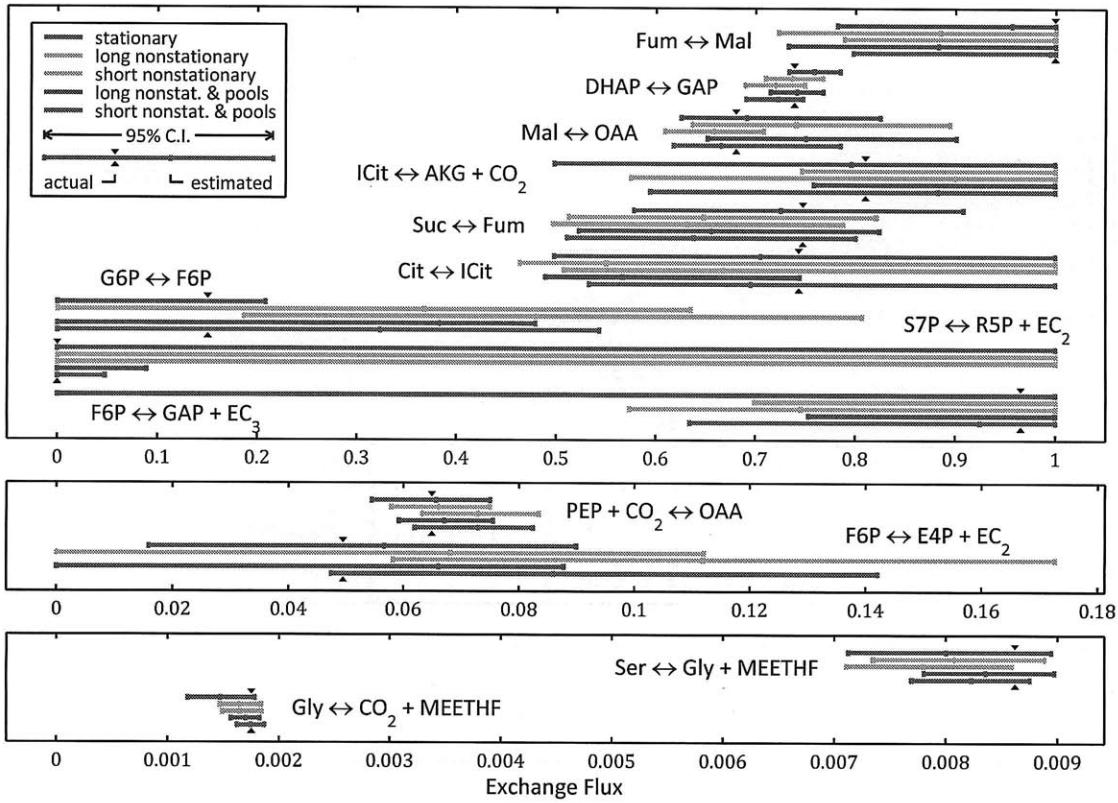

Figure 3-3: A comparison of estimated exchange fluxes in the large *E. coli* network using stationary measurements, long nonstationary measurements (with and without concentrations), and short nonstationary measurements (with and without concentrations). Actual flux values are indicated by black arrowheads. Accurate 95% confidence intervals are indicated by horizontal error bars while hash marks within these bars indicated estimated flux values. Values are dimensionless and are scaled according to Equation 3.1. Unidentifiable and nearly unidentifiable fluxes were omitted.

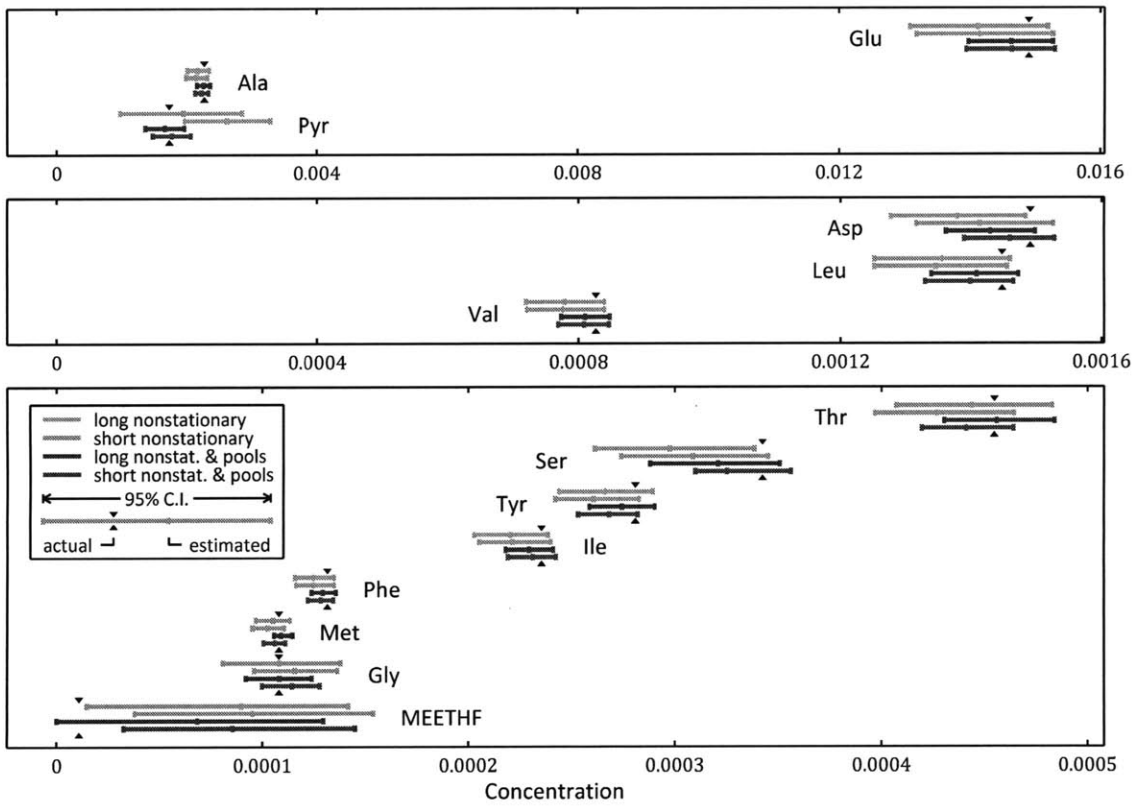

Figure 3-4: A comparison of estimated metabolite concentrations in the large *E. coli* network using long nonstationary measurements and short nonstationary measurements, both with and without concentration measurements. Actual concentration values are indicated by black arrowheads. Accurate 95% confidence intervals are indicated by horizontal error bars while hash marks within these bars indicate estimated concentration values. Values are dimensionless and are scaled such that the $\text{Gluc}_{\text{ext}} \rightarrow$ G6P flux is 100 h^{-1}. Concentrations with lower bounds of zero were omitted.

Net Flux	O	S	LNS	SNS	LNS w/ C	SNS w/ C
FBP → DHAP...	92.1	90.8 96.4 99.7	88.6 94.4 100	88.7 91.8 96.8	89.4 94.9 100	87.5 92.7 98.2
PEP → Pyr	52.3	47.4 51.7 55.9	45.3 48.8 52.4	46.6 50.3 54.1	47.6 50.1 53.0	48.6 51.5 54.4
6PG → Ru5P...	41.5	36.7 40.4 44.7	37.2 45.0 52.0	36.7 42.9 49.6	38.0 46.2 51.8	35.1 41.5 45.9
Pyr → AcCoA...	52.6	47.8 52.3 56.6	45.4 48.9 52.6	46.4 50.2 53.5	47.7 50.3 53.2	48.2 51.3 54.4
AKG → SucCoA...	48.6	44.0 48.5 52.8	41.8 45.1 48.7	42.9 46.4 50.0	43.8 46.3 49.1	44.1 47.5 50.4
Mal → Pyr...	2.8	2.3 2.8 3.4	1.6 2.1 2.8	1.2 1.8 2.5	1.7 2.2 2.8	1.4 2.1 2.7
PEP... → OAA	5.3	4.6 5.2 5.9	3.9 4.5 5.3	3.4 4.2 5.0	4.2 4.7 5.4	3.8 4.5 5.2
Glyc → HPA	140	138 149 160	139 150 161	133 143 154	139 150 160	132 143 151
Gly → CO2...	0.0	0.0 0.1 0.1	0.0 0.0 0.1	0.0 0.0 0.1	0.0 0.1 0.1	0.0 0.0 0.1
Asp... → Met...	0.1	0.1 0.1 0.1	0.1 0.1 0.1	0.1 0.1 0.1	0.1 0.1 0.1	0.1 0.1 0.1
$Gluc_{ext}$ → G6P	100	97.8 103 109	98.5 104 109	95.1 100 106	99.7 105 110	95.5 101 106

Table 3.6: A comparison of original (O) and estimated independent net fluxes in the large *E. coli* network using stationary measurements (S), long nonstationary measurements (LNS), short nonstationary measurements (SNS), long nonstationary measurements with concentrations (LNS w/ C), and short nonstationary measurements with concentrations (SNS w/ C). The estimated value is underlined and placed in between the lower and upper bounds of the 95% confidence interval for each parameter of each experiment. Values are in units of h^{-1} and are scaled such that the $Gluc_{ext}$ → G6P flux is 100 h^{-1}.

Exchange Flux	O	S	LNS	SNS	LNS w/ C	SNS w/ C
G6P ↔ F6P	18.6	0 0.0 27.5	0 61.0 182	24.0 96.6 438	0 65.1 96.5	0 50.1 125
FBP ↔ DHAP...	222	0 218 ∞	0 256 ∞	0 357 ∞	0 509 ∞	0 778 ∞
DHAP ↔ GAP	297	289 330 383	257 294 345	233 269 313	263 301 346	234 274 310
GAP ↔ 3PG	2E4	552 1E7 ∞	200 634 ∞	534 1E6 ∞	244 751 ∞	657 8E4 ∞
3PG ↔ PEP	1E3	497 8E3 ∞	254 2E3 ∞	287 2E3 ∞	286 2E3 ∞	253 629 4E3
Ru5P ↔ X5P	91.8	6.7 112 ∞	0 5.8 ∞	0 0.0 ∞	0 0.0 ∞	0 69.1 ∞
Ru5P ↔ R5P	3E4	0 3E4 ∞	0 3E4 ∞	0 3E4 ∞	0 4E3 ∞	0 8E3 ∞
X5P ↔ GAP...	31.4	6.8 39.4 ∞	0 0.0 ∞	0 0.0 ∞	0 14.2 ∞	0 0.0 ∞
F6P ↔ E4P...	5.5	1.7 6.3 10.4	0 7.7 13.2	6.5 13.2 21.9	0 7.4 10.1	5.2 9.9 17.4
S7P ↔ R5P...	0.0	0 0.0 ∞	0 0.0 ∞	0 0.0 ∞	0 0.0 10.3	0 0.0 5.2
F6P ↔ GAP...	3E3	0 2E5 ∞	244 2E5 ∞	141 307 ∞	321 1E5 ∞	182 1E3 ∞
S7P ↔ E4P...	2E4	0 2E4 ∞	0 2E4 ∞	0 2E4 ∞	0 4E4 ∞	0 5E4 ∞
Cit ↔ ICit	304	104 250 ∞	90.7 128 ∞	108 210 ∞	100 137 306	120 239 ∞
ICit ↔ AKG...	447	104 409 ∞	309 3E5 ∞	142 944 ∞	329 3E5 ∞	153 791 ∞
SucCoA ↔ Suc	1.7	0 1.8 ∞	0 29.7 ∞	0 69.7 ∞	0 23.2 ∞	0 11.0 ∞
Suc ↔ Fum	310	144 277 1E3	110 193 481	103 180 390	115 200 491	109 185 421
Fum ↔ Mal	2E5	376 2E3 ∞	274 809 ∞	392 1E5 ∞	288 791 ∞	414 2E4 ∞
Mal ↔ OAA	223	175 234 495	183 299 888	163 202 254	196 315 953	169 208 383
PEP... ↔ OAA	7.3	6.0 7.4 8.5	6.4 7.4 8.5	7.1 8.2 9.5	6.6 7.5 8.6	6.9 8.2 9.4
AcCoA ↔ Ac	1E3	0 1E3 ∞	0 1E3 ∞	0 1E3 ∞	0 310 ∞	0 2E3 ∞
DHAP ↔ Glyc3P	112	0 85.4 ∞	0 21.5 ∞	0 0.0 ∞	0 118 ∞	0 48.3
Ser ↔ Gly...	0.9	0.8 0.8 0.9	0.8 0.9 0.9	0.8 0.8 0.9	0.8 0.9 0.9	0.8 0.9 0.9
Gly ↔ CO2...	0.2	0.1 0.2 0.2	0.2 0.2 0.2	0.2 0.2 0.2	0.2 0.2 0.2	0.2 0.2 0.2
$Glyc_{ext}$... ↔ Glyc...	97.2	0 91.4 ∞	0 94.3 ∞	0 80.0 ∞	0 203 ∞	0 27.6 ∞

Table 3.7: A comparison of original (O) and estimated independent exchange fluxes in the large *E. coli* network using stationary measurements (S), long nonstationary measurements (LNS), short nonstationary measurements (SNS), long nonstationary measurements with concentrations (LNS w/ C), and short nonstationary measurements with concentrations (SNS w/ C). The estimated value is underlined and placed in between the lower and upper bounds of the 95% confidence interval for each parameter of each experiment. Values are in units of h^{-1} and are scaled such that the $Gluc_{ext} \rightarrow$ G6P flux is 100 h^{-1}.

Pool	O	LNS	SNS	LNS w/ C	SNS w/ C
AKG	9.1E-5	0 6.7E-6 5.5E-3	0 8.4E-6 5.0E-3	9.3E-5 1.1E-4 1.3E-4	9.3E-5 1.1E-4 1.3E-4
Ac	1.1E-5	0 9.7E-6 2.6E-3	0 1.3E-5 1.5E-3	0 1.2E-6 2.0E-3	0 4.4E-4 1.5E-3
AcCoA	1.9E-4	0 1.2E-3 2.6E-3	0 1.8E-4 1.5E-3	0 8.8E-4 2.1E-3	0 1.1E-4
Ala	2.3E-3	2.0E-3 2.2E-3 2.3E-3	2.0E-3 2.1E-3 2.3E-3	2.2E-3 2.3E-3 ∞ 2.4E-3	2.1E-3 2.2E-3 2.3E-3
Arg	1.9E-3	0 1.9E-3 ∞	0 1.9E-3 ∞	0 1.9E-3 ∞	0 1.9E-3 ∞
Asn	1.2E-3	0 1.2E-3 ∞	0 1.2E-3 ∞	0 1.2E-3 ∞	0 1.2E-3 ∞
Asp	1.5E-3	1.3E-3 1.4E-3 1.5E-3	1.3E-3 1.4E-3 1.5E-3	1.4E-3 1.4E-3 1.5E-3	1.4E-3 1.5E-3 1.5E-3
CO2	1.1E-5	0 7.8E-3 2.3E-2	0 2.1E-5 1.1E-2	0 3.4E-3 1.1E-2	0 4.6E-6 1.1E-2
Cit	2.9E-3	0 3.4E-4 2.5E-3	0 3.0E-3 5.4E-3	2.5E-3 3.0E-3 3.6E-3	2.6E-3 3.2E-3 3.7E-3
Cys	1.1E-4	0 2.7E-4 2.5E-2	0 1.1E-6 ∞	0 1.8E-4 8.7E-3	0 5.1E-6 ∞
DHAP	1.1E-4	0 1.2E-5	0 8.3E-6 8.2E-3	0 4.1E-5 1.2E-2	0 6.2E-5 8.1E-3
E4P	6.3E-5	0 1.9E-6 1.1E-3	0 2.1E-4 9.5E-4	0 1.7E-6 9.8E-4	0 2.2E-4 8.1E-4
EC2	1.1E-4	0 5.4E-4 3.9E-3	0 4.8E-6 2.9E-3	0 8.7E-4 3.7E-3	0 1.4E-5 3.4E-3
EC3	1.1E-4	0 2.5E-5 4.6E-3	0 3.3E-6 3.0E-3	0 3.5E-4 5.0E-3	0 8.7E-5 3.3E-3
F6P	3.9E-4	0 1.1E-5 4.4E-3	0 4.8E-6 3.9E-3	0 5.7E-5 4.4E-3	0 3.3E-5 2.8E-3
FBP	1.8E-4	0 5.8E-6 5.7E-3	0 5.6E-6 3.8E-3	0 1.4E-5 5.3E-3	0 2.2E-5 4.0E-3
FTHF	1.1E-5	0 1.1E-5 ∞	0 1.1E-5 ∞	0 1.1E-5 ∞	0 1.1E-5 ∞
Fum	5.8E-5	0 2.0E-6 1.1E-3	0 2.1E-6 6.5E-4	0 3.9E-6 1.0E-3	0 3.6E-6 3.9E-4
G6P	2.2E-3	0 4.9E-3 7.1E-3	0 2.8E-3 5.3E-3	0 3.9E-3 6.8E-3	0 1.2E-3 4.7E-3
GAP	1.4E-4	0 1.2E-5 9.0E-3	0 2.5E-5 6.5E-3	0 5.7E-5 8.2E-3	0 3.3E-5 6.6E-3
Gln	5.9E-3	1.3E-3 3.8E-3 8.0E-3	4.4E-4 3.6E-3 ∞	1.5E-3 4.2E-3 8.5E-3	2.2E-4 3.8E-3 ∞
Glu	1.5E-2	1.3E-2 1.4E-2 1.5E-2	1.3E-2 1.4E-2 1.5E-2	1.4E-2 1.5E-2 1.5E-2	1.4E-2 1.5E-2 1.5E-2
Gly	1.1E-4	8.1E-5 1.1E-4 1.4E-4	9.6E-5 1.2E-4 1.4E-4	9.2E-5 1.1E-4 1.2E-4	10.0E-5 1.1E-4 1.3E-4
Glyc	1.1E-4	0 1.1E-4 ∞	0 1.1E-4 ∞	0 1.1E-4 ∞	0 1.1E-4 ∞
Glyc3P	1.1E-5	0 6.9E-6 ∞	0 8.4E-6 ∞	0 1.9E-6 ∞	0 3.1E-6 ∞
HPA	1.1E-5	0 1.1E-5 ∞	0 1.1E-5 ∞	0 1.1E-5 ∞	0 1.1E-5 ∞
His	1.5E-3	0 1.5E-3 ∞	0 1.5E-3 ∞	0 1.5E-3 ∞	0 1.5E-3 ∞
ICit	2.9E-3	0 4.3E-3 5.4E-3	2.6E-7 2.6E-3 6.0E-3	8.7E-4 1.8E-3 2.8E-3	2.4E-3 3.3E-3 ∞
Ile	2.4E-4	2.0E-4 2.2E-4 2.4E-4	2.1E-4 2.2E-4 2.4E-4	2.2E-4 2.3E-4 2.4E-4	2.2E-4 2.3E-4 2.4E-4
KDPG	1.1E-5	0 1.7E-6 4.4E-4	0 2.0E-5 2.0E-4	0 3.8E-5 4.0E-4	0 3.7E-5 2.2E-4
LL-DAP	1.1E-5	0 4.0E-5 ∞	0 1.3E-5 ∞	0 1.1E-5 ∞	0 1.9E-5 ∞
Leu	1.4E-3	1.3E-3 1.4E-3 1.5E-3	1.3E-3 1.3E-3 1.5E-3	1.3E-3 1.4E-3 1.5E-3	1.3E-3 1.4E-3 1.5E-3
Lys	9.5E-4	0 9.5E-4 ∞	0 9.5E-4 ∞	0 9.5E-4 ∞	0 9.5E-4 ∞
MEETHF	1.1E-5	1.5E-5 9.0E-5 1.4E-4	3.8E-5 9.5E-5 1.5E-4	0 6.8E-5 1.3E-4	3.2E-5 8.5E-5 1.5E-4
METHF	1.1E-5	0 8.7E-7 8.5E-6	0 2.6E-6 7.9E-6	0 2.5E-6 8.8E-6	0 3.6E-6 9.8E-6
Mal	3.0E-4	0 1.1E-6 1.2E-3	0 2.2E-6 7.3E-4	2.0E-4 2.6E-4 3.2E-4	2.0E-4 2.6E-4 3.2E-4
Met	1.1E-4	9.7E-5 1.1E-4 1.1E-4	9.5E-5 1.0E-4 1.1E-4	1.1E-4 1.1E-4 1.1E-4	1.0E-4 1.1E-4 1.1E-4
OAA	1.1E-5	0 7.3E-4 1.4E-3	0 3.0E-6 7.3E-4	0 5.5E-4 1.2E-3	0 3.7E-6 4.3E-4
PDO	1.1E-3	0 1.1E-3 ∞	0 1.1E-3 ∞	0 1.1E-3 ∞	0 1.1E-3 ∞
PEP	1.7E-3	0 1.1E-5 8.5E-3	0 1.9E-3 6.4E-3	0 6.1E-4 7.4E-3	0 5.1E-3 ∞
3PG	1.4E-3	0 1.6E-5 9.7E-3	0 2.1E-5 6.5E-3	0 1.2E-4 7.6E-3	0 4.2E-5 6.7E-3
6PG	5.2E-6	0 8.7E-4 2.9E-3	0 1.8E-4 2.0E-3	0 1.2E-3 2.4E-3	0 2.5E-6 2.0E-3
Phe	1.3E-4	1.2E-4 1.3E-4 1.3E-4	1.2E-4 1.2E-4 1.3E-4	1.2E-4 1.3E-4 1.4E-4	1.2E-4 1.3E-4 1.3E-4
Pro	3.3E-4	0 3.3E-4 ∞	0 3.3E-4 ∞	0 3.3E-4 ∞	0 3.3E-4 ∞
Pyr	1.7E-3	9.7E-4 1.9E-3 2.9E-3	2.0E-3 2.6E-3 3.3E-3	1.4E-3 1.7E-3 2.0E-3	1.5E-3 1.8E-3 2.1E-3
R5P	2.6E-4	0 1.2E-6 2.8E-3	0 7.7E-5 1.5E-3	0 1.1E-6 2.6E-3	0 2.5E-4 1.8E-3
Ru5P	7.2E-5	0 8.8E-6 2.4E-3	0 4.0E-5 1.7E-3	0 1.0E-6 2.1E-3	0 1.7E-5 1.7E-3
S7P	1.8E-4	0 1.9E-6 9.1E-4	0 2.9E-6 9.6E-4	0 1.7E-6 8.9E-4	0 5.8E-5 1.2E-3
Ser	3.4E-4	2.6E-4 3.0E-4 3.4E-4	2.7E-4 3.1E-4 3.5E-4	2.9E-4 3.2E-4 3.5E-4	3.1E-4 3.3E-4 3.6E-4
Suc	5.9E-5	0 1.2E-6 8.0E-4	0 2.5E-4 7.1E-4	5.5E-5 6.6E-5 7.8E-5	5.5E-5 6.6E-5 7.8E-5
SucCoA	1.1E-5	0 1.2E-6 8.0E-4	0 2.5E-6 6.8E-4	0 1.1E-6 7.5E-4	0 3.8E-6 3.8E-4
Thr	4.5E-4	4.1E-4 4.4E-4 4.8E-4	4.0E-4 4.3E-4 4.6E-4	4.3E-4 4.6E-4 4.8E-4	4.2E-4 4.4E-4 4.6E-4
Trp	1.1E-4	0 1.1E-4 ∞	0 1.1E-4 ∞	0 1.1E-4 ∞	0 1.1E-4 ∞
Tyr	2.8E-4	2.4E-4 2.7E-4 2.9E-4	2.4E-4 2.6E-4 2.8E-4	2.6E-4 2.7E-4 2.9E-4	2.5E-4 2.7E-4 2.8E-4
Val	8.3E-4	7.2E-4 7.8E-4 8.4E-4	7.2E-4 7.8E-4 8.4E-4	7.7E-4 8.1E-4 8.5E-4	7.7E-4 8.1E-4 8.5E-4
X5P	8.9E-5	0 3.3E-5 3.1E-3	0 1.9E-4 2.7E-3	0 1.9E-5 3.4E-3	0 5.4E-5 1.9E-3

Table 3.8: A comparison of original (O) and estimated metabolite concentrations using the same measurement sets listed in Tables 3.6 and 3.7, excepting the stationary set. The estimated value is underlined and placed in between the lower and upper bounds of the 95% confidence interval for each parameter of each experiment. Values are dimensionless and are scaled such that the $Gluc_{ext} \rightarrow G6P$ flux is 100 h^{-1}.

a score of one indicates absolute precision (or confidence intervals of zero) over all parameters, and a score of zero indicates absolute unidentifiability for all parameters under consideration.

To simplify comparisons even further, we studied confidence intervals for the three major pathways of central carbon metabolism (glycolysis, the pentose phosphate pathway, and the TCA cycle) resulting from the five different experimental designs. Experiments were qualitatively ranked according to their effectiveness in flux estimation for each pathway. These rankings are shown in Table 3.9.

3.4 Discussion

The most striking finding arising from our experiment-by-experiment comparisons is the significant advantage created by the addition of concentrations to the measurement set. We see that concentration measurements increase precision across each parameter class (Figure 3-5) as well as each central pathway (Table 3.9). Concentration data has a dramatic effect on estimation because they reduce the increased indetermination introduced by the shift from MFA to NMFA. The parameter space of NMFA consists of fluxes and concentrations and is therefore usually larger than the typical flux-only MFA parameter space by orders of magnitude. By providing concentration measurements, we can effectively reduce this space while still maintaining the advantages of nonstationary analysis (e.g., increased sensitivity of labeling measurements at early time points).

Even when concentrations are not available, we still find that metabolic parameters can be successfully estimated in the large *E. coli* model. We were able to estimate all net fluxes, most exchange fluxes, and some metabolite concentrations with only external flux measurements and nonstationary labeling measurements. In these experiments, we see that measurements clustered earlier in isotopic transience tend to deliver more precision than those in late transience. Since this advantage disappears upon including concentration data, it follows that these early time points help resolve the competing effects of concentration and flux upon labeling. However,

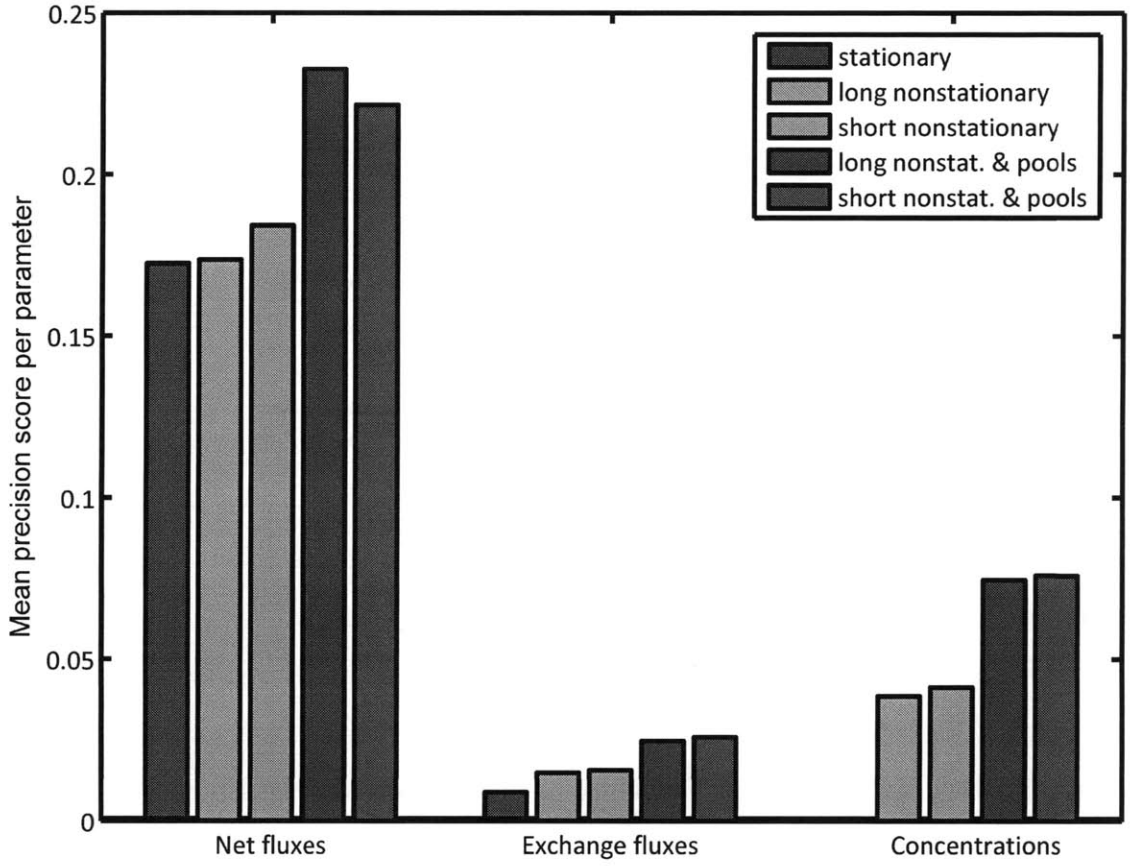

Figure 3-5: Mean precision scores for net fluxes, exchange fluxes, and metabolite concentrations for the five different *E. coli* experimental designs. A score of one indicates absolute precision over all considered parameters while a score of zero correlates to absolute unidentifiability. In general, parameters exhibit a slight improvement in precision when relying on nonstationary labeling measurements as compared to only stationary measurements and a more substantial improvement when concentrations are also included as measurements. The equations behind precision scoring can be found in Chapter 6. For this calculation, parameters α and β were set to 1 and 0.1, respectively.

	Glycolysis	PP Pathway	TCA Cycle
MFA	☆ ☆ ☆	☆ ☆ ☆ ☆ ☆	☆
Long NMFA	☆	☆	☆ ☆
Short NMFA	☆ ☆ ☆	☆ ☆ ☆	☆ ☆ ☆
Long NMFA w/ pools	☆ ☆ ☆ ☆ ☆	☆ ☆ ☆	☆ ☆ ☆ ☆ ☆
Short NMFA w/ pools	☆ ☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆

Table 3.9: A simple, qualitative comparison of flux results using stationary measurements, long nonstationary measurements, short nonstationary measurements, long nonstationary measurements with concentrations, and short nonstationary measurements with concentrations. A ranking of five stars indicates the best performance, and one star the worst.

the simulated data do indicate that there are some limitations when relying upon measurements at extremely short time scales. Confidence intervals produced from the short nonstationary experiments were the least likely to recapture the original parameter values. Across the five experiments, only six of the original flux values fell outside of their respective 95% confidence intervals and all of these intervals were generated from the two short nonstationary experiments. Intervals from the stationary and long nonstationary experiments successfully recovered all actual flux values.

Overall, our analysis of the *E. coli* network demonstrates that the potential improvements of nonstationary flux analysis (without concentration measurements) over stationary flux analyis were modest at best; in general, the nonlinear confidence intervals of almost all parameters estimated in both the short and long nonstationary experiments (without concentration measurements) were comparable to those of the stationary experiment. This indicates that the gains due to additional sensitivity in the nonstationary measurements are most likely offset by the increased number of total parameters that need to be estimated by NMFA. However, previous research has indicated that more flux information may be available when drawing measurements from the nonstationary regime of a labeling experiment (instead of during the later, isotopically stationary stage), leading to more precise confidence intervals [107]. The reasons behind this observed discrepancy are two-fold. First, our rigorous and accurate determination of nonlinear confidence intervals revealed that the linearized standard errors employed in previous studies are unreliable and often much smaller than the true parameter uncertainty. Second, the total number of measurements were not standardized across all conditions in these previously published studies. Confidence intervals for the stationary case relied upon only one set of measurements whereas the nonstationary cases with multiple time points comprised multiple sets. To ensure a fair comparison, we used replicate measurements so that the total quantity of measurements was equal in every simulated experiment.

When comparing stationary and nonstationary flux estimates across pathways, an interesting trend appears. MFA significantly outperforms NMFA (without concentration measurements) in estimating pentose phosphate pathway fluxes and also

performs better than or equal to NMFA in glycolysis estimation. However, within the TCA cycle we see the opposite: NMFA-estimated fluxes possess greater precision than their MFA counterparts. When we study the network locations of our measured intracellular metabolites in the context of these results, we find that stationary flux analysis gives superior estimates to nonstationary analysis in areas of the network dense with measurements and inferior estimates where measurements are sparse. In the TCA cycle, we measured citrate, succinate, α-ketoglutarate, and malate, while in glycolysis and the pentose phosphate combined we measured only pyruvate (which actually falls on the very border of glycolysis). Presumably, the strength of NMFA (greater measurement sensitivities) is made most manifest in measurement-rich network regions, leading to a greater local flux quality, while the weakness of NMFA (a large parameter space) dominates estimation precision in measurement-poor regions.

Taken together, these observations show that nonstationary flux analysis can be a beneficial tool for flux estimation. In many cases, NMFA will outperform stationary analyses. However, NMFA is not necessarily superior to MFA. The results of these simulated *E. coli* experiments demonstrate that there is no single, overarching experimental design that will guarantee an optimal estimation. In fact, the significantly varying individual flux confidence intervals from experiment to experiment (as seen in Figures 3-2 and 3-3) teach us that simulation can and should be a valuable tool in choosing parameters such as measurement time points. This holds true especially when targeting specific pathways or fluxes of a metabolic network.

Chapter 4

NMFA of Brown Adipocytes

4.1 Introduction

The many complexities of mammalian cells (various organelles, posttranslational processing, required rich medium components, etc.) make them considerably more difficult to model, measure, and analyze. However, these complexities are oftentimes the very things that make these organisms interesting and powerful. In particular, mammalian cells serve as mirrors to our own human cellular metabolism and function.

Brown adipocytes (brown fat cells) are mammalian cells that grow together to form brown fat, a thermogenic tissue that is most obvious in rodents and infant humans [30]. Recently, it has also been shown that even adults have substantial amounts of brown adipose tissue [129]. Adipose metabolism in general is of interest because obesity is a major risk factor for many common diseases, including type 2 diabetes, cardiovascular disease, stroke, hypertension, and many cancers [25]. Brown adipose metabolism in specific is of interest because brown adipocytes have been shown to confer anti-obesity properties, including slower weight gain, higher insulin sensitivity, lower levels of serum-free fatty acids, and protection from diabetes [31, 121]. Brown fat cells produce these effects because of their unique cellular and metabolic capabilities. Besides synthesizing and storing triglycerides, they host very high numbers of mitochondria which are used to generate large amounts of heat via mitochondrial electon transport and fuel oxidation [81].

Developing tools for measuring brown cell fluxes will increase our understanding of the metabolism behind obesity and potentially give us insight into strategies for fighting obesity-related disorders [89, 162]. Here, we have used isotopically nonstationary metabolic flux analysis to study brown adipocytes and gain a basic knowledge of their intracellular flux distribution.

4.2 Methods and Materials

Brown adipocytes were cultured as described by Yoo *et al.* [161]. After cells were mature and at a metabolic steady state, the growth medium was replaced with assay medium containing [U-^{13}C]glucose and unlabeled glutamine. At 2, 4, and 6 hours, cells were quenched and intracellular metabolites were extracted, derivatized, and analyzed with GC/MS. The metabolite fragments that were subsequently measured are listed in Table 4.1. Standard errors of at most 1.5 mol% were assumed for all GC/MS measurements. The adipocyte metabolic network is shown in Figure 4-1. A detailed list of atom transitions for each reaction can be found in Table 4.2.

4.3 Flux Estimation

Over the three time points we collected 249 mass isotopomer abundances while the model attempted to fit 12 free fluxes and 14 metabolite concentrations. Hence, the system possessed $249 - 12 - 14 = 223$ redundant measurements and the expected lower and upper bounds of the 95% confidence region were 183 and 266, assuming that the minimized sum of squared residuals in Equation 2.15 follows a χ^2 distribution. A nonstationary flux estimation was conducted to fit the model to these measurements. The resulting minimized sum of squared residuals was 251, indicating that the fit was statistically acceptable. The estimated labeling profiles are shown together with the measurements in Figure 4-2. Net and exchange flux estimates are presented in Figure 4-3. These parameters' estimated numerical values along with their accurate 95% confidence intervals are listed in Table 4.3. Exchange fluxes are calculated according

Metabolite	Mass	Carbons	Formula
Ala	232	23	$C_{10}H_{26}ONSi_2$
Ala	260	123	$C_{11}H_{26}O_2NSi_2$
Asp	316	234	$C_{15}H_{34}O_2NSi_2$
Asp	418	1234	$C_{18}H_{40}O_4NSi_3$
Cit	459	123456	$C_{20}H_{39}O_6Si_3$
Glu	330	2345	$C_{16}H_{36}O_2NSi_2$
Glu	432	12345	$C_{19}H_{42}\ O_4NSi_3$
Lac	233	23	$C_{10}H_{25}O_2Si_2$
Lac	261	123	$C_{11}H_{25}O_3Si_2$
Mal	419	1234	$C_{18}H_{39}O_5Si_3$
Palm	270	1-16	$C_{17}H_{34}O_2$
Pyr	174	123	$C_6H_{12}O_3NSi$

Table 4.1: Intracellular metabolite fragment MIDs measured in the brown adipocyte study. Metabolite extracts were treated with MBSTFA to form TBDMS-derivatized molecules that were then analyzed by GC/MS.

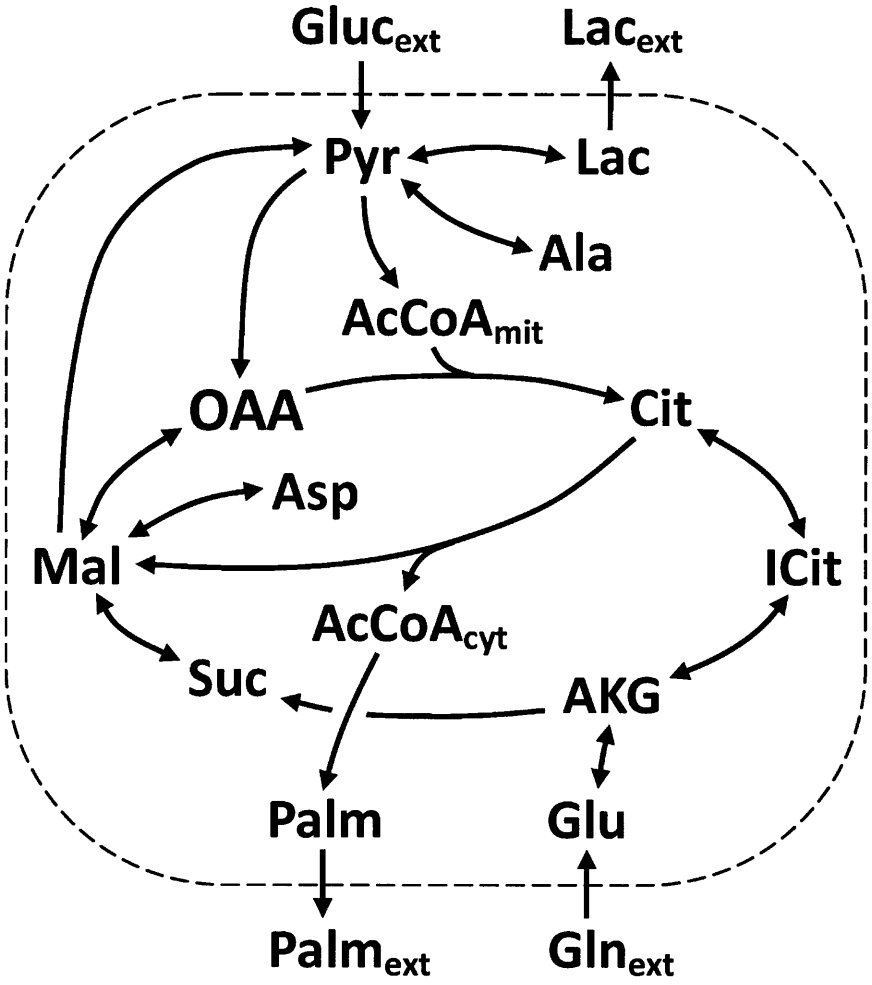

Figure 4-1: A simplified model of brown adipocyte metabolism. (Atom transitions can be found in Table 4.2.) Pentose phosphate activity was assumed negligible, and glycolysis reactions were lumped into a single $Gluc_{ext} \rightarrow$ Pyr reaction. Species present in multiple compartments were assumed to be in rapid exchange across the membrane and all reactions were assumed to occur only in one compartment, allowing us to ignore compartmentalization in the network. Acetyl-CoA was the lone exception; since it does not cross the mitochondrial membrane but is present both in the cytosol and mitochondria, it must be represented in the model by two distinct metabolite states.

Glycolysis and Lactate Reactions			
v_1	$Gluc_{ext}$ (abcdef)	→	Pyr (cba) + Pyr (def)
v_2	Pyr (abc)	↔	Lac (abc)
v_3	Lac (abc)	→	Lac_{ext} (abc)
v_4	Pyr (abc)	→	$AcCoA_{mit}$ (bc) + CO_2 (a)
Anaplerotic Reactions			
v_5	Pyr (abc) + CO_2 (d)	→	OAA (abcd)
v_6	Mal (abcd)	→	Pyr (abc) + CO_2 (d)
Tricarboxylic Acid Cycle			
v_7	$AcCoA_{mit}$ (ab) + OAA (cdef)	→	Cit (fedbac)
v_8	Cit (abcdef)	↔	ICit (abcdef)
v_9	ICit (abcdef)	↔	AKG (abcde) + CO_2 (f)
v_{10}	AKG (abcde)	→	1/2 Suc (bcde) + 1/2 Suc (edcb) + CO_2 (a)
v_{11}	Suc (abcd)	→	1/2 Fum (abcd) + 1/2 Fum (dcba)
v_{12}	Fum (abcd)	↔	1/2 Mal (abcd) + 1/2 Mal (dcba)
v_{13}	Mal (abcd)	↔	OAA (abcd)
Palmitate Synthesis			
v_{14}	Cit (abcdef)	→	$AcCoA_{cyt}$ (ed) + OAA (fcba)
v_{15}	8 $AcCoA_{cyt}$ (ab)	→	Palm (abababababababab)
v_{16}	Palm (abababababababab)	→	$Palm_{ext}$ (abababababababab)
Amino Acid Reactions			
v_{17}	Pyr (abc)	↔	Ala (abc)
v_{18}	OAA (abcd)	↔	Asp (abcd)
v_{19}	AKG (abcde)	↔	Glu (abcde)

Table 4.2: A complete list of reactions and atom transitions for the brown adipocyte model. Simplified versions of glycolysis, anaplerosis, the citric acid cycle, fatty acid metabolism, and amino acid metabolism were considered. Carbon atom transitions are indicated within parentheses. Irreversible and reversible reactions are indicated by the symbols → and ↔, respectively.

to Equation 3.1, where v_{ref} is the glucose uptake flux.

4.4 Discussion

Both the adipocyte experiment and the simulated *E. coli* experiments of the previous chapter confirm that NMFA can yield important information regarding metabolite concentrations even when those concentrations are not directly measured. Admittedly, most metabolite concentrations in the estimation are unidentifiable. The reason is twofold. First, the time intervals between measurements are too large for these pools to influence observable changes in labeling; that is, these pools appear to be at a pseudo-steady state relative to the sampling time scale. Second, because labeling is measured in only a subset of metabolites, there is considerable ambiguity with regard to the concentrations of unmeasured metabolites. At the very least, upper bounds can be determined for all concentrations. We were able estimate both upper and lower bounds for metabolites that (1) were at high enough concentrations to influence the labeling dynamics on a time scale similar to the sampling time scale and (2) were the targets of labeling measurements or were "sandwiched" between other metabolites where labeling measurements available. In the *E. coli* experiment, the concentrations of almost all metabolites whose labeling was measured could be estimated with a 95% confidence interval on the order of ±10% of the estimated value. The adipocyte experiment generated three concentration estimates with nonzero lower bounds. However, these confidence intervals were much less informative (between ±25 and ±55% of the estimated value), most likely because labeling was measured at only three widely spaced time points as opposed to 15 closely spaced time points in the short nonstationary *E. coli* case.

Several factors frame the brown adipocyte system as an ideal subject for nonstationary analysis. Studying the estimated parameters in the system, we find that the measured metabolites do not reach a stationary labeling state (99% of the steady-state value) until between 35 and 55 hours, except for palmitate, which requires 250 hours due to its slow turnover in these cells. In such a case, the utility of nonsta-

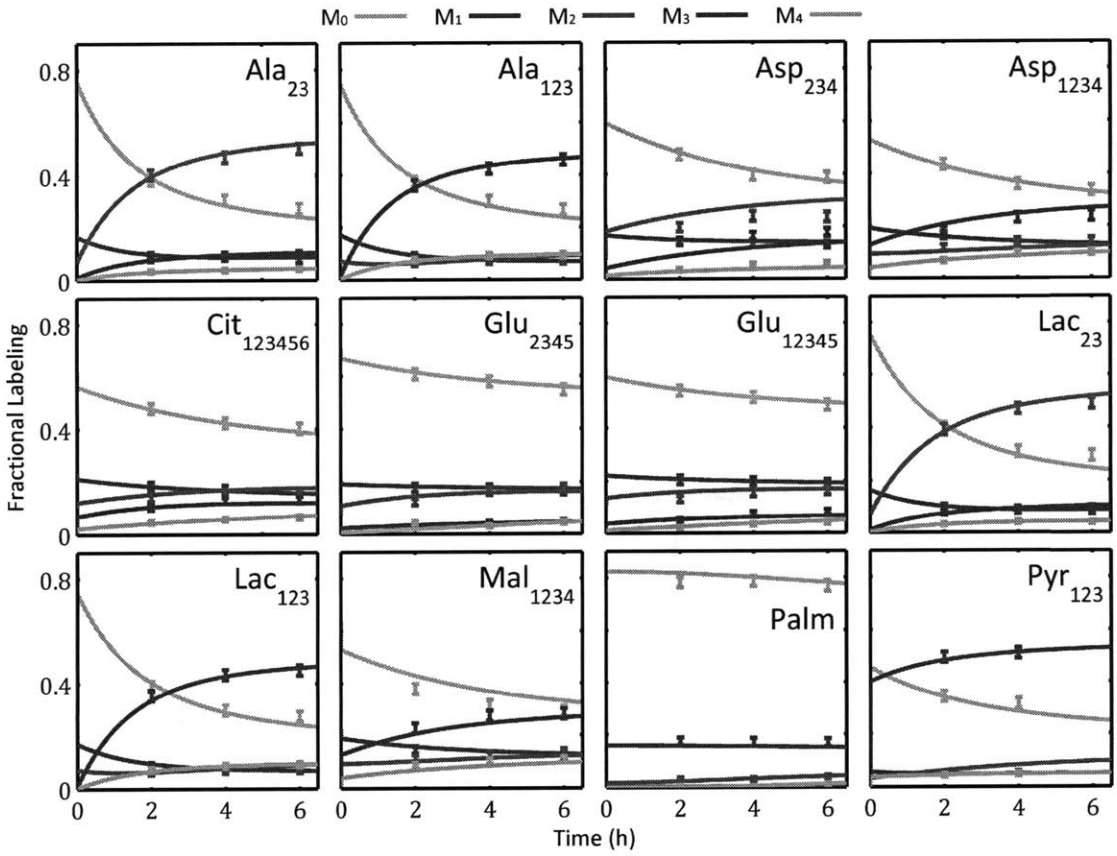

Figure 4-2: Fitted MIDs versus time for the measured metabolite fragments overlaid with actual measurements from the nonstationary brown adipocyte experiment. Only the relative abundance for the mass isotopomers M+0 through M+4 were plotted. Error bars indicate 95% confidence intervals for the measurements.

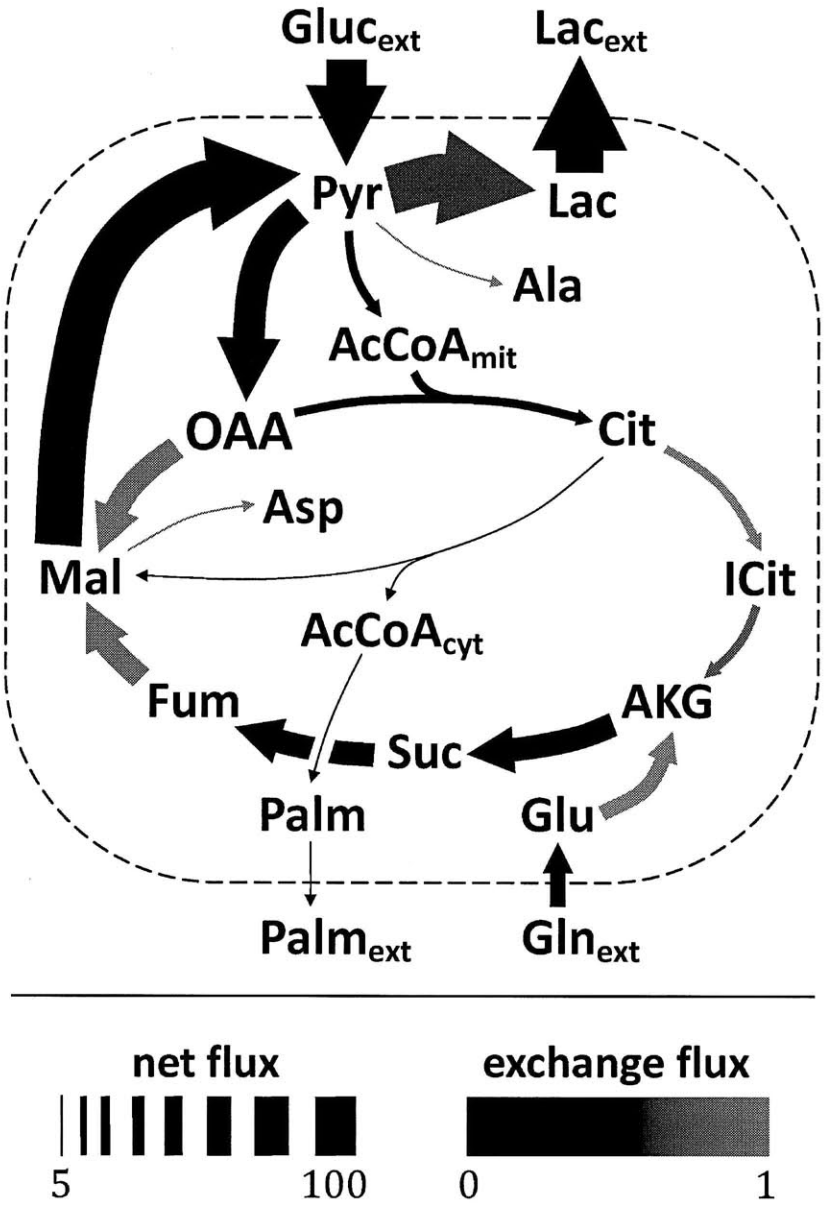

Figure 4-3: Visualization of flux estimates for nonstationary flux analysis of brown adipocytes. Net flux directionality is indicated by arrows and net flux magnitude is proportional to line thickness while exchange flux magnitude correlates to arrow color. Net flux values are in units of h^{-1} and have been scaled such that the $1/2$ $Gluc_{ext}$ $\rightarrow$ Pyr flux is 100 h^{-1}. Exchange flux values are dimensionless and have been scaled according to Equation 3.1.

Parameter	Value	Interval
Pyr → $AcCoA_{mit}$	17.0	[14.3, 22.1]
Suc → Mal	54.0	[45.6, 60.4]
Mal → OAA	-73.1	[-92.4, -55.8]
$AcCoA_{cyt}$ → 1/8 Palm	8.6e-3	[0, 11.9]
Pyr ↔ Lac	0.25	[0, 0.49]
Pyr ↔ Ala	0.00	[0, 0.46]
Cit ↔ ICit	1.00	[0.18, 1]
Glu ↔ AKG	1.00	[0.77, 1]
ICit ↔ AKG	0.23	[0.18, 1]
Fum ↔ Mal	0.53	[0.45, 0.60]
Mal ↔ OAA	1.00	[0.80, 1]
Mal ↔ Asp	0.07	[0, 1]
$AcCoA_{cyt}$ pool	1.0e-5	[0, 6.1]
$AcCoA_{mit}$ pool	126.4	[82.1, 189.9]
AKG pool	3.6e-4	[0, 23.7]
Ala pool	2.2e-3	[0, 75.0]
Asp pool	4.0e-5	[0, 40.0]
Cit pool	2.2e-4	[0, 13.0]
Glu pool	3.6e-4	[0, 23.7]
ICit pool	2.1e-4	[0, 13.0]
Lac pool	178.4	[112.9, 225.6]
Mal pool	8.1e-4	[0, 40.0]
OAA pool	6.7e-4	[0, 40.0]
Palm pool	6.3e-2	[0, 90.3]
Pyr pool	2.8e-3	[0, 64.3]
Suc pool	208.1	[111.6, 323.8]

Table 4.3: Estimated net fluxes (→), exchange fluxes (↔), and concentrations (pool) and their respective 95% confidence intervals for brown adipocyte metabolism. Concentrations are dimensionless and net fluxes have units of h^{-1}; both have been scaled such that the 1/2 $Gluc_{ext}$ → Pyr flux is 100 h^{-1}. Exchange fluxes are dimensionless and scaled according to Equation 3.1.

tionary flux analysis becomes quite apparent. Flux estimates can be obtained in only a fraction of the experimental time, leading to large cost and time savings. Moreove, even if money and time were not limiting factors, it is highly unlikely that the adipocytes could be kept metabolically or phenotypically stable for such lengthy durations, making NMFA not just a convenient tool in this situation, but an essential one.

Chapter 5

Rapid Sampling

5.1 Introduction

Metabolic flux analysis requires isotopic labeling measurements (usually by mass spectrometry) of cellular products. The metabolic products that are targeted for measurement greatly affect the overall design of the flux analysis experiment [32]. The earliest experiments relied upon measurements of proteinogenic amino acids [34, 38]. Because proteinogenic amino acids are in high abundance and because they are relatively stable, no special precautions must be taken when sampling the culture. However, the disadvantage to this method is that considerable amounts of time must pass before the protein pool receives a measureable amount of isotopic label [158]. This requires significant amounts of labeled substrate (which can be expensive) and also requires that the cells be held at metabolic steady state for long durations (which can be difficult).

A more recent experimental advance is to sample not from protein but from intracellular metabolite pools [87, 95, 143, 148]. This requires much less tracer and allows for shorter experimental durations. It also allows for a more topologically comprehensive set of measurements covering the entire metabolism. Instead of only amino acids, which exist on the periphery of central carbon metabolism, metabolites can be sampled from within the TCA cycle, the pentose phosphate pathway, and glycolysis. This improved coverage can result in more precise flux estimates.

However, these advantages come at a price. Intracellular pools are very small [46] and very labile [55, 124]. It is likely that traditional sampling and sample processing will subject cells to different conditions (e.g., reduced oxygen levels, decreased substrate concentrations, temperature shifts, etc.). These conditions can induce changes in cell metabolism which in turn will confound labeling patterns and prevent the determination of true flux values [40]. If isotopic labeling measurements are to be representative of the metabolic state of the culture, great care must be taken at each step in the sampling process:

1. **Sampling:** The sampling step must be conducted as quickly as possible, and must perturb the cell's environment as little as possible. The window in which samples must be drawn depends on the metabolic activity of the organism under study. For instance, *E. coli* pools can turn over in less than one second, and as such, sampling must be conducted at sub-second time scales in order to preserve measurement accuracy [106].

2. **Quenching:** A quenching step must be applied immediately after sampling that effectively and quickly halts all cellular metabolism, usually by inactivating enzymes. Various methods have been employed in the past to this end. Extreme heat can be applied to denature enzymes [127, 159]. Low temperatures can also be used to halt kinetics [151, 152]. Quenching must be conducted carefully so as not to permeabilize the cell membrane and allow intracellular metabolites to leak out of cells (reducing signal strength of the eventual measurement) or, even worse, to allow differently labeled extracellular metabolites to leak into cells, altering the to-be-measured labeling patterns [64].

3. **Extraction:** After quenching, the supernatant and with it all extracellular metabolites must be discarded, cell walls and membranes must be thoroughly broken down, and intracellular metabolites must be extracted from the biomass. Different solvents and schemes can be employed to target different categories of metabolites [56, 64, 91, 152].

For organisms with active metabolisms, these steps must be conducted quickly and efficiently. (Short sampling durations are also necessary when working with more delicate organisms, such as gram-positive bacteria and mammalian cells, to prevent metabolite leakage.) In many cases, manual sampling is not an option and mechanical systems that integrate sampling and quenching must be utilized. (Subsequently, these systems will be simply referred to as "rapid sampling" systems.) Many rapid sampling systems have been devoloped over the past two decades for intracellular metabolite measurement. Some rely on vacuum to power the sampling [73, 84, 142]. Some utilize impressive mechanical systems [29, 125, 127]. Most are able to able to sample with a time step between one and two seconds, and some of the systems are even able to go down to a step as low as 250 ms [125].

5.2 Rapid Sampling Apparatus

We designed and built a rapid sampler to study yeast metabolism. Our aim was to create a sampler that could operate with a time step between one and two seconds. Just as importantly, the sampler needed to be a "low threshold" device; i.e., it had to be relatively inexpensive and straightforward to construct and operate. Also important was that the sampler not require a customized bioreactor; in fact, it should be adaptable to a wide range of reactors and even flasks. Finally, the sampler cannot just sample. It also must quench rapidly, while not hindering the upcoming extraction process.

A schematic of the rapid sampling apparatus is shown in Figure 5-1. Each sample requires an individual sampling tube, shown in detail in Figure 5-2. The tube holds a 60% (v/v) methanol/water solution held at -20°C in a cold bath of ethanol and dry ice. Each tube is capped and airtight but has three exiting lines. The first leads to a manifold which in turn leads to a vacuum pump. Through this, the vacuum is established that will eventually power the sampling. The second line leads to a second manifold which in turn leads to a nitrogen line. At the tube, this nitrogen line feeds into a needle which is submerged in the quenching fluid. This line sparges gas

during the quenching process so that the culture is rapidly mixed and does not freeze upon sampling. The third line collects sample from the culture and is opened and closed by a computer-controlled solenoid valve. The valve's default setting is closed, and only opens during sampling. The free end of the sampling line is placed into the culture before sampling begins.

The timing of each valve was controlled by computer. A custom program was written in LabVIEW that allowed the timing and duration of each valve opening to be specified by the user. The graphical user interface and the block diagram for the LabVIEW program are shown in Figures 5-3 and 5-4. The minimum time step for the sampler was two seconds. The limiting factor was the time necessary for the vacuum to be reestablished after gaining pressure during a sampling.

5.3 Validation of Rapid Sampler

We tested the rapid sampling apparatus in a bioreactor environment to verify its operation. A 75/25% mixture of [1-^{13}C] and [U-$^{13}C_6$]glucose was pulsed into a reactor (previously containing only naturally labeled glucose) at time $t = 0$, immediately after which sampling commenced. Rapid sampling occurred in an approximately geometric sequence, at 2, 4, 8, 30, 60, 120, and 240 seconds. Manual samples were taken at later time points (11, 16, 31, 60, 90, 120, 180, and 240 minutes) for comparison.

Extracellular glucose labeling was measured by subjecting supernatant from each sample to aldonitrile pentapropionate derivatization [12]. 10 µL of combined cell culture and quenching solution was added to 300 µL of cold acetone, vortexed for 10 seconds, and centrifuged for 5 minutes at 18,000 g to remove proteins. The supernatant was removed and evaporated to dryness under air flow at 60°C, after which 50 µL of 2% (w/v) hydroxylamine hydrochloride in pyridine was added (Sigma-Aldrich). The mixture was heated for 60 minutes at 90°C, combined with 100 µL propionic anhydride (Sigma-Aldrich), heated for 30 minutes at 60°C, and evaporated to dryness again. Finally, the extract was combined with 200 µL of ethyl acetate (Sigma-Aldrich), vortexed, and centrifuged for 10 minutes at 18,000 g to remove any solids.

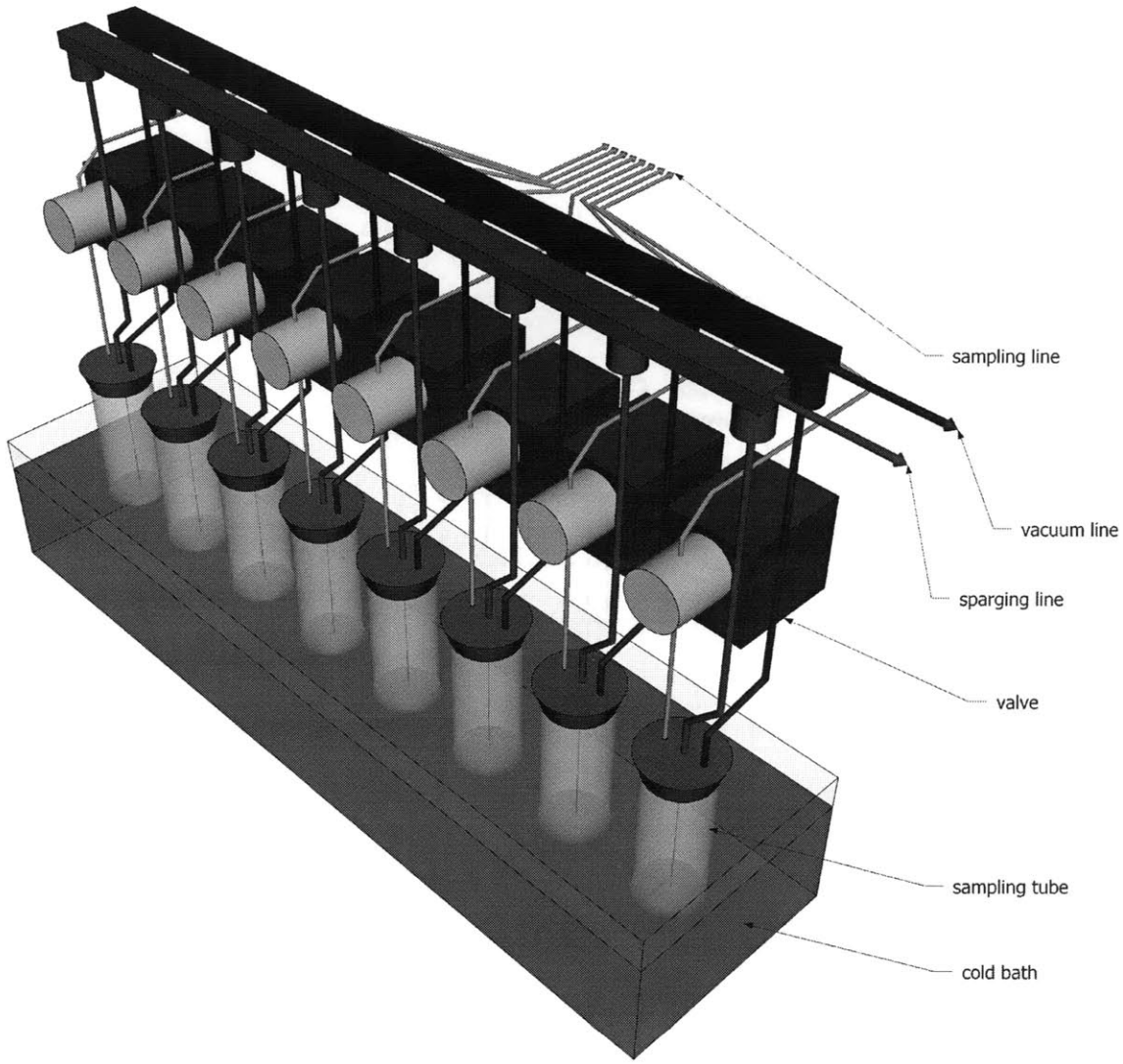

Figure 5-1: Layout of the rapid sampling apparatus. The tubes sit in a cold bath of ethanol and dry ice held at -20°C. Each sampling tube has a vacuum line, sparging line, and a valve-controlled sampling line. Sparging lines lead to a common manifold that is connected to an air or nitrogen source. Vacuum lines lead to a different manifold that is connected first to a trap and second to a vacuum source. Sampling lines are fed into the cell culture. The valves are wired to a connector black controlled by a LabVIEW program on a computer.

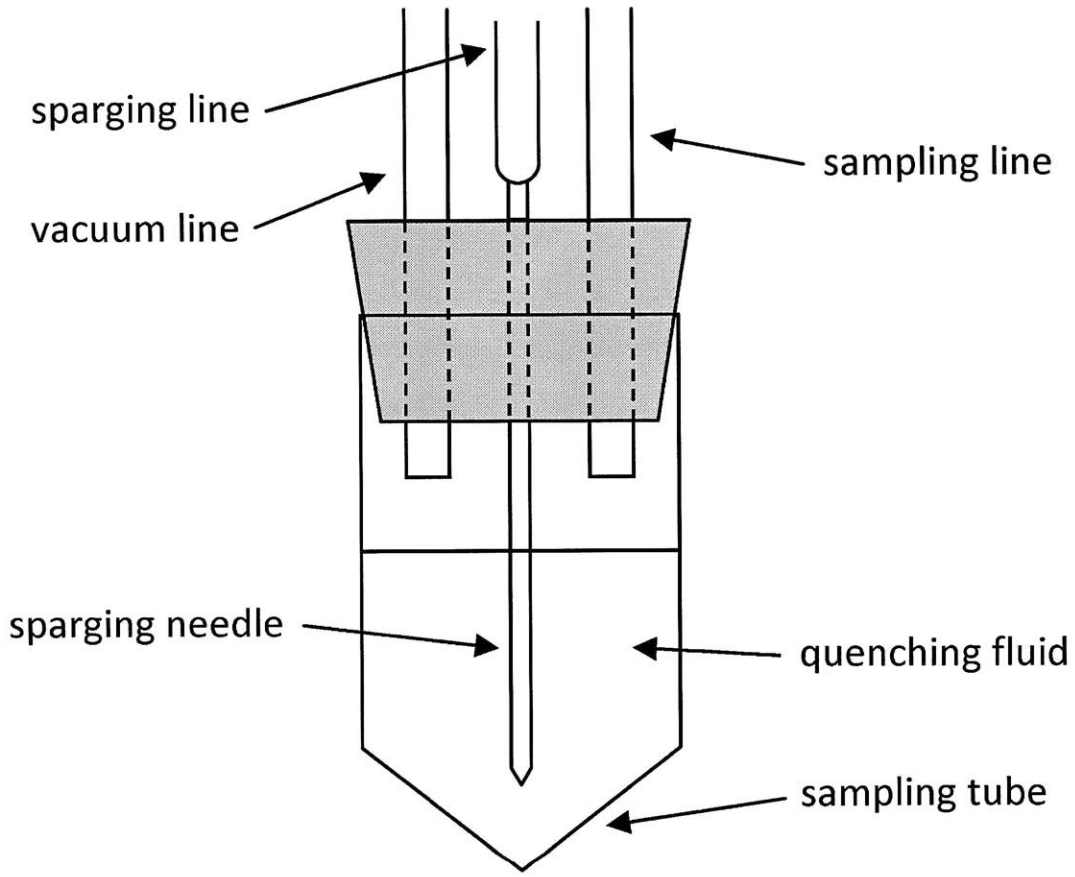

Figure 5-2: Schematic of a single sampling tube. Each tube is capped with a rubber stopper so that it is airtight, except for a vacuum line, a sparging line, and a valve-controlled sampling line that all pass through the stopper. Within the tube, the sparging line connects to a needle submerged in the quenching solution, while the vacuum and sampling lines rest above the solution (and are positioned high enough to remain so even after the addition of sample). The quenching solution consists of 20 mL of 60% (v/v) methanol/water.

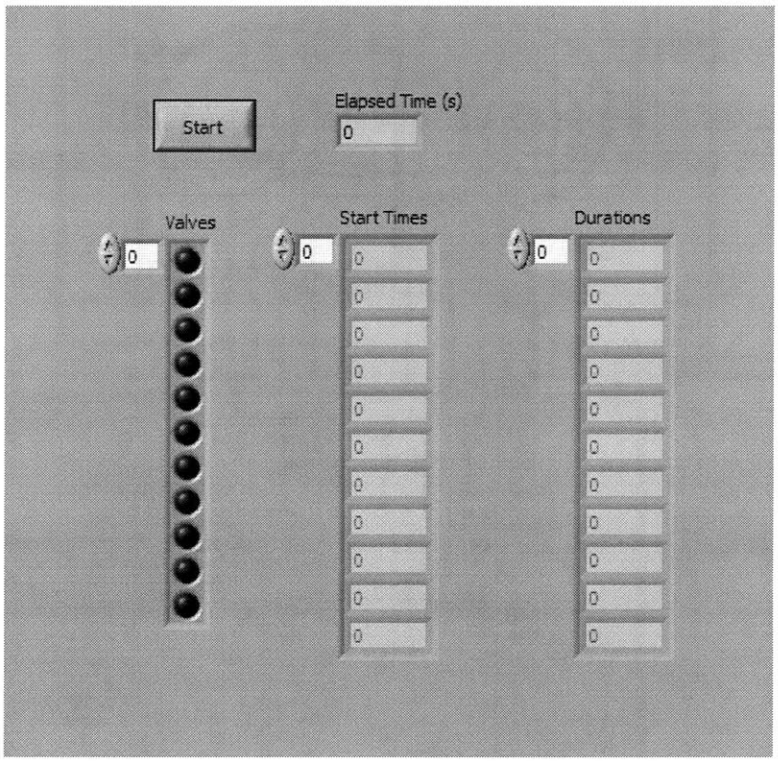

Figure 5-3: LabVIEW graphical user interface for the rapid sampling apparatus. The "Start" button is pushed at time $t = 0$, activating the timer and the sampling sequence. The current time t (in seconds) is shown in the "Elapsed Time" field. Sampling times (also in seconds) are listed in the "Start Times" column; each field corresponds to a different valve. The duration of time that each valve remains open to collect sample is listed in the "Durations" column; we found 0.75 seconds to be an appropriate value, but this will obviously depend on the length of the sampling tubes and the power of the vacuum. Finally, when a valve is currently open, its corresponding indicator in the "Valves" column will flash.

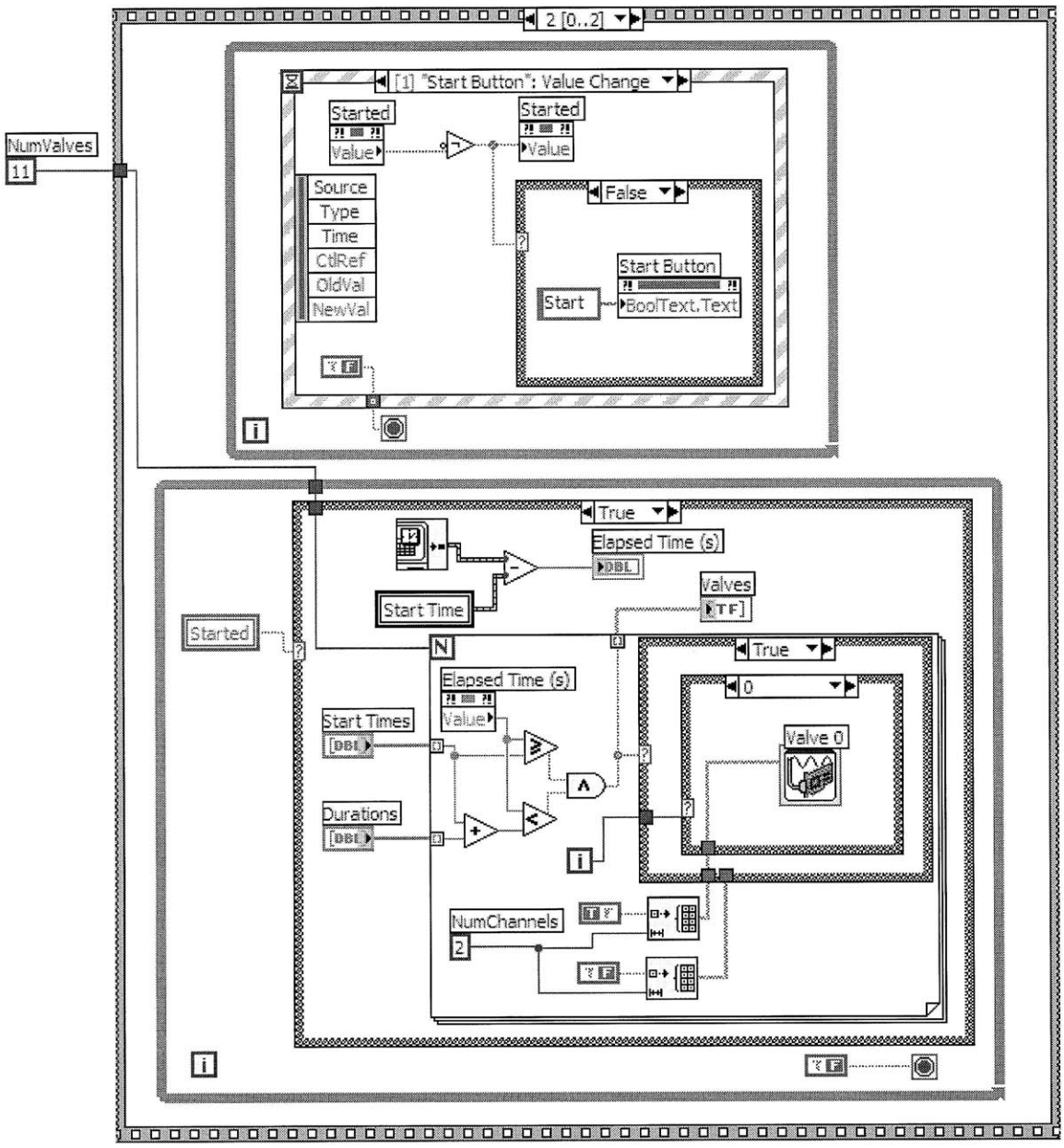

Figure 5-4: A schematic of the LabVIEW block diagram that controls the rapid sampler's valve array. The upper group of functions initates and controls the timing, while the lower group of functions activates the valves one-by-one when the current time matches the user-specified sampling times.

GC/MS analysis was performed useing an Agilent 6890 GC equipped with a 30 m DB-35MS capillary column connected to an Agilent 5975B MS operating under electron (EI) ionization at 70 eV. One µL of sample was injected in splitless mode at 250°C, using helium as the carrier gas at a flow rate of 0.88 mL min^{-1}. The GC oven temperature was held at 80°C for 1 minute, increased to 280°C at 20°C min^{-1}, and held for 4 minutes. The interface temperature was held at 300°C. Glucose labeling was determined from the ion fragment at 370 m/z, corresponding to carbon atoms C_1 through C_5 of glucose with chemical formula $C_{17}H_{24}O_8N$ [12].

We corrected our mass isotopomer distributions for natural isotope enrichments [50] and calculated fractional abundances of naturally labeled, [1-^{13}C], and [U-$^{13}C_6$]glucose by respectively comparing the resulting (corrected) M+0 and M+1 peaks and the sum of the M+4 and M+5 peaks. These isotopomer fractions over time are shown in Figure 5-5. The measurements are extremely consistent, both within the rapid and manual sampling subsets and overall. This indicates first that the mixing time of the reactor is on a much smaller time scale than the sampling and second that the rapid sampling method is not interfering with labeling measurements. Both of these results are important in verifying that the sampler can be used in flux analysis experiments.

Next we used the rapid sampler to capture intracellular isotopic transience in *Yarrowia lipolytica*. As above, a 75/25% bolus of [1-^{13}C] and [U-$^{13}C_6$]glucose was introduced into a cell culture at time $t = 0$ after which rapid sampling was employed in a geometric fashion. Metabolites were extracted, TBDMS-derivatized, and measured using gas chromatography/mass spectrometry as described in Chapter 6. The fractional abundances of M+0, M+1, and M+2 for one particular metabolite, pyruvate, are shown in Figure 5-6.

Because pyruvate is located at the end of glycolysis but before the citric acid cycle, it reaches isotopic steady state very quickly when employing a glucose tracer, presumably because glycolytic fluxes are large while glycolytic metabolite concentrations are small. In addition, the network is linear and direct (if the pentose phosphate flux is negligible). All of these factors make isotopic transience difficult to detect in glycolytic intermediates. However, as Figure 5-6 shows, the rapid sampler is able to

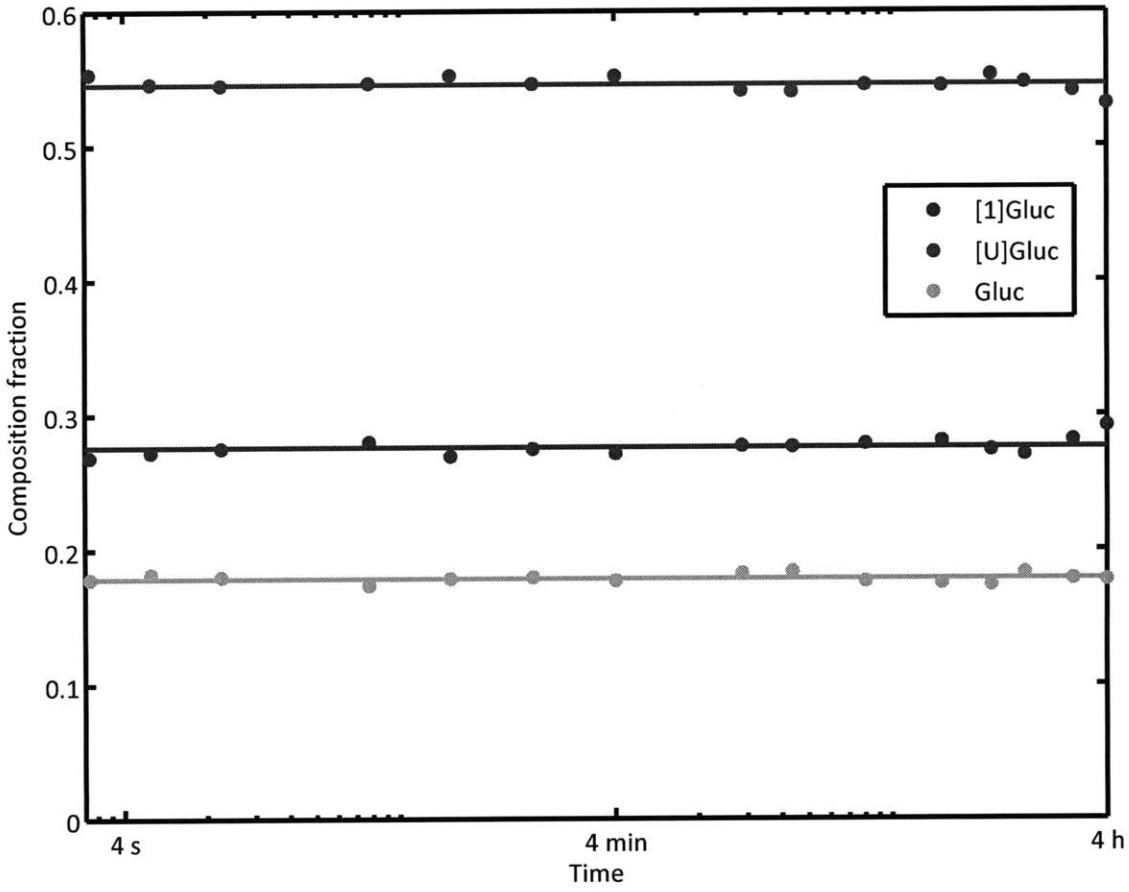

Figure 5-5: Isotopic labeling of extracellular glucose measured using the rapid sampling apparatus. At time t = 0, a 75/25% bolus of [1-^{13}C] and [U-$^{13}C_6$]glucose was introduced into a bioreactor containing only naturally labeled glucose, after which 7 rapid samples and 8 manual samples were drawn. Extracellular glucose labeling was measured for each sample and used to determine glucose's isotopomer fractions in the supernatant over time. These fractions are plotted here. Circles represent different measurements, while lines represent each isotopomer fraction's mean value over all samples.

detect sudden changes in this pathway. These measurements can then be utilized in flux analysis to gain greater confidence in our glycolytic flux and concentration estimations.

5.4 Discussion

We successfully constructed a rapid samping device that met our initial criteria. It is also an affordable, easy-to-construct piece of equipment. The entire sampler costs less than $2000 to construct (and much of that amount goes towards the LabVIEW software, which many labs and businesses already possess). A complete list of materials and more details on the construction and layout of the rapid sampler can be found in Appendix C. The apparatus also performs as originally designed, and can take multiple samples in less than 10 seconds, qualifying it for use in nonstationary flux applications for most organisms.

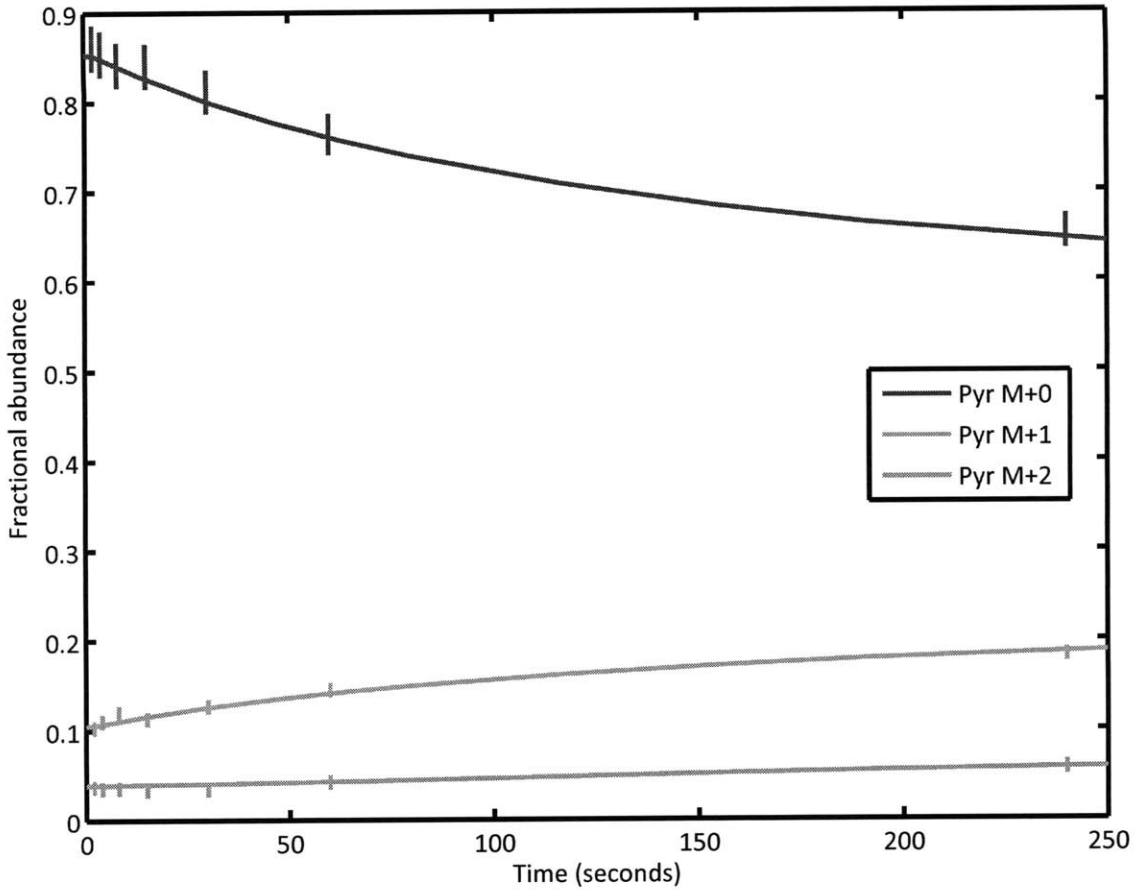

Figure 5-6: Isotopic labeling of intracellular pyruvate as measured by the rapid sampling apparatus. At time t = 0, a 75/25% bolus of [1-^{13}C] and [U-$^{13}C_6$]glucose was introduced into a bioreactor containing only naturally labeled glucose, after which 7 rapid samples were drawn. Intracellular pyruvate labeling was measured for each sample (and eventually used in an overall flux analysis). Vertical lines represent error bars for each measurement, where the height of each bar is equal to ± two standard errors.

Chapter 6

Isotopic Tracer Evaluation for Metabolic Flux Analysis

6.1 Introduction

Flux estimation can be a powerful tool for comparing and understanding metabolism and cellular phenotype. However, without accompanying ideas of precision for each estimated parameter value, flux analysis is relatively meaningless. For almost any practical analysis, parameter estimates will have wildly varying levels of confidence, and we can only assign biological meaning to those with reasonable precision.

We have been careful to report confidence intervals alongside parameter estimates throughout all results in this thesis. Unfortunately, this is not always standard practice, though it should be for the aforementioned reason. As we have explained previously in Chapter 2, calculating linear parameter standard errors using the inverse Hessian of the objective function yields poor results in most nonstationary analyses, so we compute confidence intervals via parameter continuation around the optimal solution [10]. These accurate confidence intervals allow us to correctly determine which flux values are relevant. Non-interlapping confidence intervals also indicate fluxes that are statistically significantly different when comparing two flux distributions [101, 102].

The choice of tracer plays a major role in the eventual confidence interval cal-

culations since the tracer dictates the mass isotopomer distribution (MID) of each metabolite for a given set of fluxes. The sensitivity of the MIDs, in turn, to changes in the pathway fluxes ultimately determines the confidence of flux estimates, which are as important as the flux values themselves [10]. Stationary MFA is conducted when the labeled substrate is at isotopic steady state and does not utilize pool size or transient data; as such this technique is especially reliant upon the specific tracer used. Depending on the particular bioreaction network, nonstationary MFA may be preferable and can also benefit from a informed choice of tracer [107, 108]. The issue of tracer choice is more complex in mammalian cell systems that utilize multiple carbon sources (e.g., glucose and glutamine) and are grown in complex media. To probe specific pathways, researchers have applied a wide array of isotopically labeled substrates, including glucose, glutamine, or atypical substrates such as propionate or succinate [28, 160, 161]. For example, [1,2-$^{13}C_2$]glucose is commonly employed for analysis of the pentose phosphate pathway (PPP) [23], whereas [3-^{13}C]glucose or [3-^{14}C]glucose provide information on pyruvate oxidation [72, 102]. Many tracers effectively label metabolites in the tricarboxylic acid (TCA) cycle, but the presence of anaplerotic reactions and glutamine incorporation (through glutaminolysis or reductive carboxylation) make it difficult to identify the optimal tracer for measuring all net and exchange fluxes within the network.

As we discussed in Chapter 4, the usefulness of flux analysis in mammalian systems is increasingly apparent and as such research in that area in growing [78, 103, 122]. And as our usage of flux analysis increases in mammalian cell applications, it becomes more important that we are able to design experiments to maximize meaningful information gain. Researchers have recently described methods to optimize measurement sets for flux determination [32, 117]; however, as discussed above, the choice of tracer is an equally important parameter. Because the tracer strongly influences flux estimation quality and can usually be chosen from a wide array of available isotopic substrates, judicious tracer selection is a major component of this experimental design process. As a result, a systematic analysis of available ^{13}C-labeled tracers and associated confidence intervals (i.e., sensitivities) for each estimated flux is warranted.

Optimization of tracer choice will enable researchers to more precisely measure specific fluxes in high-throughput applications that aim to screen the metabolic effects of drugs in cells [22, 130].

Some initial investigation has occurred in this area. Sriram and colleagues have conducted a Monte Carlo-based analysis of experimental precisions (using standard deviations) when using different combinations of [U-$^{13}C_6$]glucose, [1-^{13}C]glucose, and naturally labeled glucose in a mammalian network [134]. Similar studies have been performed in microbial systems, which typically consume a single carbon source [107]. However, a detailed investigation of confidence intervals generated from different, unique tracers in a two-substrate network has never been completed.

Here we have experimentally determined the flux network in a carcinoma cell line and used these data to calculate the confidence intervals for each flux when using different ^{13}C-labeled glucose and glutamine tracers. An elementary metabolite unit (EMU)-based method was used to rapidly estimate flux profiles and confidence intervals from simulated measurements in stationary MFA experiments [11, 164]. We were able to quantitatively validate the effectiveness of specific ^{13}C tracers in a cancer cell network and identify the best choice for analysis of individual reactions and pathways. A scoring algorithm was employed to determine the optimal tracer for the overall model and for subnetworks representing glycolysis, the PPP, and the tricarboxylic acid (TCA) cycle. These results may significantly improve the efficiency of MFA experiments in high-throughput applications and clinical samples where biological material is limited [24].

6.2 Cell Culture and Metabolite Extraction

The A549 lung carcinoma cell line was obtained from ATCC and maintained in high-glucose DMEM supplemented with 4 mM glutamine, 10% FBS, and 100 U mL^{-1} penicillin/streptomycin (Invitrogen). For labeling experiments, semi-confluent cells in a 10-cm dish were cultured in glucose-free DMEM (Sigma) with the above supplements and a 25 mM 1:1 mixture of [U-$^{13}C_6$]glucose and [1-^{13}C]glucose (Cambridge

Isotope Laboratories) for 6 hours to achieve isotopic steady state. Spent medium was collected and analyzed for glucose, lactate, and glutamine consumption on a YSI 7100 system (YSI Life Sciences). Cells were quenched with 1 mL ice cold methanol, an equal volume of water was added, and cells were collected with a cell scraper. Four volumes of chloroform were added, and the cells were vortexed and held on ice for 30 minutes for deproteinization. After addition of 2 mL water, samples were centrifuged at 3000 g for 20 minutes at 4°C. The aqueous phase was collected in a new tube and evaporated under airflow at room temperature.

6.3 Derivatization and GC/MS Measurements

Dried polar metabolites were dissolved in 60 µL of 2% methoxyamine hydrochloride in pyridine (Pierce), sonicated for 30 minutes, and held at 37°C for 2 hours. After dissolution and reaction, 90 µL MBTSTFA + 1% TBDMCS (Pierce) was added and samples were incubated at 55°C for 60 minutes. Gas chromatography/mass spectrometry (GC/MS) analysis was performed using an Agilent 6890 GC equipped with a 30 m DB-35MS capillary column connected to an Agilent 5975B MS operating under electron impact (EI) ionization at 70 eV. One µL of sample was injected in splitless mode at 270°C, using helium as the carrier gas at a flow rate of 1 mL min^{-1}. The GC oven temperature was held at 100°C for 3 min and increased to 300°C at 3.5°C min^{-1} for a total run time of approximately 60 min. The MS source and quadrupole were held at 230°C and 150°C, respectively, and the detector was operated in selected ion monitoring (SIM) mode. MIDs were obtained for each measured metabolite and incorporated with extracellular flux measurements for flux determination. The identity and values of these measured fragments and fluxes are listed in Tables 6.1 and 6.2.

6.4 Flux Estimation

Intracellular fluxes were estimated for a model reaction network by minimizing the lack of fit between actual and simulated flux and GC/MS measurements. The network

Reaction	Flux	Standard deviation
$Gluc_{ext} \rightarrow G6P$	45	4.5
$Lac \rightarrow Lac_{ext}$	75	7.5
$Gln_{ext} \rightarrow Gln$	10	1.0
Amino acid $\rightarrow$ Biomass	1	0.1

Table 6.1: Extracellular flux measurements for experimental metabolic flux analysis of the A549 carcinoma cell. Extracellular glucose, lactate, and glutamine concentrations were measured by YSI at the start and end of the experiment, and using these results fluxes were calculated. The biomass flux was calculated using the population's growth rate. Standard deviations were estimated to be 10% of the respective flux value. Fluxes and standard deviations are in units of nmol min^{-1} mg^{-1}.

Met	Carbons	Formula	M_0	M_1	M_2	M_3	M_4	M_5	M_6	Stdev
AKG	12345	$C_{14}H_{28}O_5NSi_2$	0.309	0.214	0.054	0.371	0.035	0.017		0.010
Ala	23	$C_{10}H_{26}ONSi_2$	0.297	0.212	0.377	0.079	0.031	0.004	0.001	0.005
Ala	123	$C_{11}H_{26}O_2NSi_2$	0.289	0.212	0.074	0.332	0.065	0.029		0.005
Asp	12	$C_{14}H_{32}O_2NSi_2$	0.571	0.186	0.191	0.037	0.013	0.002		0.005
Asp	1234	$C_{18}H_{40}O_4NSi_3$	0.563	0.188	0.074	0.137	0.027	0.011	0.002	0.005
Cit	123456	$C_{20}H_{39}O_6Si_3$	0.692	0.179	0.097	0.019	0.009	0.002	0.003	0.005
Gln	12345	$C_{19}H_{43}O_3N_2Si_3$	0.608	0.190	0.127	0.036	0.017	0.007	0.004	0.005
Glu	2345	$C_{16}H_{36}O_2NSi_2$	0.536	0.220	0.140	0.063	0.023	0.008	0.004	0.005
Glu	12345	$C_{17}H_{36}O_3NSi_2$	0.641	0.204	0.113	0.026	0.008	0.003	0.002	0.005
Glu	12345	$C_{19}H_{42}O_4NSi_3$	0.544	0.227	0.137	0.060	0.020	0.006	0.002	0.005
Lac	23	$C_{10}H_{25}O_2Si_2$	0.619	0.205	0.127	0.035	0.012	0.002	0.001	0.003
Lac	123	$C_{11}H_{25}O_3Si_2$	0.598	0.207	0.132	0.040	0.017	0.005	0.001	0.003
Mal	1234	$C_{14}H_{39}O_5Si_3$	0.552	0.223	0.149	0.049	0.019	0.006	0.002	0.005
Pyr	123	$C_6H_{12}O_3NSi$	0.624	0.233	0.107	0.025	0.006	0.001	0.001	0.005
Suc	1234	$C_8H_{25}O_4Si_2$	0.357	0.211	0.243	0.105	0.053	0.021	0.007	0.010

Table 6.2: Intracellular mass isotopomer measurements for experimental metabolic flux analysis of the A549 carcinoma cell. Organic and amino acids were TBDMS-derivatized and their relative labeling was measured with GC/MS.

contained simplified versions of glycolysis, the PPP, anaplerotic reactions, the TCA cycle, and amino acid biosynthesis. Table 6.3 lists all network reactions and atom transitions. We additionally calculated 95% confidence intervals for each flux using parameter continuation [10]. All flux simulation, estimation, and continuation in this study was conducted using Metran, a flux analysis tool built upon an EMU framework [11, 161, 164]. Some fluxes were virtually unidentifiable and approximate values were obtained from the literature [92, 102].

The following assumptions regarding the network and metabolism were made to obtain a successful fit:

1. Proliferating A549 cells were at metabolic steady state, and labeled tracers ([U-$^{13}C_6$] and [1-^{13}C]glucose) achieved isotopic steady state within the 6 hours of application. This assumption was supported by unpublished observations in our laboratory and various results in the literature [71, 102], which have demonstrated that glucose tracers achieve isotopic steady state within glycolysis, the pentose phosphate pathway, and the TCA cycle within 4 hours.

2. All CO_2 reincorporated into the system is unlabeled.

3. Succinate and fumarate are symmetric molecules and retain no particular orientation.

4. With the exception of Acetyl-CoA, a single pool for each metabolite was present in the model. Any metabolites present in multiple compartments were assumed to be at isotopic equilibrium.

5. Isotope enrichment in succinate was much less than that of adjacent TCA cycle metabolites. As such, we included a dilution flux of unlabeled succinate to the measured pool, mimicking channeling or isolated compartmentalization of this metabolite [33]. The dilution flux was isolated from the network and did not act as a source of unlabeled material to the TCA cycle.

6. Alanine labeling was significantly diluted relative to pyruvate and lactate. The cause of this discrepancy was assumed to be exchange with unlabeled amino

Glycolysis			
v_1	$Gluc_{ext}$ (abcdef)	→	G6P (abcdef)
v_2	G6P (abcdef)	↔	F6P (abcdef)
v_3	F6P (abcdef)	→	DHAP (cba) + GAP (def)
v_4	DHAP (abc)	↔	GAP (abc)
v_5	GAP (abc)	↔	3PG (abc)
v_6	3PG (abc)	→	Pyr (abc)
v_7	Pyr (abc)	↔	Lac (abc)
v_8	Lac (abc)	→	Lac_{ext} (abc)
Pentose Phosphate Pathway			
v_9	G6P (abcdef)	→	P5P (bcdef) + CO_2 (a)
v_{10}	P5P (abcde) + P5P (fghij)	↔	S7P (abfghij) + GAP (cde)
v_{11}	S7P (abcdefg) + GAP (hij)	↔	F6P (abchij) + E4P (defg)
v_{12}	P5P (abcde) + E4P (fghi)	↔	F6P (abfghi) + GAP (cde)
Tricarboxylic Acid Cycle/Anaplerosis			
v_{13}	Pyr (abc)	→	$AcCoA_{mit}$ (bc) + CO_2 (a)
v_{14}	OAA (abcd) + $AcCoA_{mit}$ (ef)	→	Cit (dcbfea)
v_{15}	AKG (abcde) + CO_2 (f)	↔	Cit (abcdef)
v_{16}	AKG (abcde)	→	Suc (bcde) + CO_2 (a)
v_{17}	Suc (abcd)	↔	1/2 Fum (abcd) + 1/2 Fum (dcba)
v_{18}	Fum (abcd)	↔	1/2 Mal (abcd) + 1/2 Mal (dcba)
v_{19}	OAA (abcd)	↔	Mal (abcd)
v_{20}	Pyr (abc) + CO_2 (d)	→	OAA (abcd)
v_{21}	Mal (abcd)	↔	Pyr (abc) + CO_2 (d)
Amino Acid Metabolism			
v_{22}	Gln_{ext} (abcde)	→	Gln (abcde)
v_{23}	Gln (abcde)	→	Glu (abcde)
v_{24}	Glu (abcde)	↔	AKG (abcde)
v_{25}	Pyr (abc) + Glu (defgh)	→	Ala (abc) + AKG (defgh)
v_{26}	OAA (abcd) + Glu (efghi)	→	Asp (abcd) + AKG (efghi)
v_{27}	3PG (abc) + Glu (defgh)	→	Ser (abc) + AKG (defgh)
v_{28}	Ser (abc)	→	Gly (ab) + MEETHF (c)
Biomass Formation			
v_{29}	P5P (abcde)	→	NTP (abcde)
v_{30}	DHAP (abc)	→	G3P (abc)
v_{31}	Cit (abcdef)	→	$AcCoA_{cyt}$ (ed) + OAA (fcba)
v_{32}	$AcCoA_{cyt}$ (ab)	→	FA (ab)
v_{33}	0.18 Asp + 0.23 Glu + 0.17 Ser + 0.11 Gly + 0.15 Ala + 0.16 Gln	→	Biomass

Table 6.3: A complete list of reactions and atom transitions for the A549 carcinoma model used in both the experimental and simulated flux analysis studies. Carbon atom transitions are indicated within parentheses. Irreversible and reversible reactions are indicated by the symbols → and ↔, respectively.

acids released from protein. A dilution flux was included to account for the observed isotopic dilution.

7. Oxidative sources of unlabeled carbon to the TCA cycle (e.g., amino acids and fatty acids) were assumed to be negligible.

8. Biomass flux measurements were included in the model fit with standard errors of 10%; these fluxes included the synthesis of nucleotides, GLP (for phospholipid production), and proteins (from selected amino acids). Values were estimated based upon known values of cellular composition (weight percent) [2] and an observed doubling time of 20 hours. Amino acid fluxes to biomass that were present in the model were scaled according to the observed amino acid ratio of mammalian cell biomass [110]. No measurement was included for fatty acid synthesis (citrate lyase) and this flux was freely fitted within the model. The fitted value for this flux fell within the expected range for biomass synthesis.

6.5 Tracer Evaluation

The effects of 18 ^{13}C-labeled tracers (11 glucose and 7 glutamine) on flux estimation precision were evaluated (see Table 6.4). Tracers were chosen if commercially available or if previously cited in literature. The effectiveness of each tracer was gauged as follows:

1. A defined set of extracellular flux measurements (Table 6.1) and GC/MS measurements (Table 6.5) was simulated for a given tracer.

2. Standard errors of 5% for external flux measurements and 0.1–1 mol% for GC/MS MIDs were introduced randomly and normally.

3. Flux values and 95% confidence intervals were determined for each reaction.

The usefulness of a tracer was assumed to be directly linked to the precision with which it was able to estimate fluxes of interest; i.e., tracers producing narrower confidence intervals have greater value.

Tracer	Abbreviation
[1-^{13}C]glucose	[1]Gluc
[1,2-$^{13}C_2$]glucose	[1,2]Gluc
[1,6-$^{13}C_2$]glucose	[1,6]Gluc
[2-^{13}C]glucose	[2]Gluc
[3-^{13}C]glucose	[3]Gluc
[3,4-$^{13}C_2$]glucose	[3,4]Gluc
[4-^{13}C]glucose	[4]Gluc
[4,5-$^{13}C_2$]glucose	[4,5]Gluc
[5-^{13}C]glucose	[5]Gluc
[6-^{13}C]glucose	[6]Gluc
[U-$^{13}C_6$]glucose	[U]Gluc
[1-^{13}C]glutamine	[1]Gln
[1,2-$^{13}C_2$]glutamine	[1,2]Gln
[3-^{13}C]glutamine	[3]Gln
[3,4-$^{13}C_2$]glutamine	[3,4]Gln
[4-^{13}C]glutamine	[4]Gln
[5-^{13}C]glutamine	[5]Gln
[U-$^{13}C_5$]glutamine	[U]Gln

Table 6.4: Glucose and glutamine tracers chosen for evaluation and their corresponding abbreviations used throughout Chapters 6 and 7. All commercially available glucose and glutamine tracers were chosen for analysis.

Metabolite	Carbons
3PG	123
AKG	12345
Ala	23
Ala	123
Asp	12
Asp	12
Asp	234
Asp	1234
Cit	123456
Gln	12345
GLP	123
Glu	2345
Glu	12345
Glu	12345
Gly	2
Gly	12
Lac	23
Lac	123
Mal	1234
P5P	12345
Pyr	123
Ser	12
Ser	23
Ser	123
Suc	1234

Table 6.5: Intracellular mass isotopomer measurements for simulated metabolic flux analysis of the A549 carcinoma cell. These simulated measurements are similar to the actual experimental measurements listed in Table 6.2, but we assumed a few additional measurements would be available, namely, 3PG, GLP, and P5P (via TMS-derivatization) and glycine (via TBDMS-derivatization).

6.6 Precision Scoring

To more easily compare estimate precision on a group basis, we created a precision scoring metric. Similar optimality criteria have been used previously for experimental design [107, 108]; however, these earlier methods have been based on the parameter covariance matrix, which assumes linearity and does not always truly capture the nonlinear, constrained systems studied in flux analysis. This precision scoring metric relies on the more robust and accurate nonlinear confidence intervals obtained via parameter continuation [10]. First, a normalized range is calculated for each flux using the formula

$$r_i = \min\left(\frac{u_i}{|v_i|}, \frac{v_i}{|v_i|} + \alpha\right) - \max\left(\frac{l_i}{|v_i|}, \frac{v_i}{|v_i|} - \alpha\right) \tag{6.1}$$

where v_i, l_i, u_i, and r_i are the estimated flux, lower bound, upper bound, and normalized range for the ith flux, and α is a cut-off parameter that prevents one excessively distant bound from overly influencing the scoring. The individual ranges are next converted into scores using a negative exponential function and summed into a final overall score via the expression:

$$S = \sum_i w_i \exp\left(-\frac{r_i}{\beta}\right) \tag{6.2}$$

where w_i is a weighting parameter for the ith flux, β is a range-scaling parameter, and S is the overall precision score. We empirically found that values of 1 and 3 for α and β result in a good dynamic range of scores. If each w_i can be simply zero or one (serving to either exclude or include fluxes in the overall score), each flux's precision score will range between zero (unidentifiable) and one (perfectly identifiable). The overall score will then range between zero and the number of fluxes under consideration.

Because the upper and lower bounds of any given confidence interval are sensitive to the random error introduced into the simulated measurements, the corresponding precision scores will also vary for simulated experiments with different random errors. To account for this, we conducted six simulated experiments for every tracer of interest

and generated a distribution of precision scores, allowing us to report a mean precision score and a precision score standard deviation.

6.7 Experimental Flux Analysis

The metabolic phenotype of a cell is a key component of its overall behavior, and evidence suggests that metabolism plays an important role in maintaining the tumor phenotype [41]. Metabolic analysis of cancer cells has again become an active area of research, and technological improvements have expanded our ability to investigate metabolism using stable isotopically labeled substrates. In this study we used the A549 cancer cell line as a model system for evaluating isotopic tracers in mammalian cells. This line is often used in molecular and metabolic studies of cancer and exhibits aerobic glycolysis, commonly known as the Warburg effect [35, 67]. To obtain baseline values for our metabolic network we estimated fluxes of semi-confluent cells using an equimolar mixture of [U-$^{13}C_6$]glucose and [1-^{13}C]glucose, a combination of tracers commonly used in the literature [107, 134]. The actual proportion of each tracer and naturally labeled glucose from serum was obtained from GC/MS measurements of extracellular glucose. Extracellular fluxes and MIDs of intracellular metabolites were used to estimate the flux distribution (see Tables 6.1 and 6.2 for all measurement data). The system possessed 101 redundant measurements and the expected upper bound of the 95% confidence region is 130, assuming that the minimized sum of squared residuals (SSR) followed a χ^2 distribution. Flux estimation resulted in a minimized SSR of 52, indicating that the fit was statistically acceptable. Estimated values and 95% confidence intervals for each independent flux are listed in Table 6.6. The overall metabolic network is depicted in Figure 6-1. Several exchange fluxes were difficult to precisely determine; in these cases, values were culled from literature.

As expected, the cancer cells displayed a high glycolytic flux and excreted most of the carbon as lactate. Approximately 15% of the glucose flux was diverted to the pentose phosphate shunt. However, these estimations did not include explicit measurements of pentose phosphate intermediates; as such, exchange fluxes within this

Glycolysis	Flux	Interval
$Gluc_{ext} \rightarrow G6P$	38.6	[34.7, 43.2]
$G6P \rightarrow F6P$	32.4	[29.7, 35.1]
$G6P \leftrightarrow F6P$	0.0	[0.0, Inf]
$F6P \rightarrow DHAP + GAP$	36.5	[32.6, 40.7]
$DHAP \rightarrow GAP$	35.5	[31.6, 39.7]
$DHAP \leftrightarrow GAP$	0.0	[0.0, 258.7]
$GAP \rightarrow 3PG$	74.0	[65.7, 82.8]
$GAP \leftrightarrow 3PG$	0.0	[0.0, Inf]
$3PG \rightarrow Pyr$	73.1	[64.8, 81.9]
$Pyr \rightarrow Lac$	73.6	[65.8, 83.3]
$Pyr \leftrightarrow Lac$	13,420	[0.0, Inf]
$Lac \rightarrow Lac_{ext}$	73.6	[65.8, 83.3]

Pentose phosphate pathway	Flux	Interval
$G6P \rightarrow P5P + CO_2$	6.2	[5.4, 7.4]
$2\ P5P \rightarrow S7P + GAP$	2.0	[1.8, 2.4]
$2\ P5P \leftrightarrow S7P + GAP$	19,290	[0.0, Inf]
$S7P + GAP \rightarrow F6P + E4P$	2.0	[1.8, 2.4]
$S7P + GAP \leftrightarrow F6P + E4P$	0.0	[0.0, 0.4]
$P5P + E4P \rightarrow F6P + GAP$	2.0	[1.8, 2.4]
$P5P + E4P \leftrightarrow F6P + GAP$	0.0	[0.0, 0.6]

TCA cycle/anaplerosis	Flux	Interval
$Pyr \rightarrow AcCoA_{mit} + CO_2$	8.6	[5.5, 11.0]
$OAA + AcCoA_{mit} \rightarrow Cit$	8.6	[5.5, 11.0]
$AKG + CO_2 \rightarrow Cit\ (abcdef)$	2.0	[0.2, 3.6]
$AKG + CO_2 \leftrightarrow Cit\ (abcdef)$	5.7	[4.6, 7.1]
$AKG \rightarrow Suc + CO_2$	8.2	[6.6, 10.7]
$Suc \rightarrow Fum$	8.2	[6.6, 10.7]
$Suc \leftrightarrow Fum$	0.2	[0.0, Inf]
$Fum \rightarrow Mal$	8.2	[6.6, 10.7]
$Fum \leftrightarrow Mal$	23.8	[5.2, Inf]
$OAA \rightarrow Mal$	1.4	[0.8, 2.9]
$OAA \leftrightarrow Mal$	106,200	[35.9, Inf]
$Pyr + CO_2 \rightarrow OAA$	0.0	[0.0, 1.7]
$Mal \rightarrow Pyr + CO_2$	9.6	[8.5, 10.7]
$Mal \leftrightarrow Pyr + CO_2$	1.3	[0.9, 1.6]

Amino acid metabolism	Flux	Interval
$Gln_{ext} \rightarrow Gln$	11.5	[10.4, 12.6]
$Gln \rightarrow Glu$	11.0	[9.9, 12.1]
$Glu \rightarrow AKG$	8.2	[7.0, 9.3]
$Glu \leftrightarrow AKG$	81.3	[35.9, 480.3]
$Pyr + Glu \rightarrow Ala + AKG$	0.5	[0.4, 0.6]
$OAA + Glu \rightarrow Asp + AKG$	0.6	[0.5, 0.7]
$3PG + Glu \rightarrow Ser + AKG$	0.9	[0.8, 1.1]
$Ser \rightarrow Gly + MEETHF$	0.4	[0.3, 0.4]

Biomass formation	Flux	Interval
$P5P \rightarrow NTP$	0.1	[0.1, 0.1]
$DHAP \rightarrow G3P$	1.0	[0.8, 1.2]
$Cit \rightarrow AcCoA_{cyt} + OAA$	10.7	[5.3, 14.5]
$AcCoA_{cyt} \rightarrow FA$	10.7	[5.3, 14.5]
$Amino\ acids \rightarrow Biomass$	3.3	[2.7, 4.0]

Table 6.6: Experimentally determined net ($\rightarrow$) and exchange ($\leftrightarrow$) fluxes and 95% flux confidence intervals for the A549 carcinoma cell.

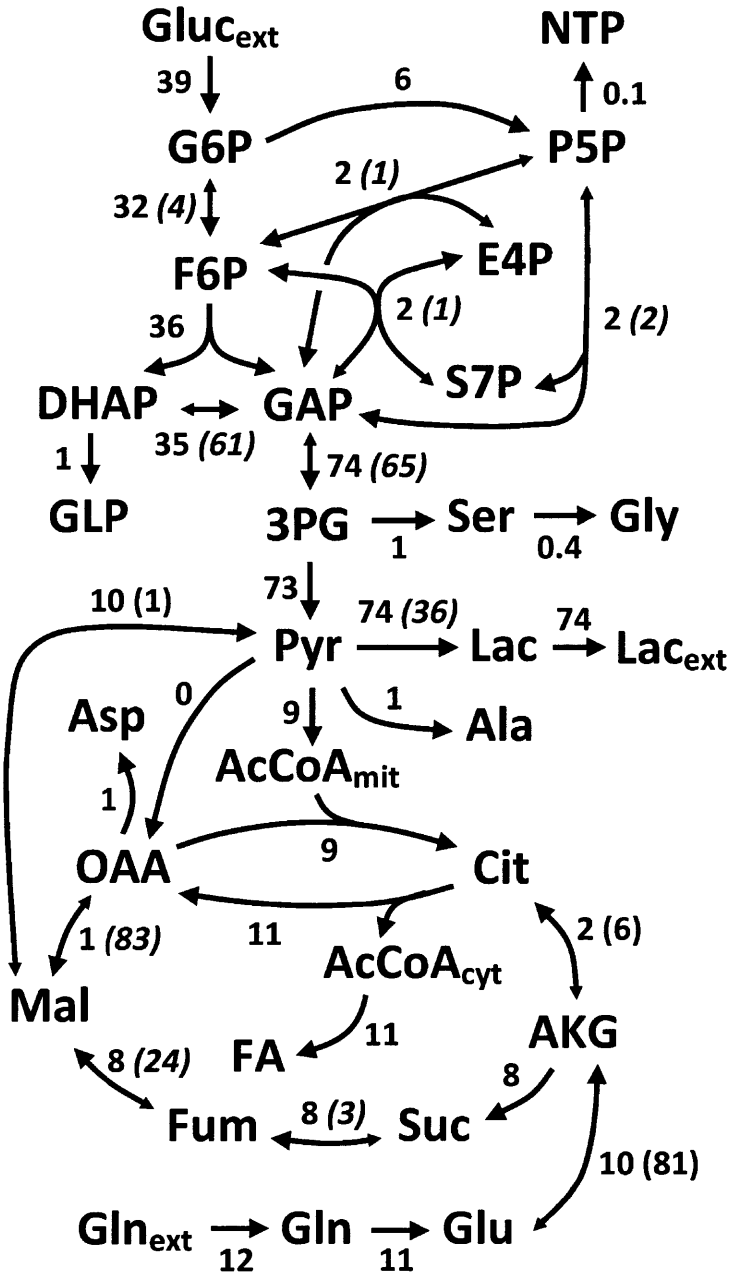

Figure 6-1: Experimentally determined fluxes representing central carbon metabolism in A549 carcinoma cells. Extracellular fluxes and MIDs were measured and incorporated into the network shown (see Tables 6.1 and 6.2). An acceptable fit was obtained with a sum of squared residuals (SSR) of 52, well under the upper bound of the 95% confidence region for a χ^2 distribution. Net fluxes are listed first for each reaction and exchange fluxes are within parentheses. Units for all fluxes are nmol min^{-1} mg protein^{-1}. Italicized numbers represent flux values that were taken from the literature since they were unidentifiable for our particular experiment.

pathway were unidentifiable (Table 6.6). TCA cycle fluxes were relatively low and largely driven by glutamine consumption. Interestingly, we observed a net flux toward citrate for the isocitrate dehydrogenase (IDH) reaction, indicating that this cell line undergoes reductive carboxylation of α-ketoglutarate (Figure 6-1). Finally, we estimated a significant flux from malate to pyruvate (malic enzyme) and a negligible flux through pyruvate carboxylase (pyruvate to oxaloacetate). Malic enzyme flux, which regenerates NADPH, can be assumed to compensate for the excessive NADPH requirements of fatty acid synthesis and any lost in the reductive carboxylation pathway [42]. Our results successfully described the metabolic phenotype of cancer cells and serve as a benchmark for our tracer analysis below.

6.8 Confidence Intervals by Tracer

We next calculated confidence intervals for every combination of tracer and flux. Results for selected fluxes are shown in Figure 6-2, and the complete set of confidence intervals over all fluxes is available in Figures 6-3 and 6-4.

Because the lower glycolytic fluxes consist primarily of stoichiometrically determined net fluxes and completely unidentifiable exchange fluxes, tracers of any kind offer no benefit in flux estimation and show little variation in confidence interval precision in this region of the network. Tracers do generate results of more diverse quality in upper glycolysis (see Figures 6-2A and B for examples). Here, glutamine tracers are completely ineffective, since there is no set of reactions by which any glutamine atom can travel to this portion of the network. Uniformly labeled glucose also gives nominal precision; because there are no other carbon sources feeding into glycolysis, all metabolites here are fully labeled by this tracer at isotopic steady state regardless of the flux distribution.

Glucose tracers labeled at some combination of the 4th, 5th, and 6th carbons give results of limited quality, primarily because these labeled atoms are trapped in a cycle. If M_i is the ith atom of metabolite M, the atomic transitions show that glucose$_i$ (where $4 \leq i \leq 6$) will distribute ^{13}C to only G6P$_i$, F6P$_i$, P5P$_{i-1}$, S7P$_{i+1}$,

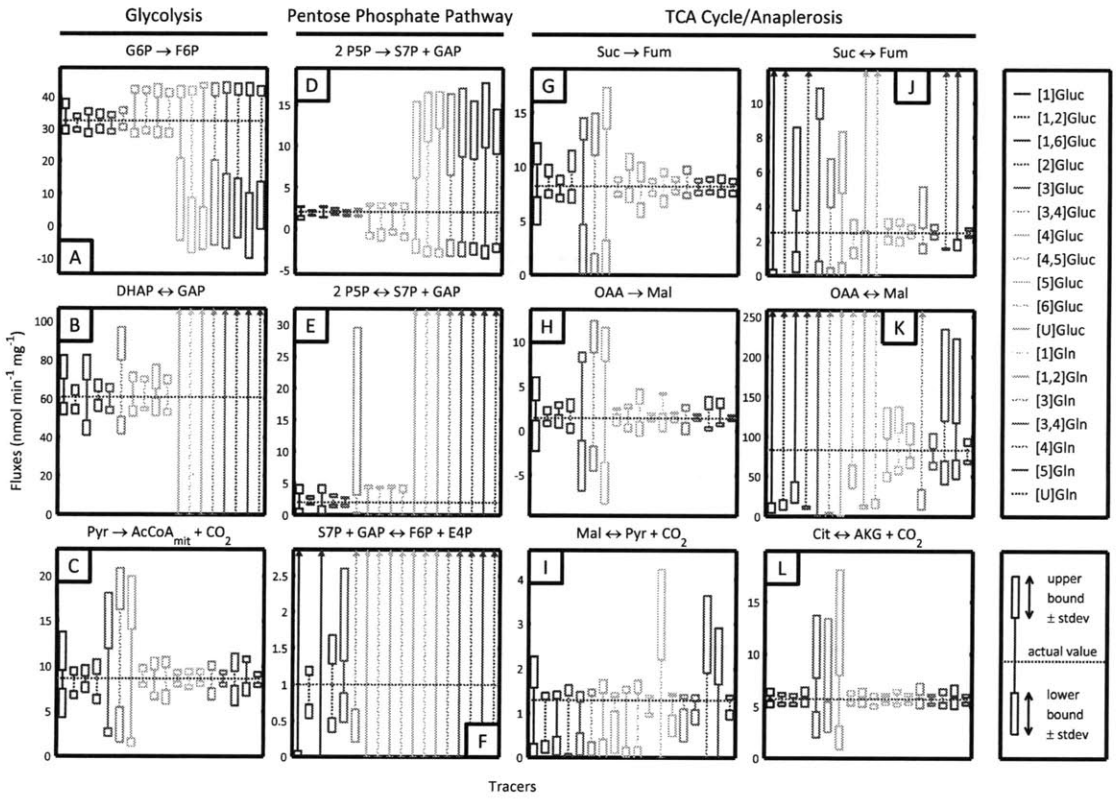

Figure 6-2: Simulated confidence intervals for selected fluxes of A549 carcinoma cell metabolism when using specific isotopic tracers. Horizontal dashed lines indicated actual fluxes. Upper and lower bounds of the 95% confidence interval are illustrated for each simulated tracer. The standard error of both upper and lower bounds is represented by the boxes at the top and bottom of each interval. (A) Glucose-6-phosphate isomerase and (B) triose-phosphate isomerase fluxes demonstrate the effectiveness of glucose tracers in estimated glycolytic fluxes. (C) Pyruvate dehydrogenase flux is most precisely estimated by most glutamine tracers and some glucose tracers. (D-F) Net and exchange fluxes within the pentose phosphate pathway are best determined with glucose tracers labeled at the 1st, 2nd, or 3rd carbon, with [1,2]Gluc performing the best. (G and H) Net fluxes and (I-L) exchange fluxes in the TCA cycle are characterized well using [U]Gluc, [1,2]Gln, [3,4]Gln or [U]Gln.

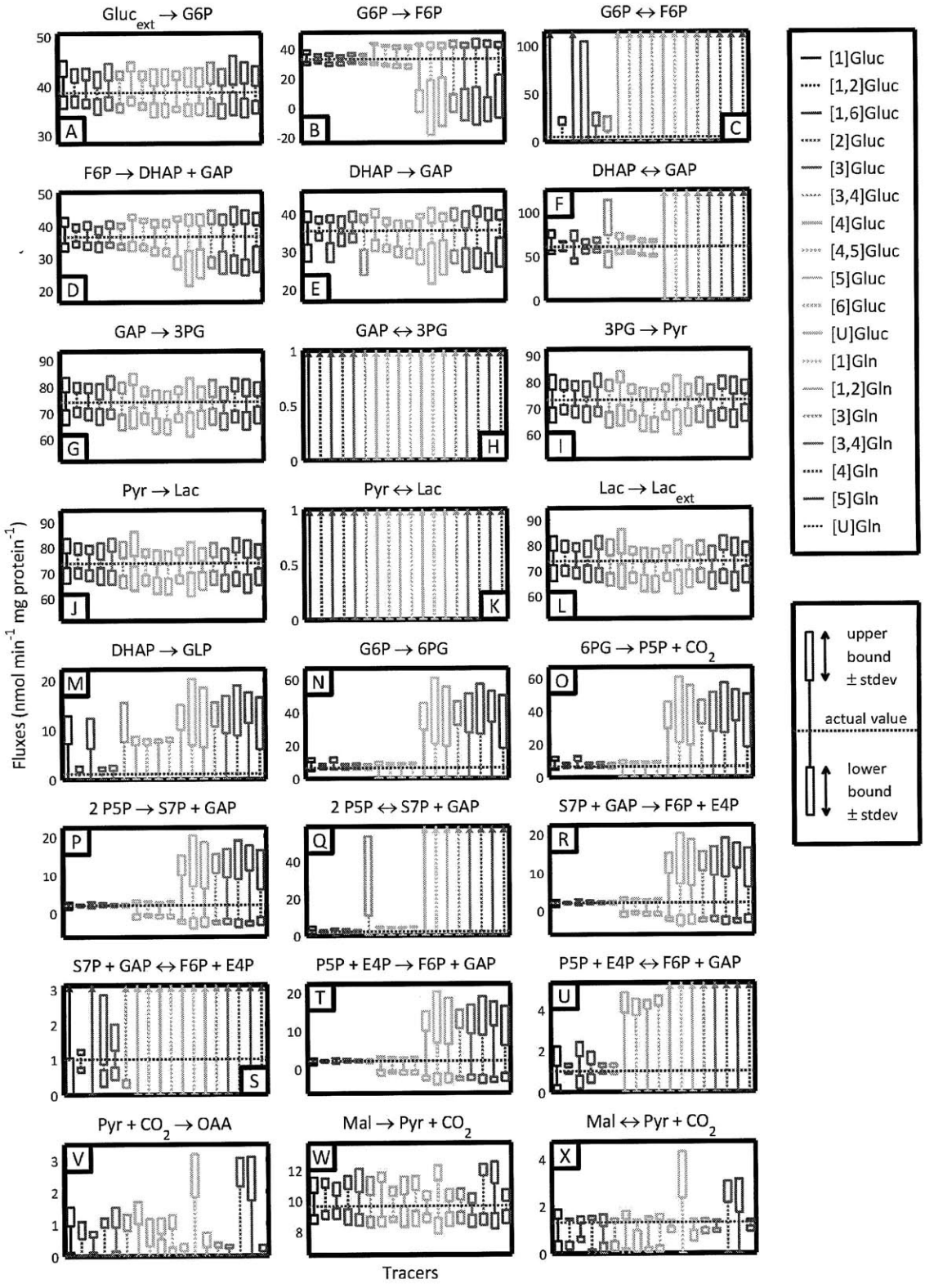

Figure 6-3: Simulated confidence intervals for fluxes #1-24 of the A549 carcinoma cell network when using various tracers.

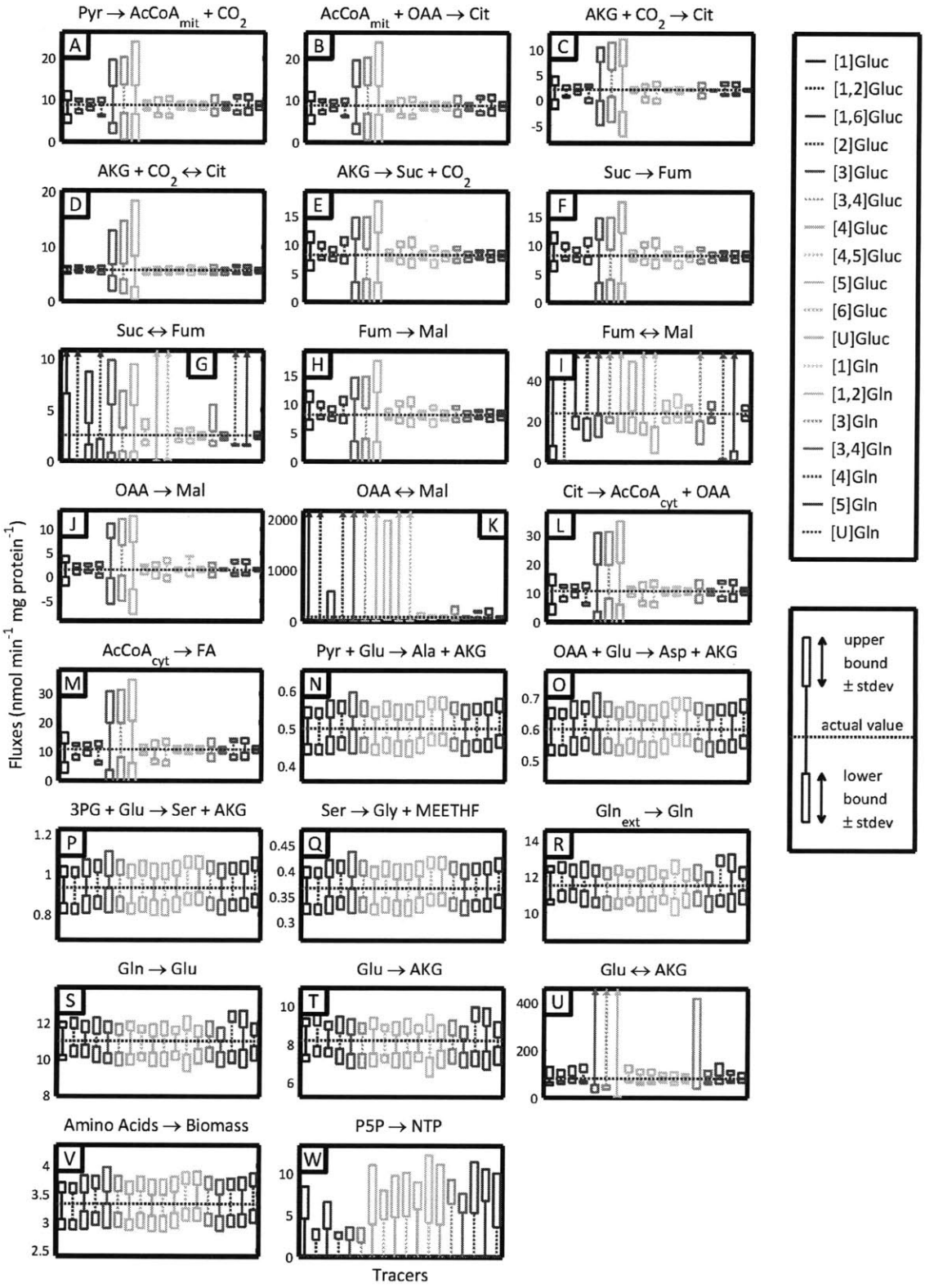

Figure 6-4: Simulated confidence intervals for fluxes #25-47 of the A549 carcinoma cell network when using various tracers.

$E4P_{i-2}$, GAP_{i-3}, and $DHAP_{i-3}$ before permanently exiting glycolysis and the PPP (see highlighted atom transitions for [4]Gluc in Figure 6-5B). This uniformity means that the key GC/MS fragment behind glycolysis flux estimation (GLP) will only be labeled at a single carbon, reducing the potential measurement diversity, which in turn reduces the sensitivity to fluxes. The 1st carbon of glucose feeds into a comparable positional carbon cycle, leading to [3-^{13}C]GLP and similarly decreased sensitivity. Tracers labeled at the 2nd or 3rd carbons ([1,2]Gluc, [2]Gluc, and [3]Gluc), however, capitalize on the atomic transitions of the subnetwork and distribute significant label to each carbon in GLP, producing the greatest sensitivity and most precise confidence intervals in glycolysis (see highlighted atom transitions for [2]Gluc in Figure 6-5A).

The estimation quality of the PPP closely reflects that of upper glycolysis. Instead of GLP, P5P is now the major contributing measurement. [U]Gluc and all glutamine tracers are again completely noninformative (for the same reasons as before). [4], [4,5], [5], and [6]Gluc are ineffective because label is once more restricted to a small subset of positions in the network. Because [1]Gluc loses its label to CO_2 in the oxidative PPP, P5P is predominantly unlabeled and demonstrates little sensitivity. Glucose labeled at the 2nd and 3rd positions are again the tracers that most confidently estimate fluxes. These behaviors are fairly consistent through all individual net and exchange fluxes of the PPP. Specific examples are shown in Figures 6-2D (transketolase net), E (transketolase exchange), and F (transaldolase exchange).

Glutamine tracers, on the other hand, generally provided better estimations of the pyruvate dehydrogenase flux, located at the junction between glycolysis and the TCA cycle (Figure 6-2C). Because both the 3rd and 4th carbons in glucose mostly transition to the 1st carbon of pyruvate and exit the system as CO_2 at this step, [3]Gluc, [3,4]Gluc, and [4]Gluc could not precisely resolve the PDH flux, despite their common use and effectiveness in previous experiments [72, 102]. This discrepancy arises because these previous studies measured labeling in CO_2 while our analysis assumes no such measurement. Importantly, one cannot exchange measurement sets of MFA experiments and necessarily expect similar results or patterns when evaluating tracers.

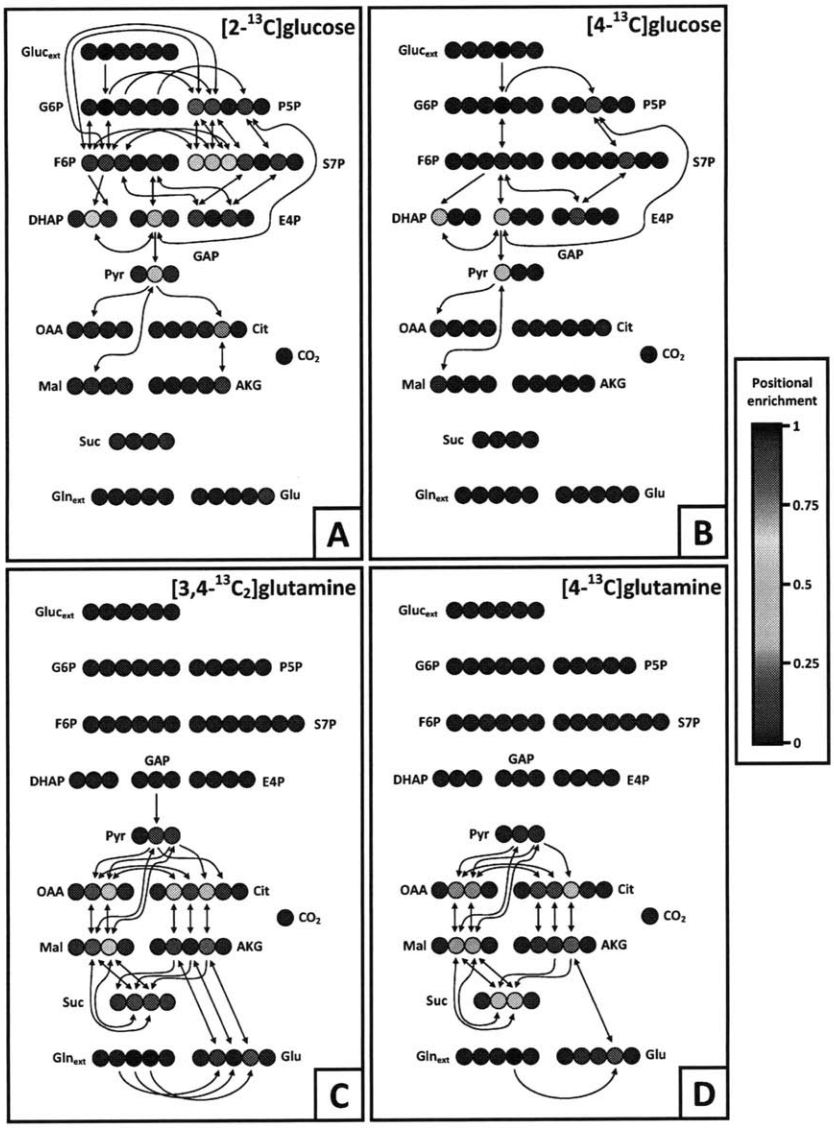

Figure 6-5: Atom transition networks and positional fractional labeling for selected glucose and glutamine tracers in A549 carcinoma cells. Fractional labeling is indicated by a color map, where dark red indicates all atoms at that position are ^{13}C and dark blue that all atoms are ^{12}C. No natural labeling was assumed in the creation of these maps. Atom transitions are indicated for all positions where fractional ^{13}C labeling exceeds 10%. (A) [2]Gluc effectively characterizes glycolytic and PPP fluxes because DHAP (and by extension GLP) is labeled in multiple positions by different combinations of fluxes, leading to greater measurement sensitivity. (B) [4]Gluc poorly identifies fluxes in glycolysis and the PPP because its sole labeled carbon is caught in a cycle and can only reach the 1st carbon of DHAP. (C and D) [3,4]Gln and [4]Gln are both able to label a majority of the carbon atoms in the TCA cycle; however, the two labeled atoms in the former lead to larger, clearer measurements and therefore more precise fluxes.

While ^{13}C glutamine and uniformly labeled glucose tracers offered minimal information for glycolysis and the PPP, they demonstrated considerable utility in estimating TCA cycle and anaplerotic fluxes. Net fluxes within the TCA cycle (succinate to fumarate and oxaloacetate to fumarate) were best characterized when using glutamine tracers with two or more carbons, specifically [1,2]Gln, [3,4]Gln, and [U]Gln (Figures 6-2G and H). Because the glutamine flux into the TCA cycle is significantly smaller than the incoming glucose flux, multiply labeled glutamine tracers are presumably more useful because they introduce greater amounts label that are less easily diluted (see Figures 6-5C and D to compare atom transitions for [3,4]Gln and [4]Gln). The effectiveness of these tracers was further highlighted by the improved confidence intervals obtained for exchange fluxes in the TCA cycle and malic enzyme reactions (Figures 6-2I through L). Exchange between succinate, fumarate, and oxaloacetate are key reactions within oxidative metabolism that also affect cofactor levels and pyruvate cycling; therefore, precise estimation of these reactions are of paramount importance for cancer research [42, 67].

To better describe the quality of data and simulations obtained from specific tracers, we calculated precision scores for each tracer, covering both subnetworks and central carbon metabolism in its entirety. Independent fluxes included in the scoring for each subnetwork are described in Figure 6-6. All glucose tracers except [U]Gluc scored well for glycolysis, with [1,2]Gluc, [2]Gluc, and [3]Gluc performing significantly better than most (Figure 6-7A). These three tracers provided the best scores (i.e., the most precise estimates) for the PPP as well (Figure 6-7B). The highest scoring glucose tracer for TCA cycle analysis was [U]Gluc (Figure 6-7C). Three glutamine tracers also generated similarly high scores and precise estimates within the TCA cycle subnetwork; in particular, those tracers labeled at two or more positions scored best ([1,2]Gln, [3,4]Gln, and [U]Gln). Finally, the best overall tracer for analyzing the entire cancer cell flux network was [1,2]Gluc, followed by other glucose tracers labeled at the 2nd or 3rd carbons (Figure 6-7D). Although these tracers did not generate the best results for the TCA cycle, their unique ability to consistently characterize fluxes throughout the entire network resulted in superior scores compared to other tracers.

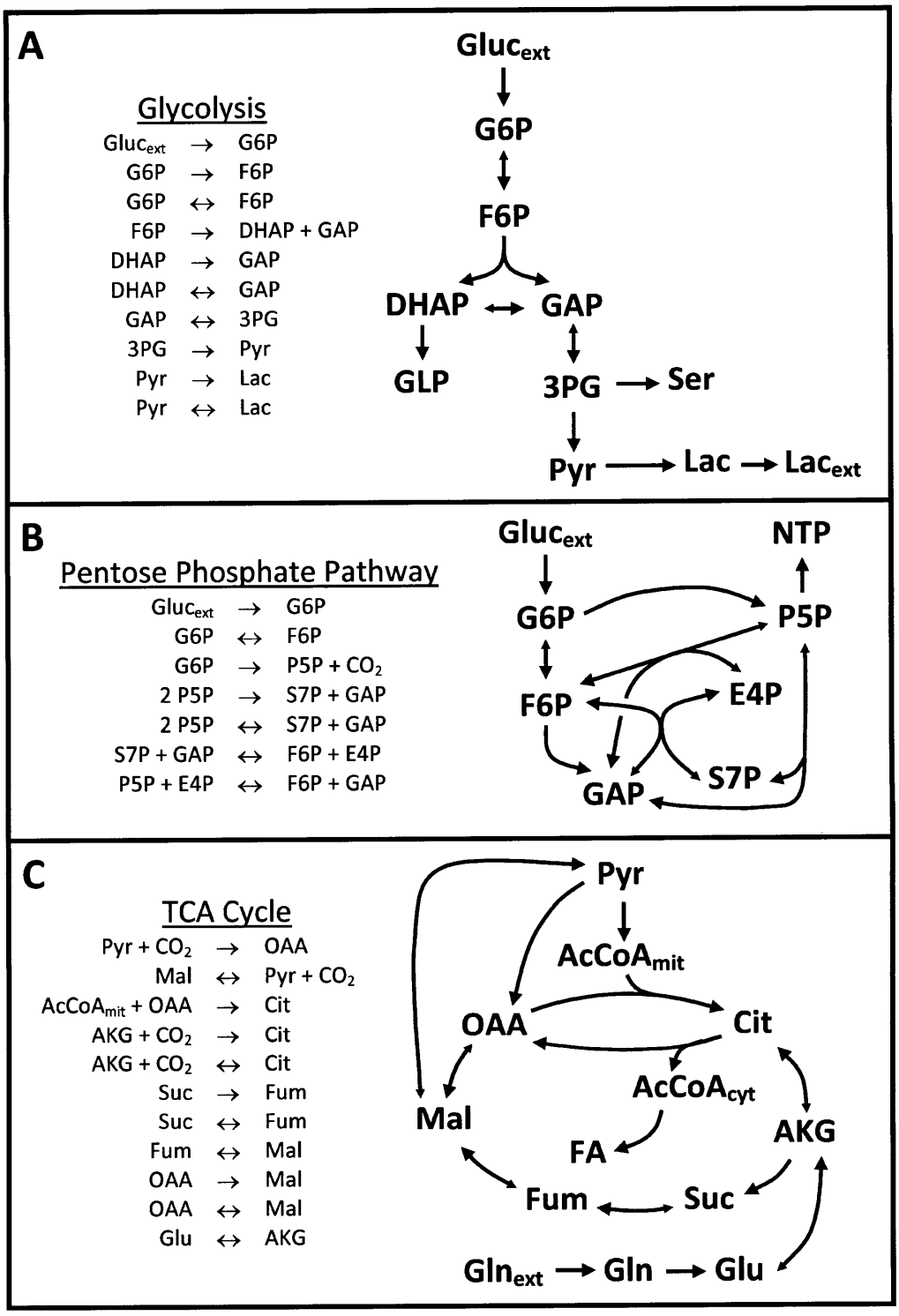

Figure 6-6: The subnetworks used to calculate precision scores for different sections of the A549 carcinoma network. The independent fluxes of the subnetworks used in the analysis are listed.

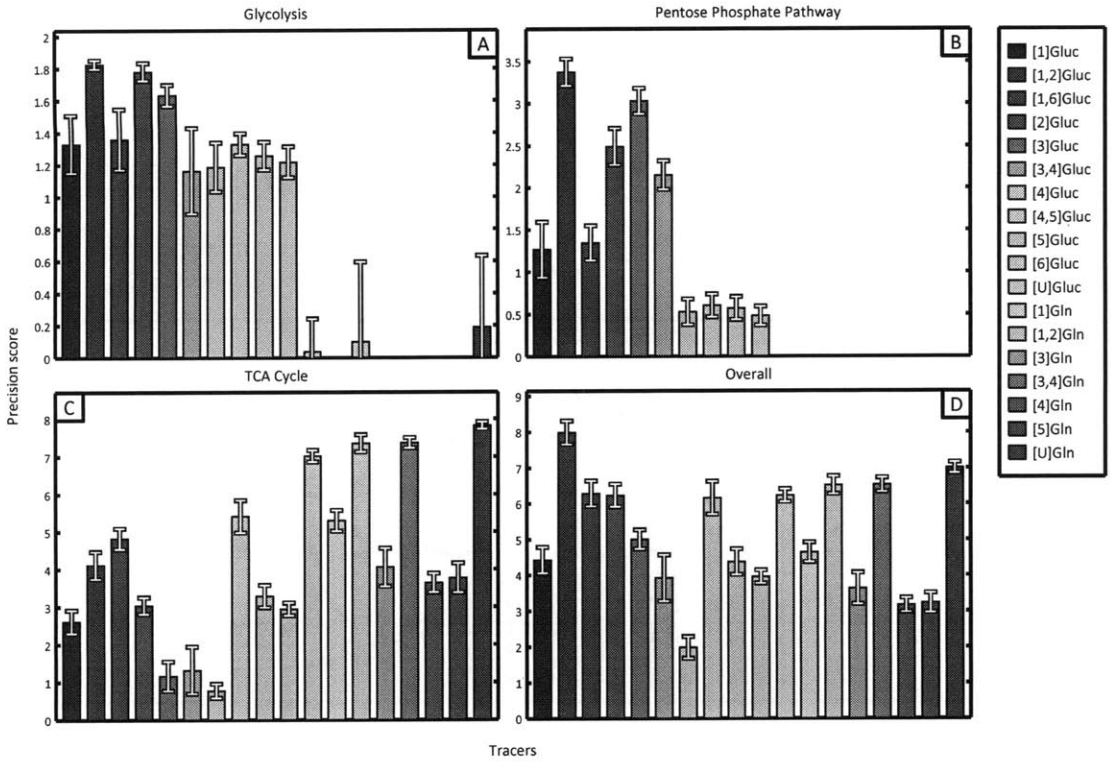

Figure 6-7: Results obtained from our precision scoring algorithm identifying the optimal tracer for the analysis of subnetworks and central carbon metabolism in A549 carcinoma cells. The precision scores resulting from simulated experiments involving only natural labeling have been subtracted from each displayed tracer score to aid in visual differentiation and comparison. (A) Glycolysis and (B) pentose phosphate subnetworks are best described by [1,2]Gluc, [2]Gluc, and [3]Gluc. (C) TCA cycle scores were highest for [U]Gluc and several glutamine tracers labeled at two or more carbons. (D) The most precise tracer for the analysis of the entire network was [1,2]Gluc.

6.9 Nonstationary Confidence Intervals

Recently, several groups have demonstrated nonstationary MFA using transient labeling measurements [106, 126, 131, 164, 165]. To observe the impact of tracer choice on the precision of nonstationary flux experiments, we simulated measurements from selected isotopic tracers at three time points (5, 10, and 15 min) and estimated fluxes and confidence intervals throughout the network. Confidence intervals derived from selected tracers are listed in Figure 6-8, and associated precision scores are presented in Figure 6-9. Relative precision from one tracer to the next was fairly consistent through almost all fluxes, and the most effective tracers for stationary MFA are also the most effictive in transient labeling experiments. One exception was the performance of [U]Gluc for estimating glycolytic fluxes. Uniformly labeled glucose generated significantly better results in our nonstationary analysis compared to our previous steady state simulations; however, tracers such as [1,2]Gluc continued to outperform this more commonly used substrate. Because the stationary and nonstationary measurements produced comparable confidence intervals, steady-state flux analysis may be the most cost- and time-effective method in mammalian cell cultures, since the transient experiments essentially require parallel experiments for each time point analyzed.

6.10 Discussion

To better demonstrate the impact of tracer selection upon flux estimation quality, we have quantitatively and comparatively described the precision of uniquely labeled ^{13}C glucose and glutamine tracers for flux determination in a mammalian cell tumor line. This area of research is particularly timely as the metabolic phenotype of tumors has reemerged as an important area of study and potential clinical target [83]. Fueled by advanced computational software and technologies, researchers can now characterize cellular metabolism in unprecedented detail. To better demonstrate the utility of isotopic tracers for cancer research, we have quantitatively and comparatively described

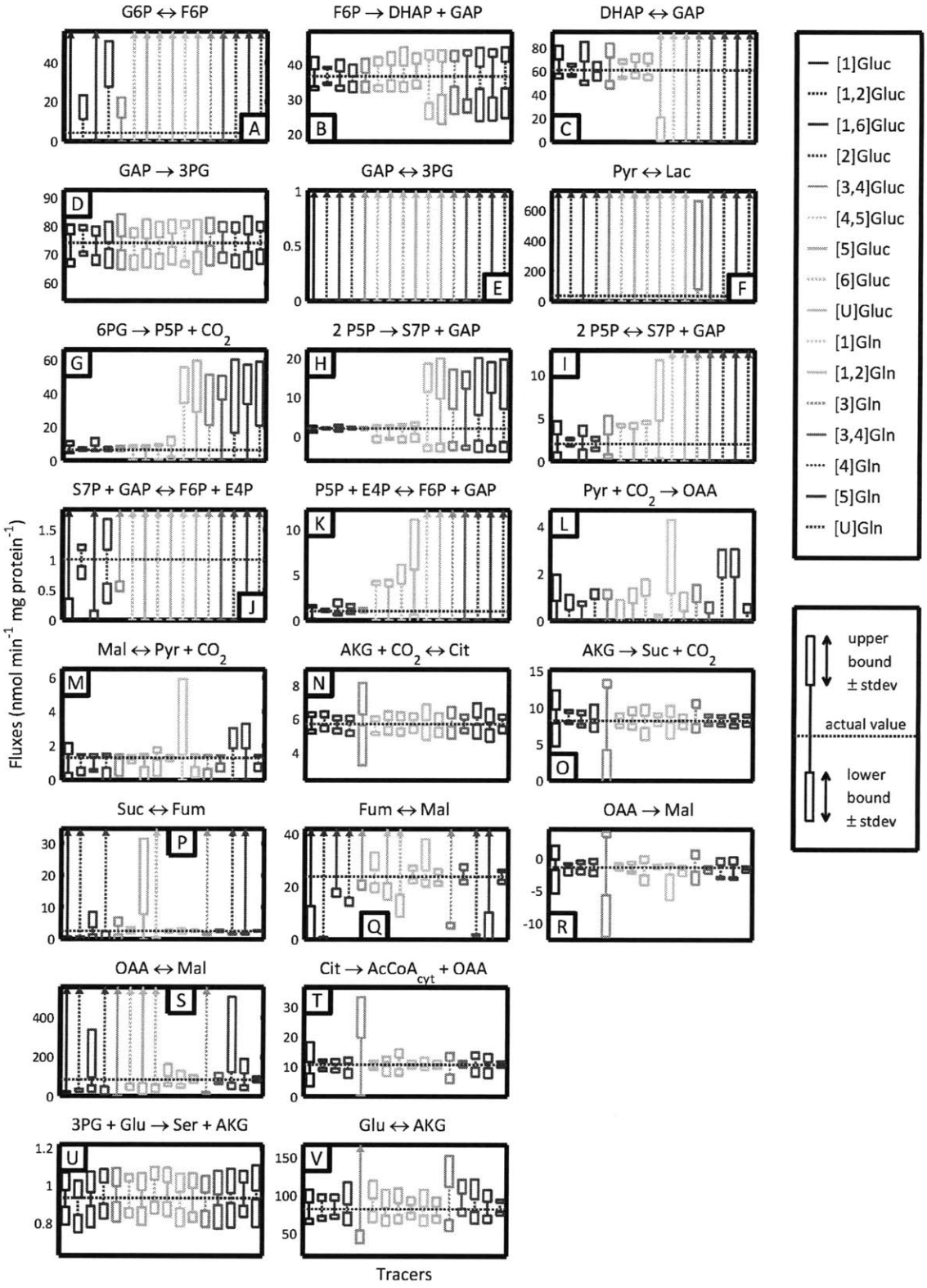

Figure 6-8: Simulated confidence intervals for all independent fluxes of the A549 carcinoma cell network when using various tracers and isotopically nonstationary measurements.

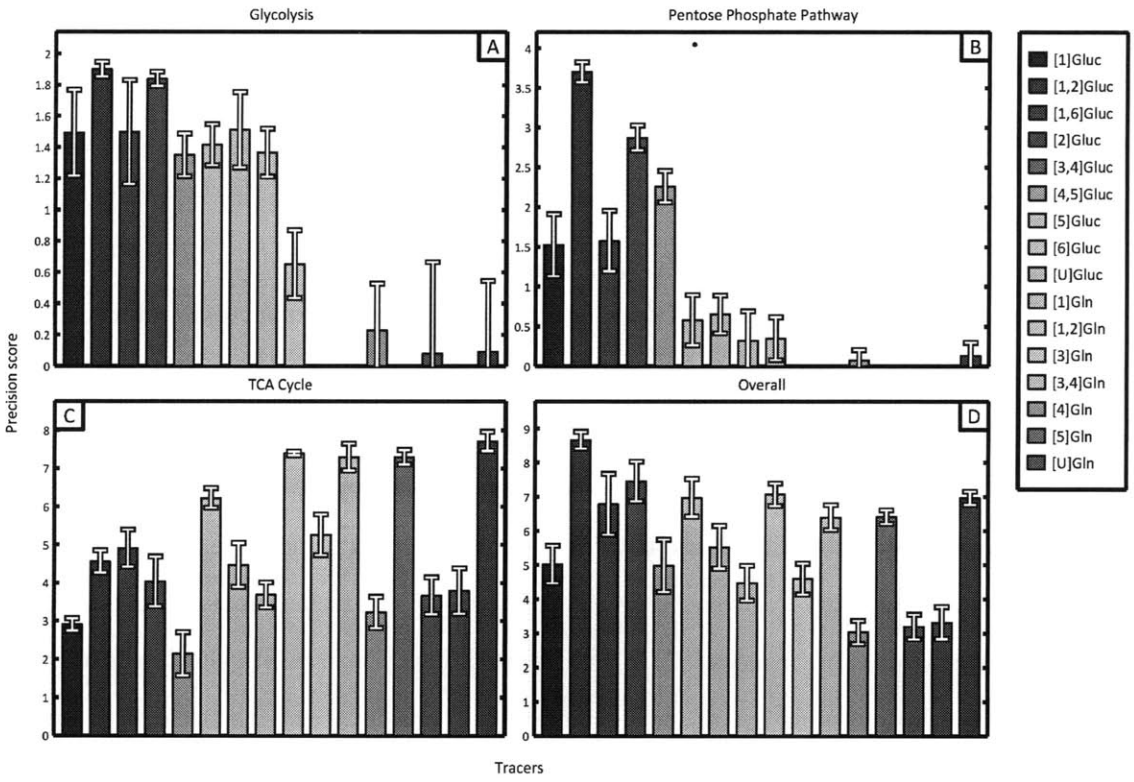

Figure 6-9: Precision scores for glycolysis, the pentose phosphate pathway, the TCA cycle, and the overall network when using various tracers and isotopically nonstationary measurements in A549 carcinoma cells.

the precision of uniquely labeled ^{13}C glucose and glutamine tracers for flux determination in mammalian cells. An EMU-based algorithm enabled the high-throughput flux estimations and confidence interval calculations required for this description. Table 6.7 summarizes the optimal tracers to use for estimating each flux, specific subnetworks, and central carbon metabolism as a whole. [1,2]Gluc provided the highest level of precision for the overall network, glycolysis, and the PPP, while [U]Gluc and multiply labeled glutamine tracers were most informative for the TCA cycle. Specific instances of their use are present in the literature [23, 102, 161]; however, we hope the demonstrable improvements of precision described here help to propagate the use of more effective tracer molecules in future MFA experiments.

The tracer evaluation process and accompanying results in this study provide a general procedure for the experimental design of isotopic tracer studies. However, it should be noted that specific cell types exhibit particular flux profiles, which in turn will affect the results generated by simulated tracer studies such as those presented here. As such, our findings are network dependent and therefore most relevant to the study of tumor cells. Furthermore, the precision of our flux estimations are dependent upon the metabolites measured. As mass spectrometry and other related technologies continue to improve, researchers will obtain richer data sets to incorporate into the estimation process [92, 102]. Nevertheless, our results should be valid for most mammalian systems given the conserved nature of atom transitions in central carbon metabolism. In cells with significantly different networks and/or flux distributions (e.g., gluconeogenic hepatocytes) these specific tracer simulations may not apply [160]. However, our methodology will still prove useful in optimizing experimental design, especially in complex systems where the best tracer cannot be determined *a priori*. These results can serve as a guide to more effectively design experiments that use flux analysis to study cellular metabolism and physiology.

Glycolysis	Best tracer(s)
$Gluc_{ext} \rightarrow G6P$	[1,2] and [2]Gluc
$G6P \rightarrow F6P$	[1,2]Gluc
$G6P \leftrightarrow F6P$	[3] and [3,4]Gluc
$F6P \rightarrow DHAP + GAP$	[1,2] and [2]Gluc
$DHAP \rightarrow GAP$	[1,2]Gluc
$DHAP \leftrightarrow GAP$	[1,2]Gluc
$GAP \rightarrow 3PG$	[1,2] and [2]Gluc
$GAP \leftrightarrow 3PG$	None
$3PG \rightarrow Pyr$	[1,2]Gluc
$Pyr \rightarrow Lac$	[1,2] and [2]Gluc
$Pyr \leftrightarrow Lac$	None
$Lac \rightarrow Lac_{ext}$	[1,2] and [2]Gluc

Pentose phosphate pathway	Best tracer(s)
$G6P \rightarrow P5P + CO_2$	[1,2]Gluc
$2\ P5P \rightarrow S7P + GAP$	[1,2] and [3]Gluc
$2\ P5P \leftrightarrow S7P + GAP$	[1,2]Gluc
$S7P + GAP \rightarrow F6P + E4P$	[1,2] and [3]Gluc
$S7P + GAP \leftrightarrow F6P + E4P$	[1,2]Gluc
$P5P + E4P \rightarrow F6P + GAP$	[1,2] and [3]Gluc
$P5P + E4P \leftrightarrow F6P + GAP$	[1,2] and [3]Gluc

TCA cycle/anaplerosis	Best tracer(s)
$Pyr \rightarrow AcCoA_{mit} + CO_2$	[U]Gln
$OAA + AcCoA_{mit} \rightarrow Cit$	[U]Gln
$AKG + CO_2 \rightarrow Cit$ (abcdef)	[U], [1,2] and [3,4]Gln
$AKG + CO_2 \leftrightarrow Cit$ (abcdef)	[U]Gln
$AKG \rightarrow Suc + CO_2$	[U] and [3,4]Gln
$Suc \rightarrow Fum$	[U] and [3,4]Gln
$Suc \leftrightarrow Fum$	[U] and [1,2]Gln
$Fum \rightarrow Mal$	[U] and [3,4]Gln
$Fum \leftrightarrow Mal$	[U]Gln
$OAA \rightarrow Mal$	[U]Gln
$OAA \leftrightarrow Mal$	[U]Gln
$Pyr + CO_2 \rightarrow OAA$	[U]Gln
$Mal \rightarrow Pyr + CO_2$	[U] and [3,4]Gln
$Mal \leftrightarrow Pyr + CO_2$	[U]Gln

Amino acid metabolism	Best tracer(s)
$Gln_{ext} \rightarrow Gln$	[3], [U] and [3,4]Gln
$Gln \rightarrow Glu$	[3], [U] and [3,4]Gln
$Glu \rightarrow AKG$	[3], [U] and [3,4]Gln
$Glu \leftrightarrow AKG$	[U]Gln

Table 6.7: Optimal tracers for net ($\rightarrow$) and exchange ($\leftrightarrow$) fluxes of the A549 carcinoma cell.

Chapter 7

A Genetic Algorithm for Tracer Optimization

7.1 Introduction

Interest in trustworthy, reliable flux analysis in mammalian cells continues to increase, both for the identification of disease mechanisms [12, 28, 102], and for the study of metabolic defects that have been implicated in cancer, mental disorders, diabetes, and related syndromes [1, 19, 35, 132]. ^{13}C metabolic flux analysis (MFA) enables researchers to quantify intracellular fluxes *in vivo* via the combined use of stable isotopic tracers, analytical methods such as NMR and mass spectrometry, and computational tools [68, 122]. While glucose tracers are effectively used to label single substrate microbial cultures and highly oxidative cells or tissues, many mammalian systems are highly compartmentalized and grow on complex media, metabolizing various substrates. For example, tumor cells divert most glucose carbon to lactate and use amino acids (e.g. glutamine) or fatty acids to contribute carbon to the tricarboxylic acid (TCA) cycle [41, 48], and hepatocytes undergoing gluconeogenesis utilize lactate, alanine, amino acids, and glycerol to produce glucose through various pathways [116, 160]. As a result, individual tracers are often only effective for characterizing individual pathways such as the TCA cycle or pentose phosphate pathway.

We discussed previously (in Chapter 6) how confidence interval calculations can

be used to select an optimal, individual tracer for a given metabolic network and flux distribution. To maximize the information obtained from a given experiment, an alternative approach is to apply multiple tracers simultaneously. This ability is particularly advantageous in situations where sample size is limited, as in the testing of clinical materials [24, 69]. However, the exact combination of tracers must be chosen with care, as information may be lost if different tracers generate the same labeling pattern or mass isotopomer distribution (MID). This problem cannot be addressed experimentally given the high cost of uniquely labeled tracers and the large search space of potential mixtures. Also, both flux estimation and sensitivity analysis must be performed to ensure that the experimental significance of each flux is maintained.

Here we employed a genetic algorithm to generate mixtures of uniquely labeled ^{13}C glucose and glutamine. At each generation of the evolution an elementary metabolite unit (EMU)-based flux analysis method was used to perform flux estimation and calculation of confidence intervals. Tracer combinations were selected to advance via tournament selection. By coupling our tracer evaluation algorithm to an evolutionary algorithm we identified optimal tracer mixtures that provide the most statistically significant flux values in a non-small cell lung carcinoma cell line. To gauge the robustness of one tracer set ([1,2-$^{13}C_2$]glucose + [U-$^{13}C_5$]glutamine) we conducted simulated experiments in which the flux values were perturbed significantly from that of the original network. Finally, we validated the optimized tracer combination experimentally in tumor cells, demonstrating improvements over two other commonly used tracer sets.

7.2 Genetic Algorithm

We employed a genetic algorithm to search the space of possible tracer mixtures [88, 155]. (The algorithm was implemented in Matlab and the code can be found in Appendix D.) In order to apply our evolutionary approach, we need to define (1) a phenotypic search space, (2) a method of encoding and decoding these phenotypes to and from chromosomes, and (3) a measure of each phenotype's fitness.

We define phenotype as a vector of tracer composition fractions,

$$\mathbf{c} = (f_{1,1}, f_{1,2}, \ldots, f_{1,N_1}, \ldots, f_{i,1}, f_{i,2}, \ldots, f_{i,N_i}, \ldots, f_{M,1}, f_{M,2}, \ldots, f_{M,N_M}) \tag{7.1}$$

where $\mathbf{c}$ is the tracer composition vector and $f_{i,j}$ is the jth isotopomer fraction of ith substrate. There are M total substrates and each substrate has N_i total isotopomers. Phenotypes can be converted into chromosomes using the expression

$$h = \sum_k 2^{b(k-1)} \lfloor 2^b \cdot c_k \rfloor \tag{7.2}$$

where c_k are the elements of the tracer composition vector, b is the number of desired chromosome bits per tracer fraction, and h is in the decimal form of the chromosome. (To apply crossover and mutation events, h is converted to its binary representation to form a bit string.) Chromosomes can be decoded back to tracer fractions as follows,

$$c_k = \frac{h \bmod 2^{bk} - h \bmod 2^{b(k-1)}}{2^{b(k-1)}} \Big/ \sum_{l \epsilon S_k} \frac{h \bmod 2^{bl} - h \bmod 2^{b(l-1)}}{2^{b(l-1)}} \tag{7.3}$$

where S_k is the set of indices of all fractions in $\mathbf{c}$ representing the same substrate as c_k so that the denominator of Equation 7.3 will renormalize isotopomer fractions in the event that their overall sum was altered during crossover and mutation. Fitness is expressed by a tracer mixture's precision score, which is inversely related to the magnitude of the parameter confidence intervals calculated from simulated MFA experiments [98].

Figure 7-1 describes the evolutionary process. First, a set of chromosomes are randomly generated. We designated 8 bits per isotopomer, allowing a resolution within less than 0.5% for each fraction. In creating the initial population, the on/off probability for each bit can be adjusted to control the initial average number of tracers participating in any given mixture. For all of the initial populations used in this study, an initial population of size 500 and an on/off probability of 10% was found to give realistic and practical results.

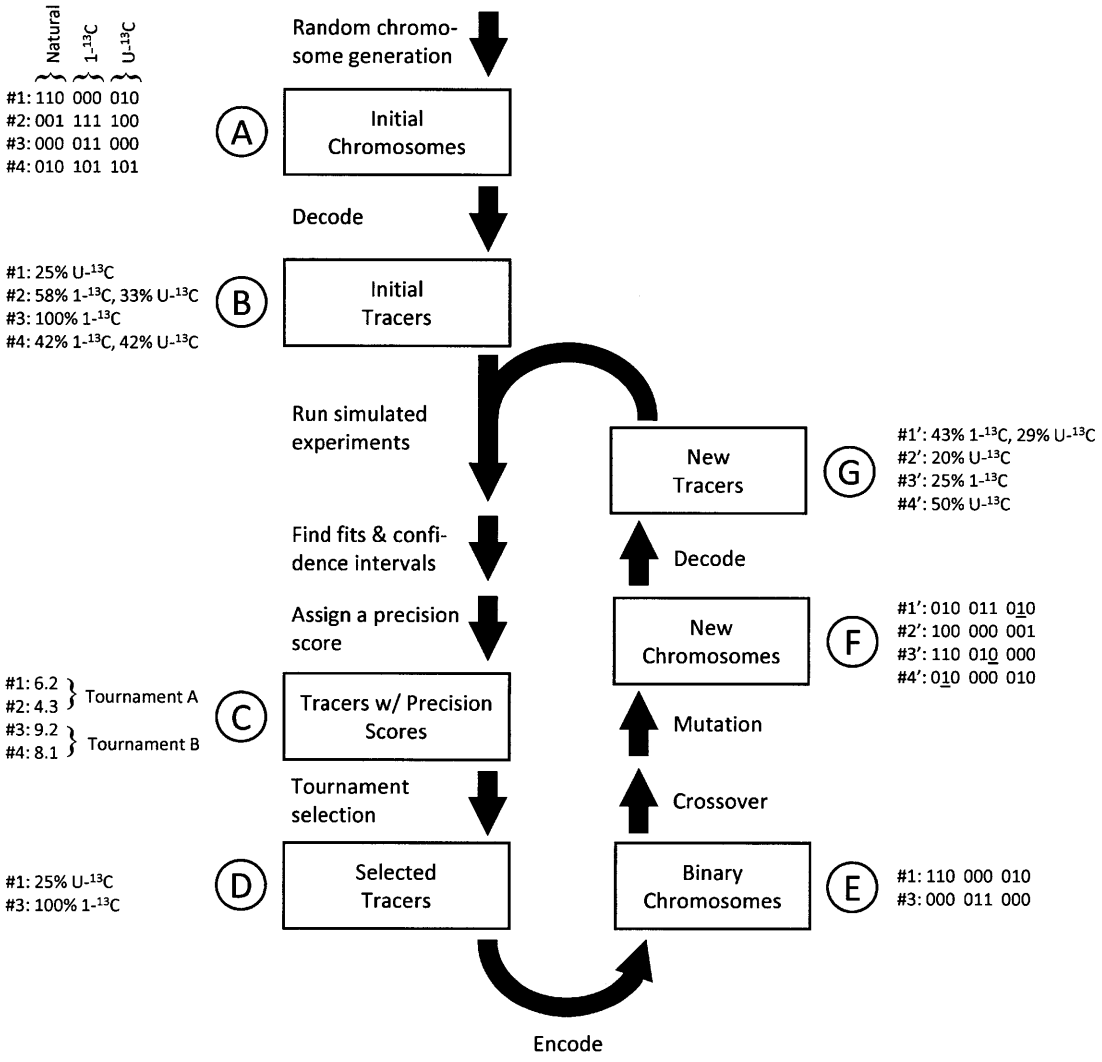

Figure 7-1: Tracer optimization by means of a genetic algorithm. The central flowchart illustrates the algorithms steps. A simple example is also provided where three isotopic forms of the tracer (naturally labeled, 1st-carbon labeled, and uniformly labeled) are considered. First, a random set of chromosomes is decoded to produce an initial population of tracer mixtures (A,B). Our selection criterion, the precision score, is generated for each mixture from confidence intervals calculated via simulated flux analysis experiments (C). Tracer mixtures are chosen for the next evolutionary round by tournament selection (D). The mixtures are encoded back to chromosomes (E), after which mutation and crossover events are applied to the selected to generate a new population (F). These new chromosomes are decoded into tracer mixtures and can participate in a new round of evolution if desired (G).

The initial chromosomes are next decoded using Equation 7.3 to give tracer composition vectors. Simulated flux analysis experiments are then conducted for each vector, producing parameter confidence intervals from which precision scores can be calculated. Tournament selection is then applied to choose successful phenotypes for recombination while still retaining some diversity [61]. The surviving phenotypes are then encoded back to chromosomes and are modified by two-point crossover and point mutation events to form a new population (which can serve as a starting point for another iteration of the evolutionary process). For each of the evolutionary rounds in this study, a tournament size of four was used to select 25% of the current population for survival. The survivors were then mated using two-point crossover to create a completely new population equal in size to the previos population.

7.3 Cell Culture and Metabolite Extraction

The A549 lung carcinoma cell line (ATCC) was maintained in high-glucose DMEM supplemented with 4 mM glutamine, 10% FBS, and 100 U mL^{-1} penicillin/streptomycin (Invitrogen). Prior to labeling, cells were maintained in custom DMEM lacking amino acids, glucose, and pyruvate (Hyclone) supplemented with 25 mM glucose, 4 mM glutamine, 1X MEM essential amino acids (Invitrogen), 10% FBS, and antibiotics for one passage. For flux experiments semi-confluent cells were cultured with [1,2-$^{13}C_2$]glucose + [U-$^{13}C_5$]glutamine, a 1:1 mixture of [U-$^{13}C_6$]glucose and [1-^{13}C]glucose, or [U-$^{13}C_5$]glutamine alone (Cambridge Isotope Laboratories) for 24 hours to achieve isotopic steady state. Spent medium was collected and analyzed for glucose, lactate, glutamine, and glutamate on a YSI 7100 system (YSI Life Sciences). Extracellular fluxes were calculated from these measurements and cell counts on parallel plates. Cells were quenched with 400 µL of −20°C methanol, an equal volume of water was added, and cells were collected via scraping. Two volumes of chloroform were added, and the cells were vortexed at 4°C for 30 minutes. Samples were centrifuged at 10,000 *g* for 10 minutes, and the aqueous phase was collected in a new tube for evaporation under airflow.

7.4 Derivatization and GC/MS Measurements

Dried polar metabolites were dissolved in 50 µL of 2% methoxyamine hydrochloride in pyridine (Pierce) and held at 37°C for 2 hours. After reaction the solution was split equally to two tubes and 45 µL of either MBTSTFA + 1% TBDMCS or MSTFA + 1% TMCS (Pierce) was added. Samples were incubated for 60 minutes at 55°C or 37°C, respectively. Gas chromatography/mass spectrometry (GC/MS) analysis was performed using an Agilent 6890 GC equipped with a 30m DB-35MS capillary column connected to an Agilent 5975B MS operating under electron impact (EI) ionization at 70 eV. One µL of sample was injected in splitless mode at 270°C, using helium as the carrier gas at a flow rate of 1 mL min^{-1}. The GC oven temperature was held at 100°C for 3 min and increased to 300°C at 3.5°C min^{-1} for a total run time of approximately 60 min. The MS source and quadrupole were held at 230°C and 150°C, respectively. The detector was operated in scanning mode for TBDMS derivatizations and in selected ion monitoring (SIM) mode for metabolites reacted with TMS. MIDs were obtained for each measured metabolite and incorporated with extracellular flux measurements for flux determination. The identity and values of these measured fragments and fluxes are listed in Table 6.5. (Note that these measurements are the same as those assumed for the simulated carcinoma metabolism in Chapter 6.)

7.5 Tracer Optimization

We applied our tracer optimization algorithm to the A549 carcinoma metabolism previously studied in Chapter 6. The major features of this flux network are a high glycolytic flux (from gluxose) that is almost completed excreted as lactate and a TCA cycle flux that is driven primarily by the extracellular glutamine uptake flux. See Figure 6-1 for the network and Table 6.3 for the atom transitions.

When considering potential components for our tracer mixtures, we returned to the same list of tracers that were evaluated independently in Chapter 6 (see Table 6.4). To this list we also added the naturally labeled versions of glucose and glutamine

to bring the total number of potential participants to 20. A set of 500 randomized tracer mixtures was created as an initial population and subjected to ten rounds of evolution (Figure 7-2). The mean precision score of the population progressed from 11.3 to 13.9 and the maximum score from 13.3 to 14.3. Hierarchical clustering revealed two primary motifs in the final selected population of tracer mixtures (Figure 7-3). One group of mixtures chiefly used [1,2]Gluc as the glucose tracer, while the other group used a combination of [3]Gluc and [3,4]Gluc. All selected mixtures relied almost exclusively upon [U]Gln as the glutamine tracer.

We simplified these two selected motifs into one mixture of 100% [1,2]Gluc and 100% [U]Gln and another 50% [3]Gluc, 50% [3,4]Gluc, and 100% [U]Gln. (With regards to precision score, these simplifed versions perform comparably to the original selected tracers.) Although we originally only calculated one lumped precision score for each tracer mixture, we can now break that down into individual precision scores for each flux. To demonstrate the strengths and weaknesses of the selected tracers, we also calculated flux-by-flux precision scores for three traditional stand-alone isotopic tracers ([1]Gluc, [U]Gluc, and [U]Gln) as well as one traditional and simple tracer mixture (a 1:1 mixture of [1]Gluc and [U]Gluc). A heat map of precision-score ratios for each pairing of selected tracer mixture and traditional tracer for each flux is shown in Figure 7-4. Confidence intervals for the tracer mixtures and traditional tracers are also shown for a selected set of fluxes in Figure 7-5.

7.6 Precision Score Sensitivity

One potential stumbling block to the reliability of precision scoring is the sensitivity of scores with respect to the originally assumed flux distribution. If confidence intervals and precision scores vary widely with only small changes in fluxes, then our new method of tracer evaluation and optimization would have little use since the originally assumed flux distribution will most frequently only be an approximation. To probe this issue, we simulated precision scores for [1,2]Gluc/[U]Gln for a large number of flux distributions varied randomly around the original to see how quickly the precision

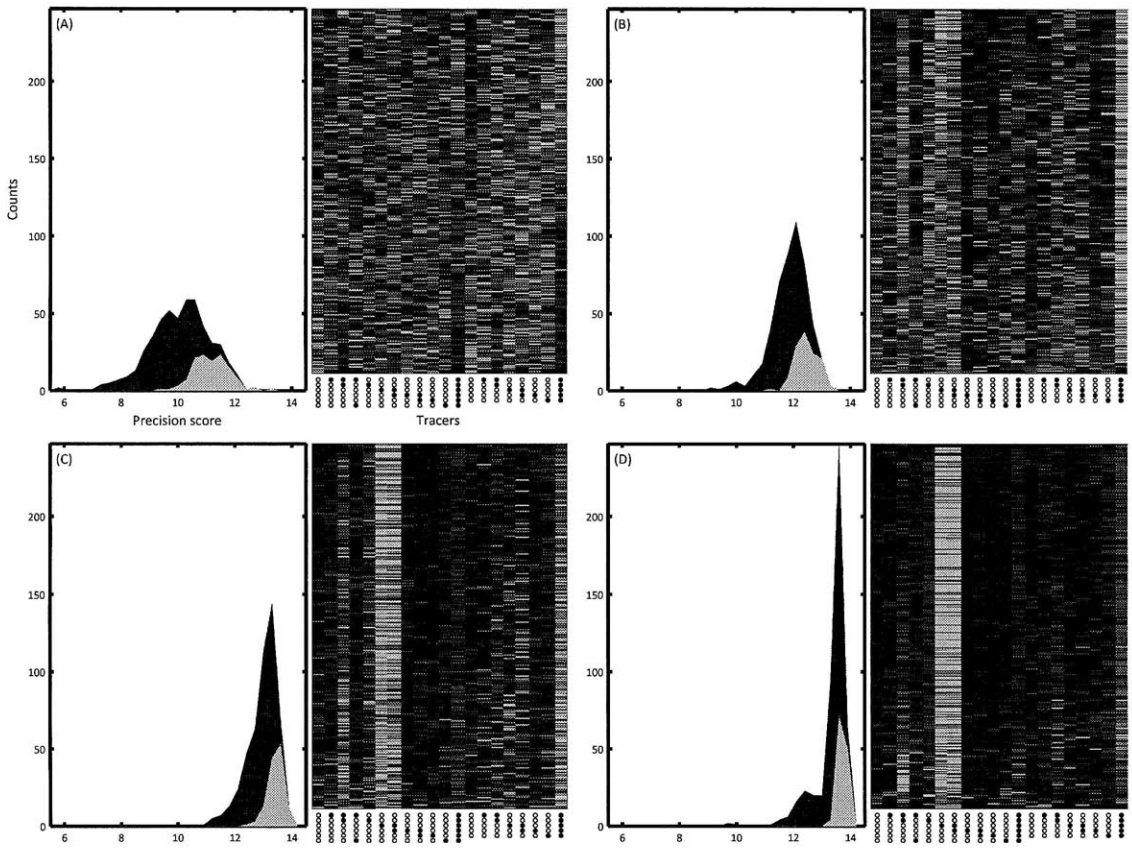

Figure 7-2: Selected tracer mixtures after rounds 1,4,7, and 10 of tracer optimization are shown in subfigures (A), (B), (C), and (D), respectively. The left subplot of each subfigure shows the general (dark blue) and selected (cyan) tracer mixture populations distributed by precision score. The right subplot shows the tracer fractions of each mixture (sorted with highest scoring mixtures at the top), where columns correspond to tracer fractions and rows to tracer mixtures. A dark red slice corresponds to a tracer fraction of 100%, while dark blue corresponds to zero. By the final round of evolution, we see that the distribution of precision scores shift upwards by about two units and the initial randomly distributed set of tracer mixtures has been narrowed down to two different tracer mixtures: [1,2]Gluc/[U]Gln and [3]Gluc/[3,4]Gluc/[U]Gln.

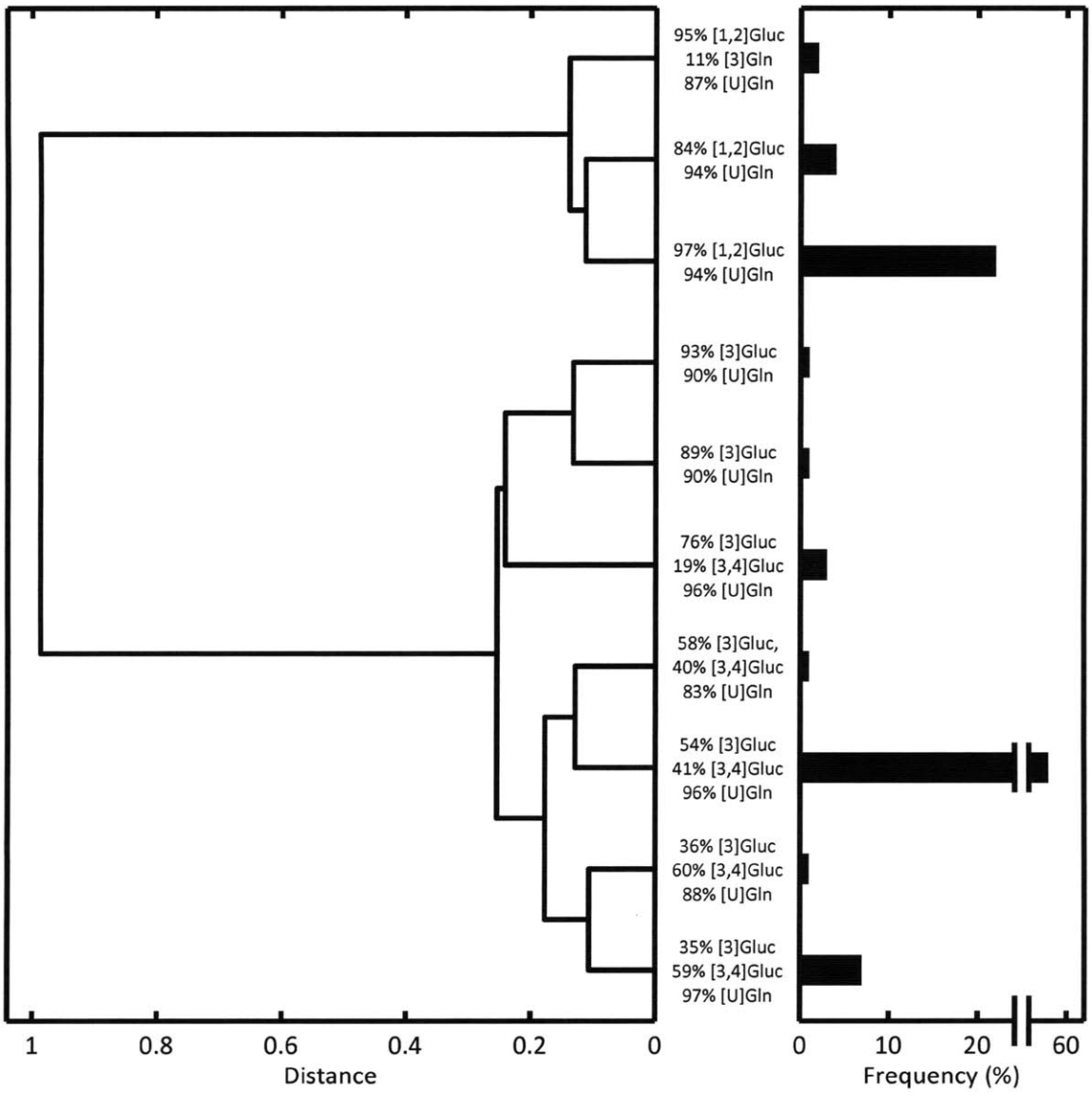

Figure 7-3: Hierarchical cluster tree of high-scoring tracer mixtures. The 100 highest-scoring tracer mixtures (over all rounds of evolution) were taken and clustered by composition fraction. The length of each tree on the left axes represents the distance (or dissimilarity) between the two tracer mixtures (or groups of tracer mixtures) being connected. The major tracers in each mixture are listed in the vertical direction. The bar lengths on the right axes represent the frequency of each cluster among the selected 100. The tree reveals that there are two general high-scoring mixtures: those using [1,2]Gluc and those using [3] and [3,4]Gluc. All mixtures predominantly utilize [U]Gln as a glutamine tracer.

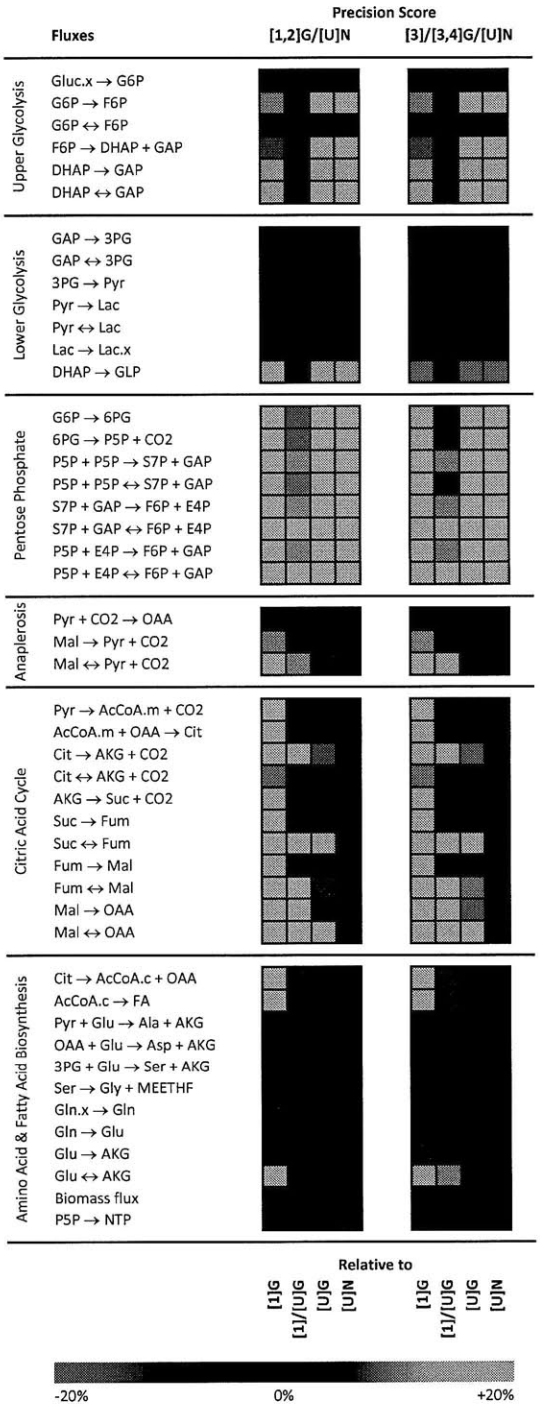

Figure 7-4: A heat map comparing flux-by-flux precision scores for traditional tracers ([1]Gluc, [1]/[U]Gluc, [U]Gluc, and [U]Gln) and evolved tracer mixtures ([1,2]Gluc/[U]Gln and [3]/[3,4]Gluc/[U]Gln). Green indicates a superior precision score for the evolved tracer and red indicates a superior score for the traditional tracer mixture.

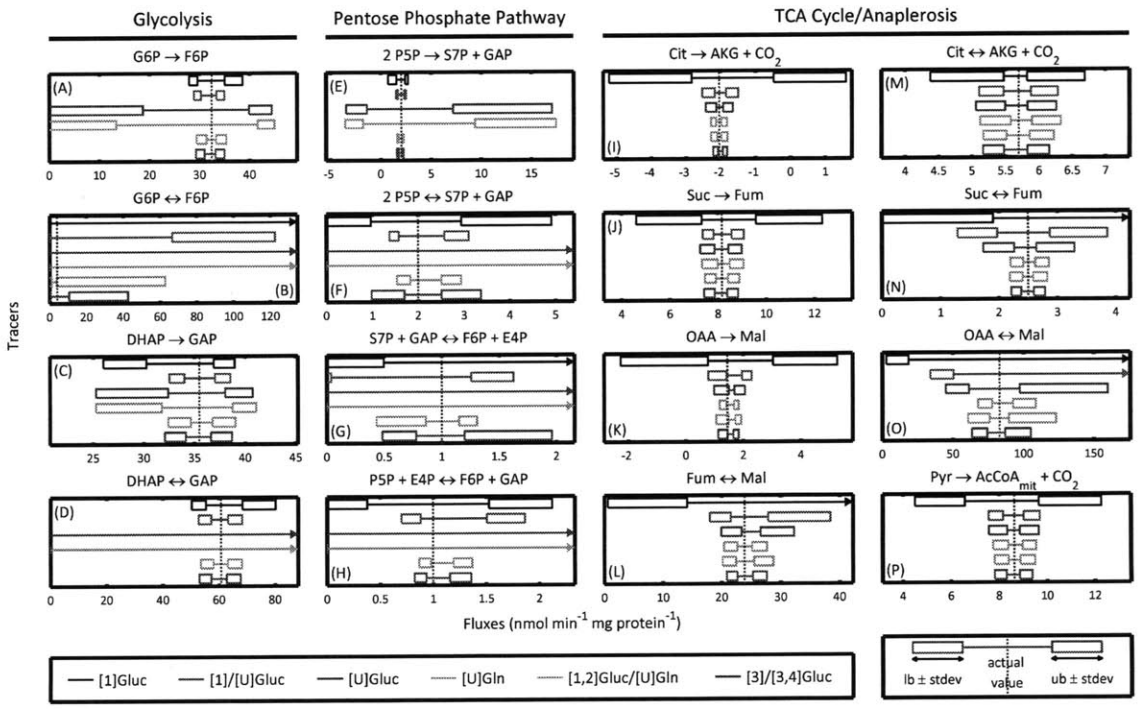

Figure 7-5: Simulated confidence intervals for selected fluxes when using common ^{13}C tracers compared to the high-scoring [1,2]Gluc/[U]Gln and [3]/[3,4]Gluc/[U]Gln tracer mixtures. Whereas glucose- and glutamine-only tracers perform poorly in different portions of the network, the tracer mixtures perform consistently well across all of central carbon metabolism, including glycolysis (A)-(D), the pentose phosphate pathway (E)-(H), and the TCA cycle (I)-(P).

score would change. For the sake of comparison, we also varied the tracer composition in a separate set of simulations to determine the sensitivity of the precision score with respect to tracer composition. The results are shown in Figure 7-6. The precision score sensitivity to the flux distribution is much less than the sensitivity to tracer composition, so we can have confidence in our selected tracer mixtures even if we are not absolutely certain of the actual metabolic flux distribution upon which the optimization was based.

7.7 Analysis of Tracer Behavior

The evolved tracer mixtures matched or outperformed all of the standard tracers throughout central carbon metabolism in almost every way. The precision score heat map (Figure 7-4) quickly highlights the pentose phosphate pathway as the major area of improvement. The optimized tracers also strongly outperformed [1]Gluc/[U]Gluc and [1]Gluc tracers in the TCA cycle and the [U]Gluc and [U]Gln tracers in glycolysis. To investigate our results further, we can return to actual confidence intervals and compare fluxes tracer-by-tracer in Figure 7-5.

1. For most glycolytic fluxes, the evolved mixtures demonstrated more precise confidence intervals than all standard tracers (Figures 7-5A, C and D). In a few instances, the evolved mixtures led to confidence intervals comparable to those of [1]/[U]Gluc but still better than the rest (Figure 7-5B).

2. For most fluxes of the pentose phosphate pathway, the evolved mixtures again had more precise confidence intervals (Figures 7-5E, G and H). In a few cases, [1]/[U]Gluc performed similarly well (Figure 7-5F).

3. For the majority of TCA cycle fluxes, the two evolved mixtures together with the [U]Gln tracer produce the narrowest confidence intervals (Figures 7-5I, K, L, N and O). For some fluxes, the evolved mixtures outperform all standard tracers (Figures 7-5J, M and P).

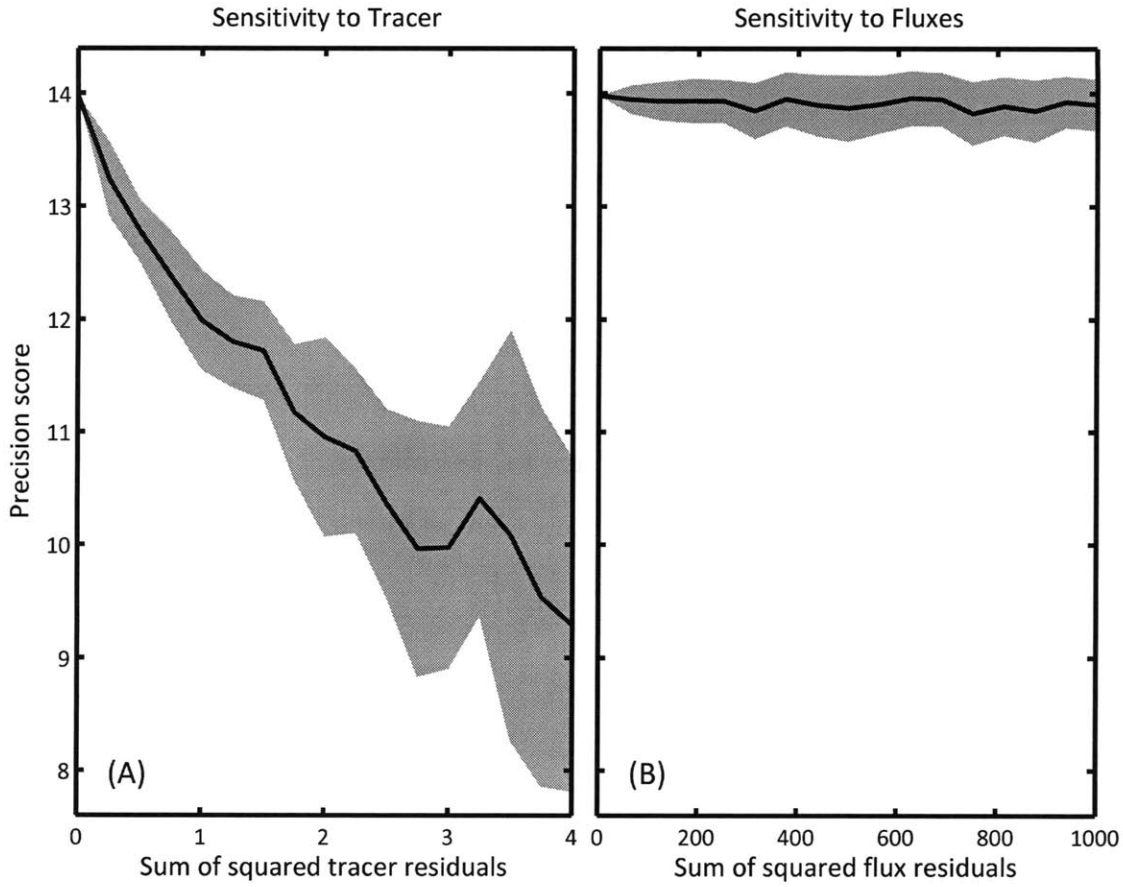

Figure 7-6: The sensitivity of precision score with respect to (A) tracer composition and (B) flux distribution. Solid lines indicate the average precision score for a particular sum of squared residuals while shading indicates one standard deviation above and below the average. In both subfigures, the initial tracer is a mixture of 100% [1,2]Gluc and 100% [U]Gln and and the initial flux distribution is the carcinoma distribution used in the tracer optimization. In subfigure (A), the flux distribution is held constant and the tracer composition is varied from a sum of squared residuals of 0 (original tracer) up to 4 (completely different tracers). In subfigure (B), the tracer composition is held constant and the flux distribution is varied from a sum of squared residuals of 0 (original tracer) up to 1000 (a simultaneous 15% change for all independent fluxes). On these scales, the precision score is much more sensitive to tracer, meaning that a high-scoring tracer will serve well even if the actual flux distribution does not match the hypothetical distribution used in the genetic algorithm.

The high precision score of our selected [1,2]Gluc/[U]Gln mixture is to be expected, since [1,2]Gluc performed well in glycolysis and the PPP and [U]Gln performed well in the TCA cycle, both for reasons previously noted in Chapter 6 (Figure 6-7). The appearance of the [3]Gluc/[3,4]Gluc/[U]Gln mixture as an optimal tracer is more surprising, and highlights the power of our evolutionary approach to find good tracers that could not easily have been chosen by intuition. To shed light on the reasons behind this mixture's efficacy, we can compare the precision of this optimal tracer with three of its close relatives: [3]Gluc, [3,4]Gluc/[U]Gln, and [4]Gluc/[U]Gln. A flux-by-flux precision score heat map of the selected tracer mixture versus these three similar mixtures is shown in Figure 7-7. (TCA cycle fluxes are not shown since precision scores there were very similar due to the common [U]Gln tracer.)

We first notice that [4]Gluc/[U]Gln scores lowest in the group, primarily because of poor precision in the pentose phosphate pathway and secondarily in upper glycolysis. As previously discussed (and shown in Figure 6-5B), the fourth atom of glucose is trapped in a cycle and fails to spread other positions, regardless of the flux distribution. When glucose is labeled both at the third and fourth positions, we see that pentose phosphate precision improves; the third labeled carbon, not limited to a futile cycle, is able to spread through the PPP atoms and as a result exhibits more sensitivity to fluxes there. The solitary [3]Gluc tracer represents an additional improvement as the cycle of atoms that was previously fully labeled (and hence insensitive) is now able to show gradations in labeling as the third glucose atom spreads through the network.

Moving from [3]Gluc to our optimized glucose tracer of [3]Gluc/[3,4]Gluc represents the final step in the progression. The difference in precision is simple. The precision of the S7P + GAP $\leftrightarrow$ E4P + F6P exchange flux increases greatly when we partially label the fourth glucose atom. This increase is most likely due to the atom transition $E4P_2 \leftrightarrow F6P_4$ within this reaction, which brings label out of the pentose phosphate pathway into glycolysis which then directly leads to the GLP measurement.

As one final demonstration of the mechanics behind the improved precision of the evolved tracers, positional atom labeling is shown for these selected mixtures in

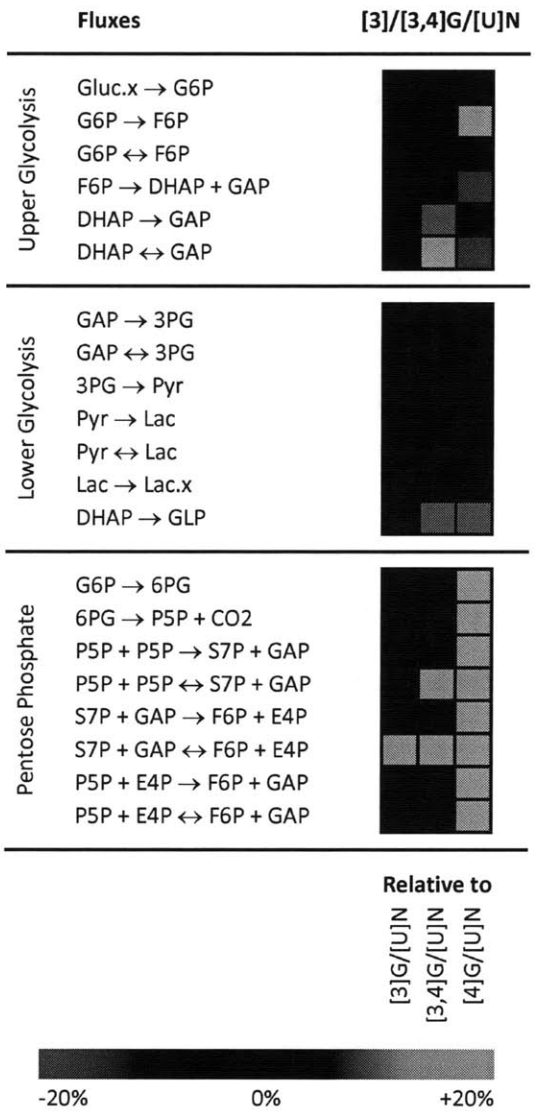

Figure 7-7: A heat map comparing flux-by-flux precision scores for the optimized [3]Gluc/[3,4]Gluc/[U]Gln tracer mixture to scores for the simpler [3]Gluc/[U]Gln, [3,4]Gluc/[U]Gln, and [4]Gluc/[U]Gln tracer mixtures. Green indicates a superior precision score for the evolved tracer and red indicates a superior score for the traditional tracer mixture.

Figure 7-8. Both of these mixtures distribute label generously among multiple atoms of measured compounds, endowing their labeling pattern with high sensitivity to its flux distribution.

7.8 Experimental Validation

To validate these simulated and admittedly hypothetical results of our genetic algorithm, we conducted actual isotopic metabolic flux analyses on the original A549 carcinoma cell line with one of our optimized tracer mixtures ([1,2]Gluc/[U]Gln), one average tracer mixtures ([1]Gluc/[U]Gluc), and one poor tracer ([U]Gln). We compared the resulting experimental precision scores to the previously simulated precision scores for each of the same tracers (Figure 7-9).

We see that, most importantly, the general trend holds steady. In both simulation and experiment, [1,2]Gluc/[U]Gln scores best, followed by [1]Gluc/[U]Gluc, with [U]Gln scoring worst, validating our tracer optimization study. There are two major differences in the two sets of scores; however, both can be accounted for.

First, overall, the experimental scores were lower. This is because the measurement standard error used for the simulations was optimistically too low. In actuality, the standard error was higher across the board, leading to increased confidence intervals and decreased precision scores. Since this was a general effect, this did not decrease the efficacy of the original optimization since all trends remain the same.

The second discrepancy is that the improvement in precision the [1,2]Gluc/[U]Gln mixture offers over the [1]Gluc/[U]Gluc mixture is lessened when we move from simulation to experiment. This can be traced to the assumptions we originally made regarding the P5P measurement. This measurement turned out to be less reliable than expected and as a result was assigned a higher-than-average standard error. Since much of the optimized mixture's advantage arises from gains made in pentose phosphate pathway precision, and since many of the confidence intervals in this pathway depend heavily on the P5P measurement, the optimized tracer's overall precision score suffered somewhat relative to the other scores.

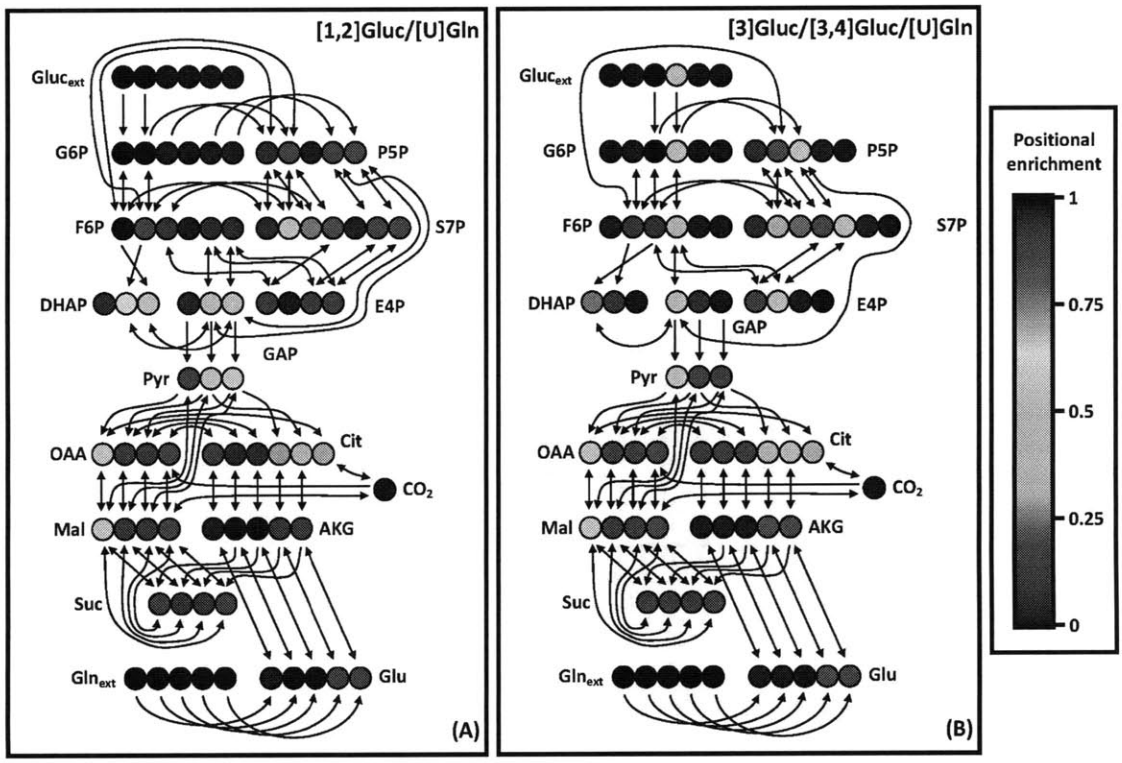

Figure 7-8: Atom transition networks and positional fractional labeling for evolved tracer mixtures of glucose and glutamine. Fractional labeling is indicated by a color map, where dark red indicates all atoms at that position are ^{13}C and dark blue that all atoms ^{12}C. No natural labeling was assumed in the creation of these maps. Atom transitions are indicated for all positions where fractional ^{13}C labeling exceeds 10%.

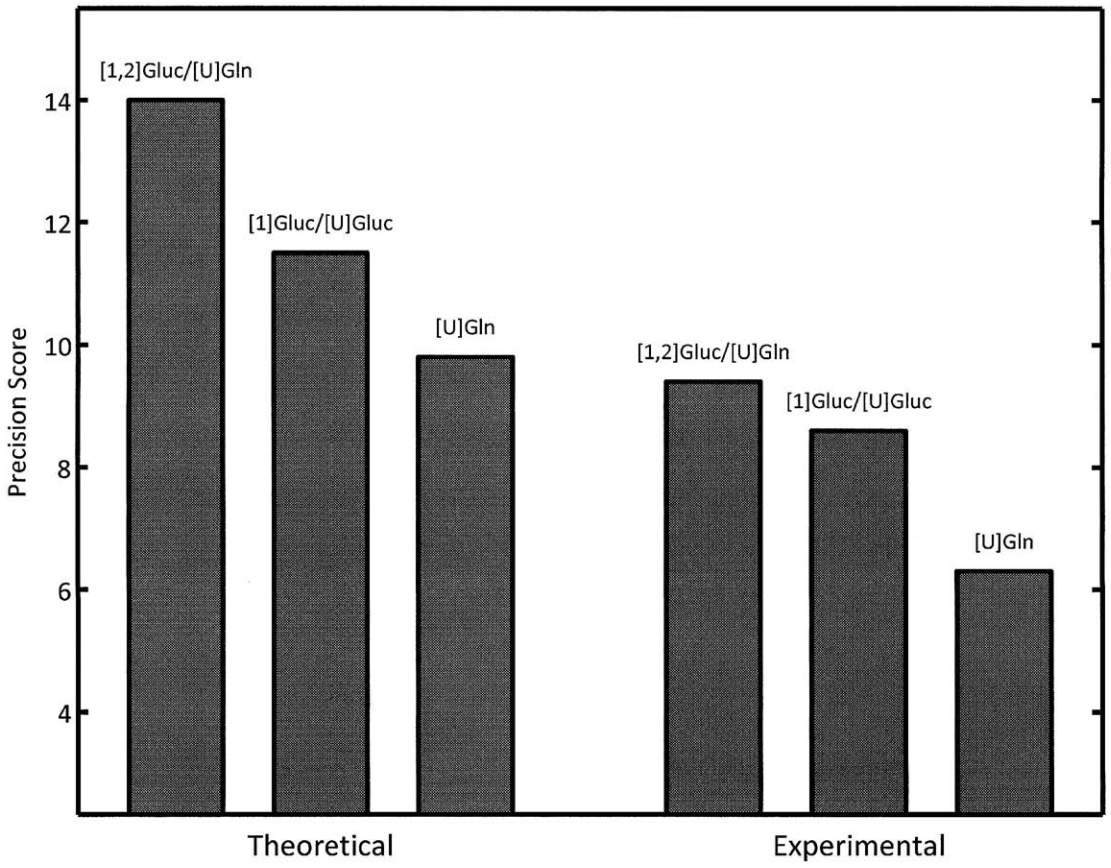

Figure 7-9: Simulated and experimental precision scores for experiments on A549 carcinoma metabolism utilizing three different tracers: [1,2]Gluc/[U]Gln, [1]Gluc/[U]Gluc, and [U]Gln. Experimental precision scores follow the same trend predicted in theory, validating our method.

7.9 Discussion

We have created a genetic algorithm that can search through a large, unpredictable space of tracer combinations to find prominently performing mixtures for high-precision metabolic flux analysis. We applied this algorithm to tumor metabolism and discovered two optimal tracer mixtures, [1,2]Gluc/[U]Gln and [3]Gluc/[3,4]Gluc/[U]Gln. These tracers lead to flux estimates of superior quality in both glycolysis and the TCA cycle, where most typical, single-substrate glucose and glutamine tracers only target the former or the latter, respectively. These evolved mixtures also offer especial improvement in pentose phosphate pathway analysis. We have confirmed our theoretical results by experimenting with selected tracer mixtures and demonstrating that they do indeed generate more precise flux confidence intervals.

These methods greatly strengthen experimental design capabilities for metabolic flux analysis. In particular, with this method of tracer selection we can customize tracers for targeted flux studies. Although much of current MFA research is directed at measuring metabolism on large scales, some of the most powerful and meaningful applications occur when we focus on specific pathways, and even specific reactions. Our genetic algorithm for tracer optimization is a powerful tool in these kinds of targeted studies.

Our evolutionary selection proved its usefulness when it highlighted a tracer containing [3]Gluc and [3,4]Gluc as an important tracer in cancer metabolism. Even though we had extensive experience with this particular network and flux distribution, this optimal tracer mixture was unexpected and nonintuitive and would have most likely remained undiscovered in a less systematic search. By studying these newly discovered tracers we can gain a deeper understanding of the complex network of atom transitions in metabolism that will aid us in further experimental design.

Chapter 8

NMFA of *Yarrowia lipolytica*

8.1 Introduction

Petroleum-based fuels have generated concerns due to rising prices, environmental damages, and supply instabilities. As a result, interest has grown in alternative energy sources, and in particular, liquid biofuels [137]. However, many issues still wait to be addressed before biofuels can substantially cut into petroleum consumption. Most of today's biofuels are produced from food crops (corn and soybeans), but it is unlikely that we can significantly increase production of such crops without negatively impacting food supplies [70]. To produce biofuels in bulk, we need to begin with a non-food based feedstock, such as corn stover, switch grass, or some other sort of cellulosic biomaterial [20].

Ethanol is by far the largest biofuel currently produced in the United States. Unfortunately, ethanol has many serious drawbacks as a biofuel compared to its main competitor, biodiesel. Ethanol requires significant amounts of energy for separation and purification since it is miscible with water, and overall yields only 25% more energy than that consumed in its production, while biodiesel produces 93% more [70]. Biodiesel also generates less pollution [113].

One promising solution that we are currently studying is lipid production by the oleaginous yeast, *Yarrowia lipolytica*. This organism has great potential as a biofuels producer. It is able to consume a wide variety of substrates, including sugars,

n-alkenes, 1-alkenes, acetate, and alcohols [17, 75, 82, 109, 144]. *Y. lipolytica* is able to produce a variety of metabolic byproducts in large quantities, including citrate, isocitrate, α-ketoglutarate, and most promisingly, triglycerides [76, 96, 111, 112]. Triglycerides formed intracellularly by *Yarriowia's* metabolism can be harvested and transformed into biodiesel via base-catalyzed transesterfication, a well-established method for commercial-scale production (Figure 8-1) [147]. If *Yarrowia* can be engineered to both consume cellulosic carbon sources and produce sufficient amounts of lipids, it could be a powerful catalyst for biofuels production.

Our efforts have been focused on understanding and increasing lipid production in *Y. lipolytica*, using both genetic and environmental means. We worked with a new strain and analyzed it for lipid production under different bioreactor conditions. In particular, we varied aeration rates to see if increased oxygen availability allowed for greater lipid production. We measured incoming and outgoing extracellular fluxes and observed how carbon utilization changed both over time and at different levels of aeration. We conducted NMFA experiments at different growth phases and studied the results in an attempt to understand how lipid production might be optimized. We saw that by tripling our aeration, fatty acid yields increased by as much as 60%. These increases were accomplished by rerouting carbon through the pentose phosphate pathway.

8.2 Bioreactor Materials and Methods

Yarrowia was grown in medium containing of 3.4 g L^{-1} YNB (MP Biomedicals), 20 g L^{-1} glucose (Mallinckrodt), and 3 g L^{-1} ammonium sulfate (Mallinckrodt). Cultures were grown in shake flasks at 30°C to an optical density of greater than 4 and then were used to inoculate 1.3-L batch reactors (New Brunswick Scientific) with an initially targeted optical density of 0.5. The initial reactor working volume was 900 mL (before inoculation). Glucose boluses were given to reactors when necessary to keep concentrations above 1 g L^{-1}. Bioreactors were controlled at a pH of 5.5 and a temperature of 28°C. One set of bioreactors was held to a sparging rate of 2.5 L

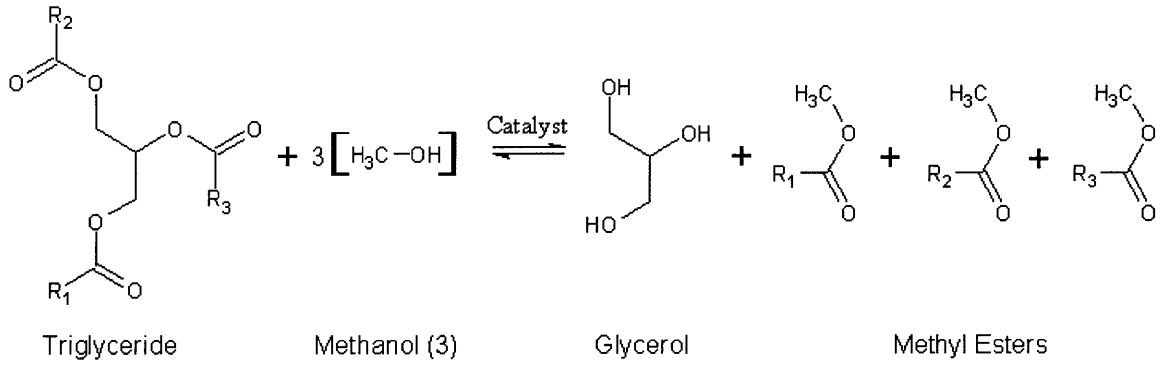

Figure 8-1: The transesterification of a triglyceride, where R_1, R_2, and R_3 are alkyl groups.

min^{-1} of air with a maximum agitation rate of 600 rpm, while another set was held to a sparging rate of 7.5 L min^{-1} and an agitation rate of 1000 rpm.

Samples were taken approximately once per doubling time (about every three hours). Glucose and ammonia were measured on a YSI 7100 system (YSI Life Sciences). Optical density was measured at 600 nm on an Ultraspec 2100 pro spectrophotometer (Amersham Biosciences). Dry cell weights were calculated using a previously determined correlation (0.37 g L^{-1} OD^{-1}). Extracellular citrate was measured in a high-performance liquid chromatography (HPLC) system with a Waters 2690 separations module connected to a Waters 410 refractive index detector (Waters). The samples were separated on a BioRad Aminex HPX-87H ion exclusion column for organic acid analysis with 14 mM sulfuric acid as the mobile phase at a flow rate of 0.7 mL min^{-1} with an organic acid column. The percentage of CO_2 in the offgas was measured with an Agilent 300A Micro GC (Agilent). The offgas flow rate was also monitored and recorded.

8.3 Fatty Acid Quantification

To measure intracellular lipid amounts (primarily palmitate, stearate, and oleate), we first subjected a defined amount of biomass to 30 minutes of vortexing in a mixture of 5 mL chloroform, 2.5 mL methanol, and 2 mL of water to lyse cells and extract lipids [56, 152]. The resulting mixture was centrifuged at 2000 g for 10 minutes. The chloroform layer was removed, dried, and treated with with Methyl-8 reagent (Agilent) at 60°C for 60 minutes to transesterify fatty acids.

Gas chromatography/mass spectrometry (GC/MS) analysis for fatty acid identification and quantification was performed using an Agilent 6890 GC equipped with a 30m DB-35MS capillary column connected to an Agilent 5975B MS operating under electron impact (EI) ionization at 70 eV. One µL of sample was injected in a 9:1 split mode at 280°C, using helium as the carrier gas at a flow rate of 1 mL min^{-1}. The GC oven temperature was held at 100°C for 5 min and ramped to 175°C (rate: 15°C min^{-1}, hold: 1 min), 205°C (rate: 3°C min^{-1}, hold: 5 min), 215°C (rate: 5°C min^{-1},

hold: 4 min), 220°C (rate: 2.5°C min^{-1}, no hold), 255° (rate: 5°C min^{-1}, no hold), and finally 290°C (rate: 15°C min^{-1}, no hold) for a total run time of approximately 44 min. The MS source and quadrupole were held at 280°C and 150°C, respectively, and the detector was operated in scan mode.

Palmitate, oleate, and stearate peaks were identified at retention times of 17.3, 21.7, and 22.0 minutes, respectively. Ion ranges of 270-274, 264-268, and 284-288 were respectively integrated for each of the three fatty acids and compared with similar peaks in a defined standard (Sigma-Aldrich) to arrive at intracellular concentration measurements.

8.4 NMFA Materials and Methods

NMFA experiments were conducted at three different growth phases under low aeration and twice under high aeration (see Figure 8.1). For each experiment, 10 mL of 450 g L^{-1} [1-^{13}C]glucose and 150 g L^{-1} [U-$^{13}C_6$]glucose (Cambridge Isotope Laboratories) were added to the reactor culture after which sampling was conducted in a geometric sequence. One manual sample was taken before labeling. Eight samples were collected using the rapid sampling apparatus at 2, 4, 8, 15, 30, 60, 120, and 240 seconds after labeling, and seven additional samples were collected manually at 8, 15, 30, 60, 120, 180, and 240 minutes (Figure 8-2). Some sampling irregularities occurred from experiment to experiment (e.g., small timing differences and occasional missed samples), but since each flux analysis was self-contained, this did not significantly affect the overall results and our ability to compare different experiments. Table 8.2 lists the intracellular organic and amino acid fragments whose labeling was measured via GC/MS and used as inputs for NMFA.

Samples were immediately quenched in a 60% (v/v) methanol solution held at −25°C in an ethanol cold bath. Quenched samples were quickly centrifuged (at 2000 g and −10°C for 10 min), decanted, and rinsed with cold pure methanol. A mixture of 5, 2.5, and 2 mL of cold chloroform, methanol, and water (respectively) was next added to the cell pellet, and the mixture was vortexed for 30 minutes to lyse cells

Experiment	L1	L2	H1	H2
Aeration (L min^{-1})	250	250	750	750
Growth Phase	Late Linear	Stationary	Late Linear	Stationary
Start Time (h)	21.58	38.07	22.63	29.68
End Time (h)	25.58	41.98	25.65	33.68
Samples	13	12	12	13

Table 8.1: A list of the bioreactor-based NMFA experiments we conducted on *Yarrowia lipolytica*. Nonstationary flux analysis was conducted during the late linear and stationary growth phases in both experiments under low (L1 and L2) and high (H1 and H2) aeration.

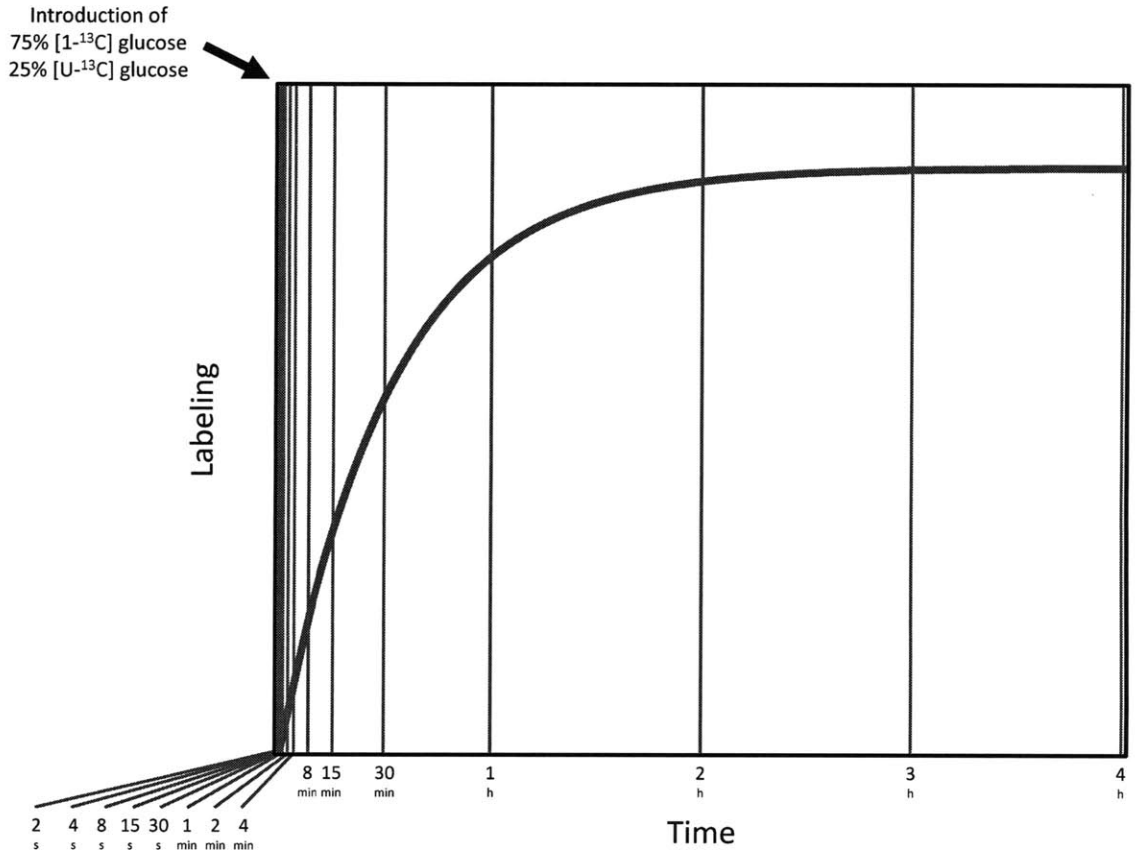

Figure 8-2: Typical bioreactor sampling times for NMFA experiments of *Y. lipolytica*. We attempted to pattern sampling geometrically and thus capture both fast and slow labeling dynamics of the system. Rapid samples were collected at 2, 4, 8, 15, 30, 120, and 240 seconds. Manual samples were collected at 8, 15, 30, 60, 120, 180, and 240 minutes. Acceptable variations in sampling times occurred from experiment to experiment due to differences in timing and occasionally missed samples.

Metabolite	Mass	Carbons	Formula
AKG	346	12345	$C_{14}H_{28}O_5NSi_2$
Ala	232	23	$C_{10}H_{26}ONSi_2$
Ala	260	123	$C_{11}H_{26}O_2NSi_2$
Asp	302	12	$C_{14}H_{32}O_2NSi_2$
Asp	376	12	$C_{16}H_{38}O_3NSi_3$
Asp	390	234	$C_{17}H_{40}O_3NSi_3$
Asp	418	1234	$C_{18}H_{40}O_4NSi_3$
Cit	459	123456	$C_{20}H_{39}O_6Si_3$
Gln	431	12345	$C_{19}H_{43}O_3N_2Si_3$
Glu	330	2345	$C_{16}H_{36}O_2NSi_2$
Glu	432	12345	$C_{19}H_{42}O_4NSi_3$
Gly	218	2	$C_9H_{24}ONSi_2$
Gly	246	12	$C_{10}H_{24}O_2NSi_2$
Ile	200	23456	$C_{11}H_{26}NSi$
Ile	274	23456	$C_{13}H_{32}ONSi_2$
Leu	274	23456	$C_{13}H_{32}ONSi_2$
Mal	419	1234	$C_{18}H_{39}O_5Si_3$
Pyr	174	123	$C_6H_{12}O_3NSi$
Ser	288	23	$C_{14}H_{34}NOSi_2$
Ser	302	12	$C_{14}H_{32}O_2NSi_2$
Ser	362	23	$C_{16}H_{40}O_2NSi_3$
Ser	390	123	$C_{17}H_{40}O_3NSi_3$
Suc	289	1234	$C_{12}H_{25}O_4Si_2$
Thr	376	234	$C_{17}H_{42}O_2NSi_3$
Thr	404	1234	$C_{18}H_{42}O_3NSi_3$
Val	260	2345	$C_{12}H_{30}ONSi_2$
Val	288	12345	$C_{13}H_{30}O_2NSi_2$

Table 8.2: Free organic and amino acid fragment MIDs measured in the *Y. lipolytica* study. Intracellular metabolites were derivatized to form TBDMS functional groups and then were fragmented and measured by GC/MS.

and extract intracellular metabolites.

The vortexed mixtures were centrifuged again at the same settings. The upper aqueous layer was removed and set aside. Cold methanol (2 mL) and water (2 mL) were added to the remaining chloroform layer, and the new mixture was vortexed for 30 seconds and centrifuged a third time at the same settings. The new upper aqueous layer was removed and pooled with the previous. This combined methanol/water extraction solution was dried under air.

Dried polar metabolites were dissolved in 60 µL of 2% methoxyamine hydrochloride in pyridine (Pierce), sonicated for 30 minutes, and held at 37°C for 2 hours. After dissolution and reaction, 90 µL MBTSTFA + 1% TBDMCS (Pierce) was added and samples were incubated at 55°C for 60 minutes. Gas chromatography/mass spectrometry (GC/MS) analysis was performed using an Agilent 6890 GC equipped with a 30m DB-35MS capillary column connected to an Agilent 5975B MS operating under electron impact (EI) ionization at 70 eV. One µL of sample was injected in a 9:1 split mode at 270°C, using helium as the carrier gas at a flow rate of 1 mL min $^{-1}$. The GC oven temperature was held at 100°C for 3 min and increased to 300°C at 3.5°C min^{-1} for a total run time of approximately 60 min. The MS source and quadrupole were held at 230°C and 150°C, respectively, and the detector was operated in scan mode.

Extracellular analysis was conducted from each rapid and manual sampling time point to measure glucose MIDs via aldonitrile pentapropionate derivatization and GC/MS as explained in Chapter 5. MIDs were corrected for natural abundances and used to calculate the fractional abundances of [1-^{13}C] and [U-$^{13}C_6$]glucose in the media over the course of each NMFA experiment.

8.5 Extracellular Flux Fitting

The measurement of extracellular fluxes is very important to the overall flux estimation, since flux measurements serve as stoichiometric anchors for the remainder of the analysis. Because fluxes are derivatives of extracellular concentration measurements,

small errors in the primary data can lead to large amounts of noise in calculated fluxes, dramatically changing the final estimated flux distribution. In order to utilize the full set of our extracellular parameters in a systematic way to obtain flux measurements at the times of our NMFA experiments, we first observed distinct phases of growth over the course of the experiment and second applied a simple parameter fitting scheme to give each measurment equal weight in the final results.

Cell growth is exponential over the first 17 hours of the experiment, as expected. However, instead of quickly leveling out to a stationary phase, the system enters a prolonged "linear" phase where biomass and metabolic byproducts are formed at a linear rate. After about 30 hours, the system then enters a stationary phase where cell growth ends but carbon dioxide and citrate production continue. We assumed that for exponential growth,

$$\begin{aligned} b(t) &= b_{\text{exp},0} \cdot \exp\left(k_{b,\text{exp}} \cdot t\right) \\ x(t) &= k_{x,\text{exp}} \cdot b_{\text{exp},0} \left[\exp\left(k_{b,\text{exp}} \cdot t\right) - 1\right] \end{aligned} \tag{8.1}$$

where b and x are the biomass and byproduct concentrations in time, $b_{\text{exp},0}$ is the initial biomass concentration, $k_{b,\text{exp}}$ and $k_{x,\text{exp}}$ are parameters to be estimated in the fitting process, and t is time. During the linear and stationary phases, we modeled the system with the following, simple equations:

$$\begin{aligned} b(t) &= b_{\text{lin},0} + k_{b,\text{lin}} \cdot t \\ x(t) &= x_{\text{lin},0} + k_{x,\text{lin}} \cdot t \end{aligned} \tag{8.2}$$

where $b_{\text{lin},0}$ and $x_{\text{lin},0}$ are the initial biomass and byproduct concentrations (in the linear phase of interest) and $k_{b,\text{lin}}$ and $k_{x,\text{lin}}$ are parameters to be estimated.

"Effective concentrations" are numerically calculated for carbon dioxide for Equations 8.1 and 8.2 according to the expression below:

$$c_{CO_2}(t) = \frac{\rho}{V(t)} \int_0^t F(\tau) \cdot f_{CO_2}(\tau) \cdot \left[1 + \min\left(0, \frac{V'(\tau)}{V(\tau)}\right)\right] d\tau \tag{8.3}$$

where c_{CO_2} is the effective concentration of carbon dioxide, ρ is the density of carbon dioxide, V is the culture volume, F is the bioreactor's offgas flow rate, f_{CO_2} is the

mass fraction of carbon dioxide in the offgas, and t is time.

Both the low- and high-aeration experiments underwent separate fitting procedures with the assumption that different metabolic behavior might be occurring under the different environmental conditions. Both fits were overdetermined, so a least squares regression was performed, giving us the constants shown in Table 8.3. However, because there were very few degrees of freedom in both fits (i.e., the fits were only slightly overdetermined), our estimated measurements closely matched the actual concentrations (see Figures 8-3 and 8-4). This allowed us to obtain smooth profiles and consistent extracellular flux calculations without deviating far from our measurements. The extracellular fluxes arising from this fitting process are shown in Table 8.4.

8.6 Carbon Balances

We measured all major carbon-containing species entering and exiting the bioreactors and then compared the overall molar carbon flux into and out of the system to determine if the inputs and outputs were correctly balanced. Glucose is the main input to the system and carbon dioxide, citrate, biomass lipids, and residual biomass the outputs. To convert biomass to carbon-moles, we assumed that the elemental composition of *Yarrowia* was similar to previously reported values for *Saccharomyces cerevisiae* ($CH_{1.94}O_{0.52}N_{0.25}P_{0.025}$) [138]. Carbon balances for both the low-aeration and high-aeration runs are shown in Figures 8-5A and B, respectively. A comparison of the balances yields some interesting observations:

1. The measured amounts of carbon entering and exiting the system show very good agreement. This strong correlation validates our extracellular measurement methods.

2. Citrate production is slightly up until the stationary phase under both aeration conditions, when it becomes the major metabolic product.

Product	$k_{\text{exponential}}$		k_{linear}		$k_{\text{stationary}}$	
	Low	High	Low	High	Low	High
Residual biomass	0.2867	0.2551	0.6179	0.2875	0.0000	0.0677
Glucose	1.9792	2.7468	2.1881	1.9250	1.4553	1.6824
Carbon dioxide	0.9949	1.9355	1.1771	1.2876	0.6811	0.6790
Citrate	0.0000	0.0000	0.1644	0.1785	1.0487	0.8221
Palmitate	0.0451	0.1565	0.1641	0.2166	0.0000	0.0000
Stearate	0.0810	0.3861	0.2482	0.3597	0.0226	0.0232
Oleate	0.0558	0.0000	0.0021	0.0420	0.0105	0.0071

Table 8.3: Fitted rate constants for metabolic byproducts of *Y. lipolytica* over exponential, linear, and stationary phases calculated using Equation 8.1 and 8.2. Residual biomass constants in the exponential phase have units h^{-1} and byproducts in the exponential phase have units (g product) (g residual biomass)$^{-1}$. All constants in the linear and stationary phases have units of (g product) L^{-1} h^{-1}.

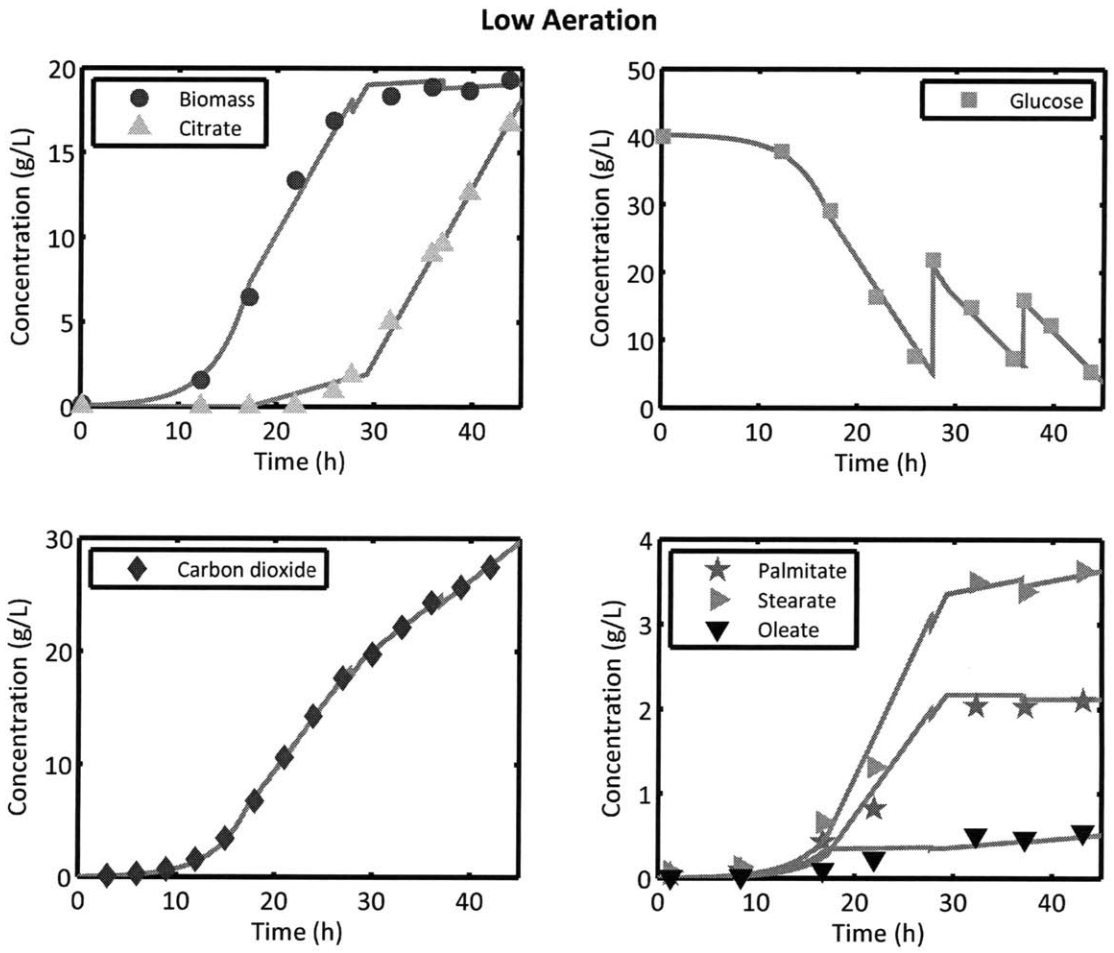

Figure 8-3: Fitted parameters for metabolic byproducts under low-aeration conditions. Biomass, citrate, glucose, carbon dioxide, and lipid measurements were used to model exponential, linear, and stationary growth phases according to Equations 8.1 and 8.2. Carbon dioxide values shown here are in the form of "effective concentrations", calculated from offgas flow rates, offgas CO_2 concentrations, and cell culture volume changes.

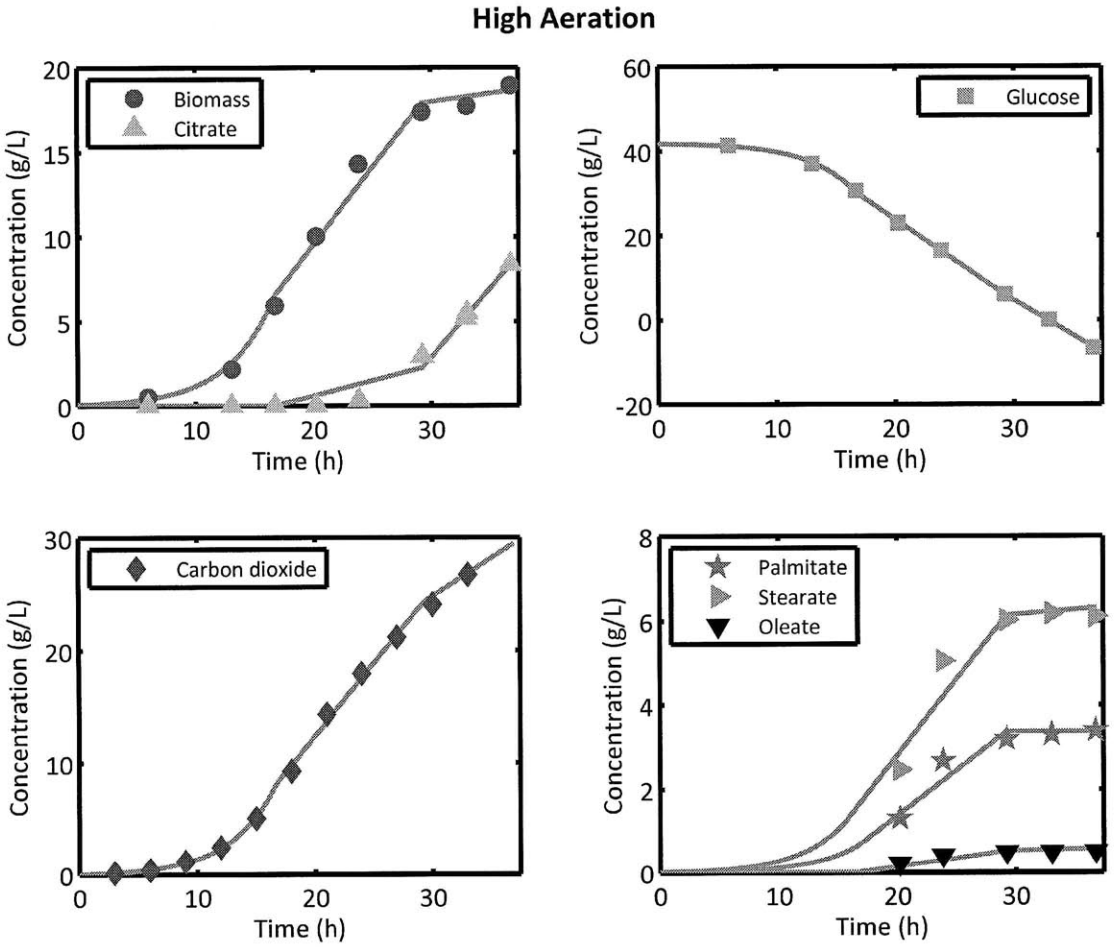

Figure 8-4: Fitted parameters for metabolic byproducts under high-aeration conditions. Biomass, citrate, glucose, carbon dioxide, and lipid measurements were used to model exponential, linear, and stationary growth phases according to Equations 8.1 and 8.2. Carbon dioxide values shown here are in the form of "effective concentrations", calculated from offgas flow rates, offgas CO_2 concentrations, and cell culture volume changes.

	$v_{\text{exponential}}$		v_{linear}		$v_{\text{stationary}}$	
Product	Low	High	Low	High	Low	High
Residual biomass	1.00	1.00	21.67	9.27	0.00	2.84
Glucose	4.22	9.34	12.84	12.39	8.23	8.47
Carbon dioxide	2.09	4.58	24.90	25.05	15.15	17.16
Citrate	0.00	0.00	0.80	0.80	5.35	4.76
Palmitate	0.02	0.06	0.60	0.72	0.00	0.00
Stearate	0.03	0.14	0.81	1.08	0.08	0.09
Oleate	0.02	0.00	0.01	0.13	0.04	0.03

Table 8.4: Extracellular fluxes for metabolic byproducts of *Y. lipolytica* over exponential, linear, and stationary phases. Fluxes in the exponential phase have units (mmol product) (mmol residual biomass)$^{-1}$ h^{-1} and fluxes in the linear and stationary phases have units (mmol product) L^{-1} h^{-1}.

3. High aeration appears to induce higher aerobic activity, as the fraction of carbon outputed in the form of CO_2 increases from 34 to 41% when aeration is increased.

4. Approximately the same amount of biomass is produced under both aeration conditions and over all growth phases. However, a significantly larger fraction of biomass is devoted to intracellular lipids (presumably in the form of triacylglycerides) under high-aeration conditions.

We can further investigate the last point mentioned above by calculating lipid yields (grams per gram glucose) over time for each level of bioreactor aeration. In Figure 8-6, both instantaneous and cumulative yields are plotted with respect to time. We used our fitted extracellular fluxes in these calculations, which is why three the three growth phases are distinctly and smoothly represented. Lipid yields are significantly higher under high aeration in both the exponential phase (0.09 versus 0.20 g g^{-1}) and linear phase (0.19 versus 0.32 g g^{-1}). Under high aeration, lipid yields approach but never exceed the maximum stoichiometric yield of 0.38 g g^{-1}.

8.7 Flux Estimation

A network of central metabolism was created to model *Yarrowia lipolytica*. The complete set of reactions and atom transitions can be found in Tables 8.5 and 8.6. Reactions representing glycolysis, the pentose phosphate pathway, the citric acid cycle, amino acid biosynthesis, anaplerosis, the transhydrogenase cycle, one-carbon metabolism, and extracellular transport were all included. Several metabolites were assumed to be present in both the cytosol and the mitochondria (namely, malate, oxaloacetate, and acetyl-CoA). Other metabolites known to be present in both compartments (e.g., pyruvate) were assumed to have rapid exchange rates and therefore were lumped and treated as one. The proteinogenic amino acid composition of *Saccharomyces cerevisiae* was used to set the stoichiometry of amino acids in the biomass formation reaction [138]. The stoichiometric ratio of all other metabolites going to biomass is patterned after reactions in a previous study of *E. coli* [12]. We assumed

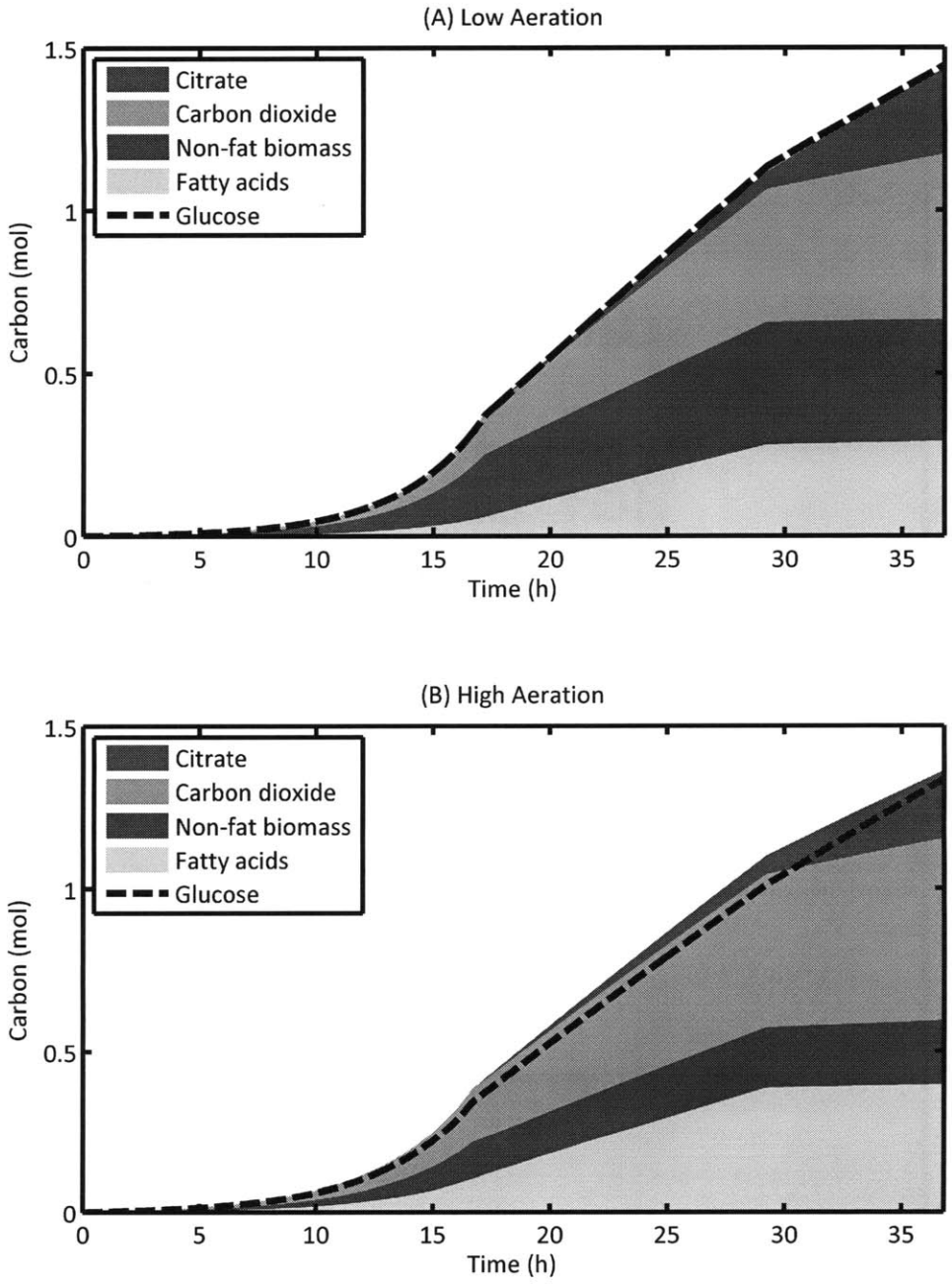

Figure 8-5: Carbon balances for *Yarrowia lipolytica* bioreactors under (A) low aeration and (B) high aeration. The carbon-containing input was glucose, represented by the dashed line, while the major carbon-containing products were intracellular fatty acids, other biomass, carbon dioxide, and citrate. The cumulative molar carbon consumed or produced as any of these substrates or products, respectively, is represented by the area of the "slice" corresponding to the particular substance.

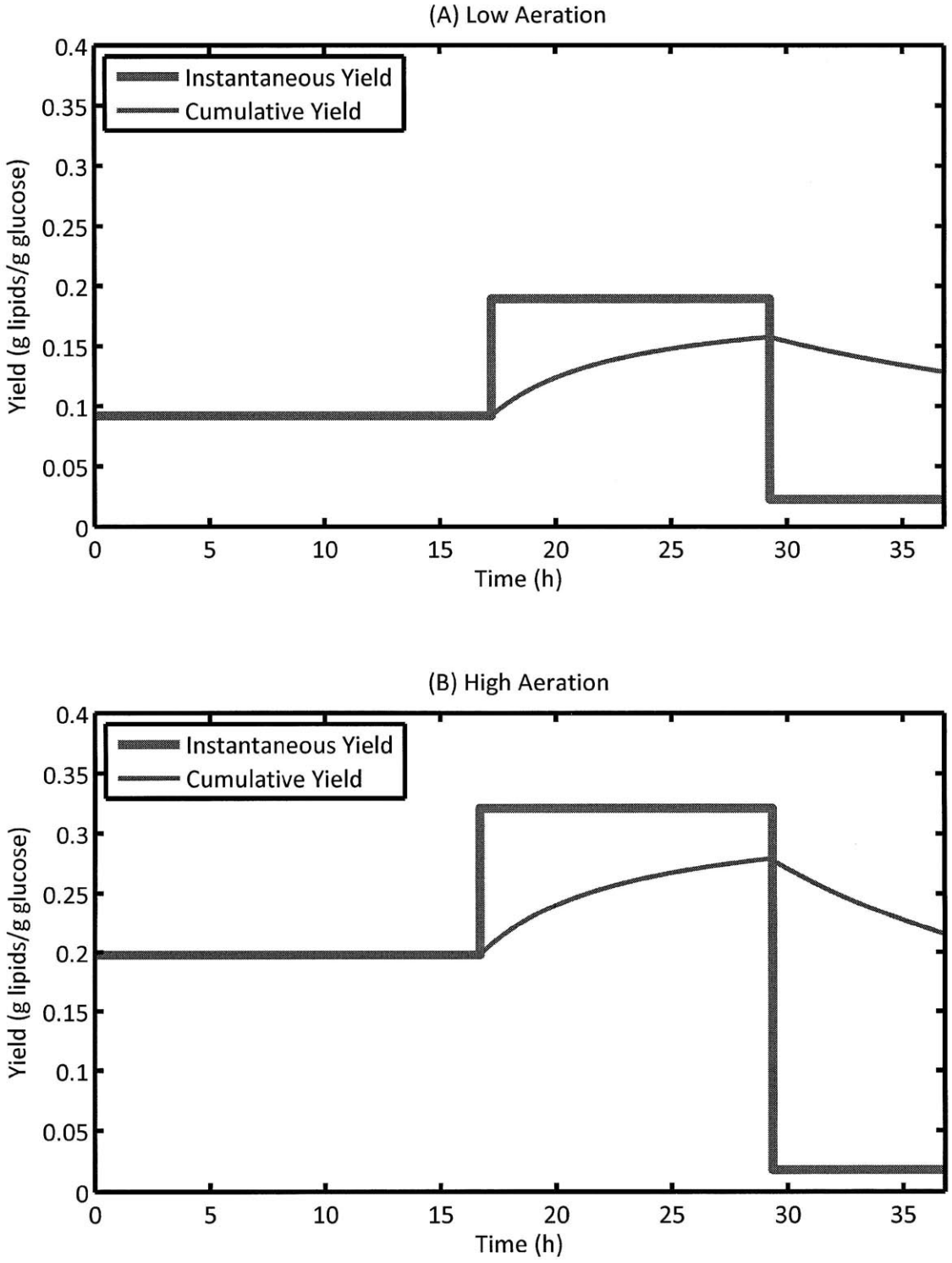

Figure 8-6: Instantaneous and cumulative yields for lipids on glucose by *Yarrowia lipolytica* in bioreactors under low- and high-aeration conditions. Yields are calculated using the fitted fluxes in Table 8.3.

that NADPH production and consumption were balanced since NADPH is critical to lipid production and since it only has a limited number of sources and sinks in metabolism [26]. The pentose phosphate pathway and the transhydrogenase cycle were the major NADPH sources while lipid and amino acid synthesis were the major sinks. NADPH is not able to transverse membranes was modeled solely in the cytosol.

Transient GC/MS measurements were successfully taken; Tables E.1 through E.8 in Appendix E list every mass isotopomer measurement for every time point of each experiment. Fits were successfully obtained for each NMFA experiment. Table 8.7 shows that each estimation was highly overdetermined, and the resulting sum of squared residuals fell well within the expected lower and upper bounds. Simulated measurement profiles resulting from the final estimated flux distribution are plotted over measurements in Figures E-1 through E-4 in Appendix E. The simulated values almost all completely fall within the standard error of the experimental measurements. Studying the measurement profiles more closely, we see that the system behave dynamically as expected. The metabolites farthest upstream (pyruvate and alanine), increase in labeling the most quickly, while the metabolites downstream in the TCA cycle display slower labeling behavior. Some peripheral metabolites, such as leucine, isoleucine, and threonine, never label completely even though they come to a steady state, suggesting that their pools are being diluted either during metabolism via proteinogenic amino acid degradation or during measurement via sequestered "cold" pools. As a result, a handful of dilution fluxes of both kinds were inserted into the model to account for these effects.

Fluxes were successfully estimated for each NMFA experiment. Flux distributions for each experiment are presented in Figure 8-7. (Values shown have been normalized to glucose input fluxes of 100 for each experiment.) A comprehensive list of flux values and confidence intervals for the entire network can be found in Appendix E in Tables E.9 through E.12. Confidence intervals were excellent for almost all estimated fluxes; in most cases the 95% intervals for net fluxes were between ±5% and ±10% of the flux value.

Upper Glycolysis			
v_1	$Gluc_{ext}$ (abcdef)	$\leftrightarrow$	G6P (abcdef)
v_2	G6P (abcdef)	$\leftrightarrow$	F6P (abcdef)
v_3	F6P (abcdef)	$\rightarrow$	DHAP (cba) + GAP (def)
v_4	DHAP (abc)	$\leftrightarrow$	GAP (abc)
v_5	DHAP (abc)	$\rightarrow$	Glyc (abc)
Lower Glycolysis			
v_6	GAP (abc)	$\leftrightarrow$	3PG (abc)
v_7	3PG (abc)	$\leftrightarrow$	PEP (abc)
v_8	PEP (abc)	$\rightarrow$	Pyr (abc)
v_9	Pyr (abc)	$\rightarrow$	AcCoA (bc) + CO_2 (a)
Pentose Phosphate Pathway			
v_{10}	G6P (abcdef)	$\rightarrow$	P5P (bcdef) + CO_2 (a) + 2 NADPH
v_{11}	P5P (abcde) + P5P (fghij)	$\leftrightarrow$	S7P (abfghij) + GAP (cde)
v_{12}	S7P (abcdefg) + GAP (hij)	$\leftrightarrow$	F6P (abchij) + E4P (defg)
v_{13}	P5P (abcde) + E4P (fghi)	$\leftrightarrow$	F6P (abfghi) + GAP (cde)
Citric Acid Cycle			
v_{14}	$AcCoA_{mit}$ (ab) + OAA_{mit} (cdef)	$\rightarrow$	Cit (fedbac)
v_{15}	Cit (abcdef)	$\leftrightarrow$	AKG (abcde) + CO_2 (f)
v_{16}	AKG (abcde)	$\rightarrow$	Suc (bcde) + CO_2 (a)
v_{17}	Suc (abcd)	$\leftrightarrow$	1/2 Fum (abcd) + 1/2 Fum (dcba)
v_{18}	Fum (abcd)	$\leftrightarrow$	1/2 Mal_{mit} (abcd) + 1/2 Mal_{mit} (dcba)
v_{19}	Mal_{mit} (abcd)	$\leftrightarrow$	OAA_{mit} (abcd)
Transhydrogenase Cycle			
v_{20}	Pyr (abc) + CO_2 (d)	$\rightarrow$	OAA (abcd)
v_{21}	OAA (abcd)	$\leftrightarrow$	Mal (abcd)
v_{22}	Mal (abcd)	$\rightarrow$	Pyr (abc) + CO_2 (d) + NADPH
v_{23}	Mal (abcd)	$\leftrightarrow$	Mal_{mit} (abcd)
v_{24}	OAA (abcd)	$\leftrightarrow$	PEP (abc) + CO_2 (d)
One-Carbon Metabolism			
v_{25}	MEETHF (a)	$\rightarrow$	METHF (a)
v_{26}	MEETHF (a)	$\rightarrow$	FTHF (a)

Table 8.5: A list of reactions and atom transitions within the *Y. lipolytica* network for glycolysis, the pentose phosphate pathway, the citric acid cycle, and others. Carbon atom transitions are indicated within parentheses. Irreversible and reversible reactions are indicated by the symbols $\rightarrow$ and $\leftrightarrow$, respectively.

Amino Acid Biosynthesis

v_{27}	AKG (abcde) + NADPH	→	Glu (abcde)
v_{28}	Glu (abcde)	→	Gln (abcde)
v_{29}	Glu (abcde) + 2 NADPH	→	Pro (abcde)
v_{30}	Glu (abcde) + CO_2 (f) + Gln (ghijk) + Asp (lmno) + NADPH	→	Arg (abcdef) + AKG (ghijk) + Fum (lmno)
v_{31}	OAA (abcd) + Glu (efghi)	→	Asp (abcd) + AKG (efghi)
v_{32}	Asp (abcd)	→	Asn (abcd)
v_{33}	Pyr (abc) + Glu (defgh)	→	Ala (abc) + AKG (defgh)
v_{34}	3PG (abc) + Glu (defgh)	→	Ser (abc) + AKG (defgh)
v_{35}	Ser (abc)	↔	Gly (ab) + MEETHF (c)
v_{36}	Gly (ab)	↔	CO_2 (a) + MEETHF (b)
v_{37}	Thr (abcd)	→	Gly (ab) + $AcCoA_{mit}$ (cd)
v_{38}	Ser (abc) + 4 NADPH	→	Cys (abc)
v_{39}	Asp (abcd) + Pyr (efg) + Glu (hijkl) + 2 NADPH	→	Lys (abcdgf) + CO_2 (e) + AKG (hijkl)
v_{40}	Asp (abcd) + 2 NADPH	→	Thr (abcd)
v_{41}	Asp (abcd) + METHF (e) + Cys (fgh) + 2 NADPH	→	Met (abcde) + Pyr (fgh)
v_{42}	Pyr (abc) + Pyr (def) + Glu (ghijk) + NADPH	→	Val (abcef) + CO_2 (d) + AKG (ghijk)
v_{43}	$AcCoA_{mit}$ (ab) + Pyr (cde) + Pyr (fgh) + Glu (ijklm) + NADPH	→	Leu (abdghe) + CO_2 (c) + CO_2 (f) + AKG (ijklm)
v_{44}	Thr (abcd) + Pyr (efg) + Glu (hijkl)	→	Ile (abfcdg) + CO_2 (e) + AKG (hijkl)
v_{45}	PEP (abc) + PEP (def) + E4P (ghij) + Glu (klmno) + NADPH	→	Phe (abcefghij) + CO_2 (d) + AKG (klmno)
v_{46}	PEP (abc) + PEP (def) + E4P (ghij) + Glu (klmno) + NADPH	→	Tyr (abcefghij) + CO_2 (d) + AKG (klmno)
v_{47}	Ser (abc) + P5P (defgh) + PEP (ijk) + E4P (lmno) + PEP (pqr) + Gln (stuvw) + NADPH	→	Trp (abcedklmnoj) + CO_2 (i) + GAP (fgh) + Pyr (pqr) + Glu (stuvw)
v_{48}	P5P (abcde) + FTHF (f) + Gln (ghijk) + Asp (lmno)	→	His (edcbaf) + AKG (ghijk) + Fum (lmno)

Product Formation

v_{49}	CO_2 (a)	→	$CO_{2,ext}$ (a)
v_{50}	Cit (abcdef)	→	Cit_{ext} (abcdef)
v_{51}	Glyc (abc)	→	$Glyc_{ext}$ (abc)
v_{52}	Cit (abcdef)	→	AcCoA (ed) + OAA (fcba)
v_{53}	Glyc + 24 AcCoA	→	$C16_{tag}$ (ab)
v_{54}	Glyc + 27 AcCoA	→	$C18_{tag}$ (ab)
v_{55}	0.607 Ala + 0.723 Arg + 0.819 Asp + 0.819 Asn + 0.116 Cys + 0.865 Glu + 0.865 Gln + 0.723 Gly + 1.174 His + 0.671 Ile + 1.020 Leu + 0.400 Lys + 0.607 Met + 0.516 Phe + 0.555 Pro + 0.671 Ser + 0.181 Thr + 0.516 Trp + 0.839 Tyr + 0.219 Val + 0.591 G6P + 0.205 F6P + 2.175 P5P + 0.372 GAP + 1.786 3PG + 0.147 PEP + 0.239 Pyr + 0.251 AKG + 0.981 OAA + 1.278 MEETHF	→	100 Biomass

Table 8.6: A list of reactions and atom transitions within the *Y. lipolytica* network for amino acid biosynthesis and product formation. Carbon atom transitions are indicated within parentheses. Irreversible and reversible reactions are indicated by the symbols → and ↔, respectively.

Experiment	L1	L2	H1	H2
Aeration	Low	Low	High	High
Growth	Linear	Stationary	Linear	Stationary
SSE lower bound	1686	1582	1547	1475
SSE actual	1787	1690	1715	1486
SSE upper bound	1922	1810	1773	1696
Measurements	1872	1764	1728	1654
Free parameters	70	70	70	70
Degrees of freedom	1802	1694	1658	1584

Table 8.7: Fits for each of the five *Yarrowia* experiments. A statistically significant fit was obtained for every set of conditions using the same model network.

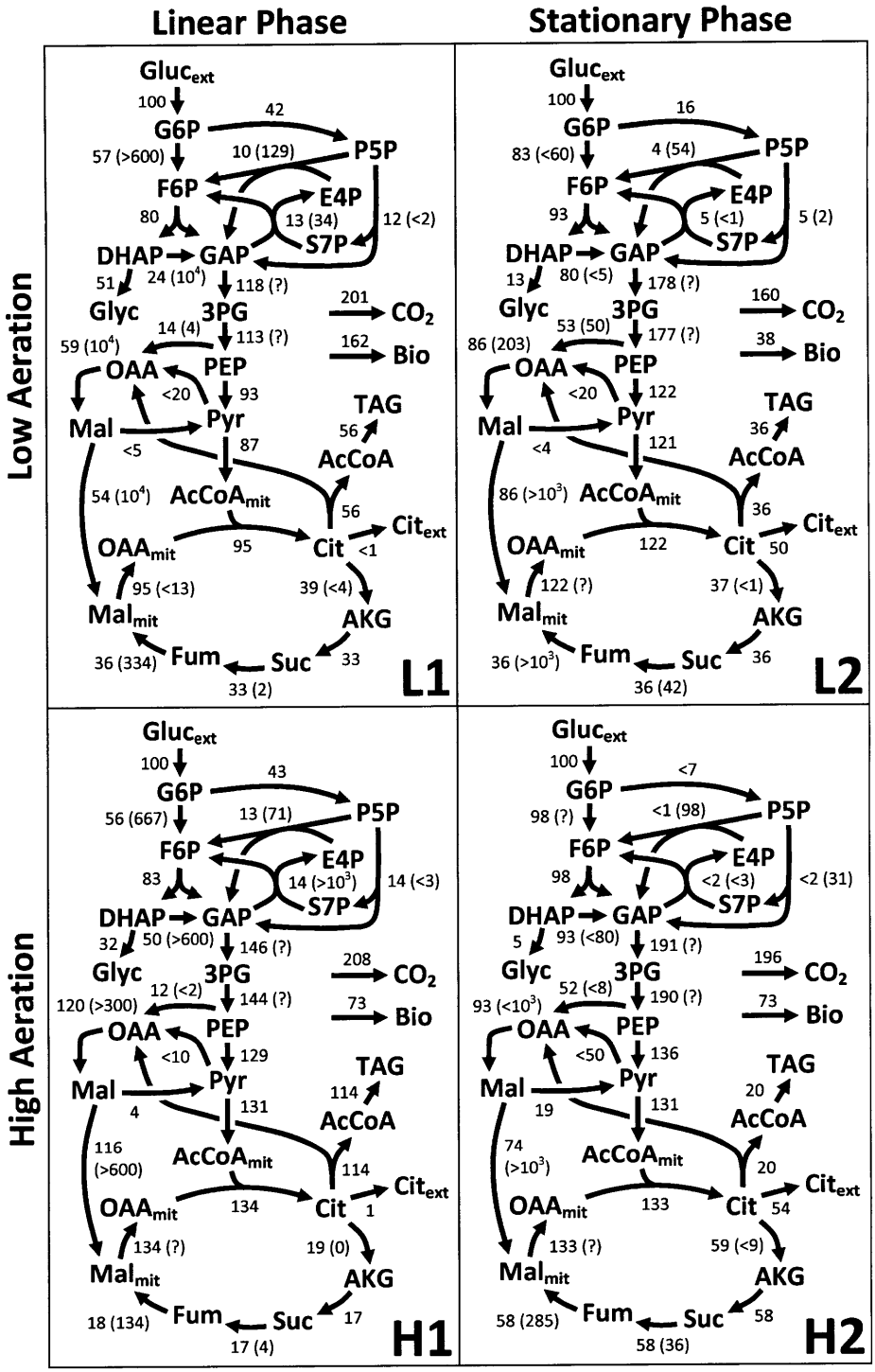

Figure 8-7: Flux distributions for *Y. lipolytica* during the late linear and stationary growth phases and at low and high aeration rates. Values are normalized to a glucose input flux of 100. Undetermined values are indicated with question marks. Net and exchange fluxes are listed with and without parentheses, respectively. Complete sets of estimated flux values, lower bounds, and upper bounds can be found in Appendix E in Tables E.9 through E.12.

8.8 Discussion

A review of the flux estimation reveals that we can place confidence in the results. For almost all net fluxes, the 95% confidence intervals fell between ±5% and ±10% of the flux values. Exchange fluxes, as usual, were more difficult to resolve. However, we were able to calculate reasonable confidence intervals for some, including isocitrate dehydrogenase exchange (Cit ↔ AKG), fumarate hydratase exchange (Fum ↔ Mal), transketolase exchanges (2 P5P ↔ S7P + GAP and P5P + E4P ↔ F6P + GAP), transaldolase exchange (S7P + GAP ↔ F6P + E4P), and glycine hydroxymethyltransferase exchange (Ser ↔ Gly). These tightly defined net and exchange fluxes were possible because of our experimental design. The large number of isotopic measurements conducted over multiple time scales together with the comprehensive measurement of extracellular carbon fluxes are the major reasons for the overall precision we observe in these fluxes. These narrow confidence intervals are especially important in this comparative study so that differences in metabolic states are detectable.

We can compare linear-phase (L1 and H1) and stationary-phase (L2 and H2) fluxes to identify key shifts in metabolism that occur as cells transition out of growth. We observe generally higher net and exchange fluxes in the TCA cycle during stationary phase, most likely because less material needs to be diverted out into amino acid biosynthesis. We also see lower carbon dioxide production in the stationary phase, which corresponds to a decreased reliance on the pentose phosphate pathway (which in turn we will discussed in greater detail in the following paragraphs). Flux through lower glycolysis and into the upper TCA cycle remains fairly consistent from the linear to stationary growth phases. In fact, regardless of the cells' growth, citrate seems to be a fairly important and central node in *Y. lipolytica* metabolism. During times of high lipid production, carbon flows through citrate and is then converted to acetyl-CoA (by cytosolic citric lyase) and from there enters fatty acid biosynthesis. During times of low growth and low lipid production, large amounts of carbon are still routed through citrate and then secreted extracellularly.

The most striking difference between linear and stationary metabolism is the need (or lack thereof) for the cofactor NADPH. *Y. lipolytica* cells in the linear growth phase are producing large amounts of amino and fatty acids, both of which require NADPH. Stationary cells, meanwhile, build insignificant amounts of these byproducts and use much less NADPH. The metabolic demands of growing *Y. lipolytica* for NADPH are manifested in the pentose phosphate pathway. Two molecules of NADPH are produced for every molecule of glucose-6-phosphate converted into ribulose. (The specific enzymes responsible for NADPH production are glucose-6-phosphate 1-dehydrogenase and 6-phosphogluconate dehydrogenase.) Metabolism in both the L1 and H1 experiments routed 42 and 43% of glucose entering the system into the pentose phosphate pathway to generate this valuable NADPH. (The remaining glucose was sent to glycolysis.) However, in the L2 and H2 experiments, only 16 and 7% of glucose was converted into ribulose. Pentose phosphate exchange rates were generally higher during linear phase, also indicating an increased level of enzymatic activity in this pathway during growth.

The transhydrogenase flux has been proposed as an important and primary source of NADPH for fatty acid synthesis [118, 119]. Figure 8-8 diagrams this network in the context of lipid biosynthesis. In this cycle, pyruvate is converted to oxaloacetate and oxaloacetate to malate (by the enzymes oxaloacetate carboxylyase and malate dehydrogenase, respectively). The former reaction consumes NADH. Finally, malic enzyme catalyzes the reaction of malate to pyruvate, producing NADPH and completing the cycle. The transhydrogenase cycle was included our model network for *Y. lipolytica*, but interestingly, was only marginally active across all scenarios. Further inspection reveals that malic enzyme activity is not ubiquitous across all oleaginous microorganisms and may be absent in some oleaginous yeasts [118].

Another flux analyis study of several species of yeast also validates our observations of high pentose phosphate pathway activity and low malic enzyme activity. Blank *et al* has reported a relative pentose phosphate flux of 40% in *Yarrowia lipolytica*, and has found that a variety of yeasts, including *S. exiguus*, *S. cerevisiae*, *K. thermotolerans*, *C. tropicalis*, and *P. angusta* all rely primarily on the pentose phosphate

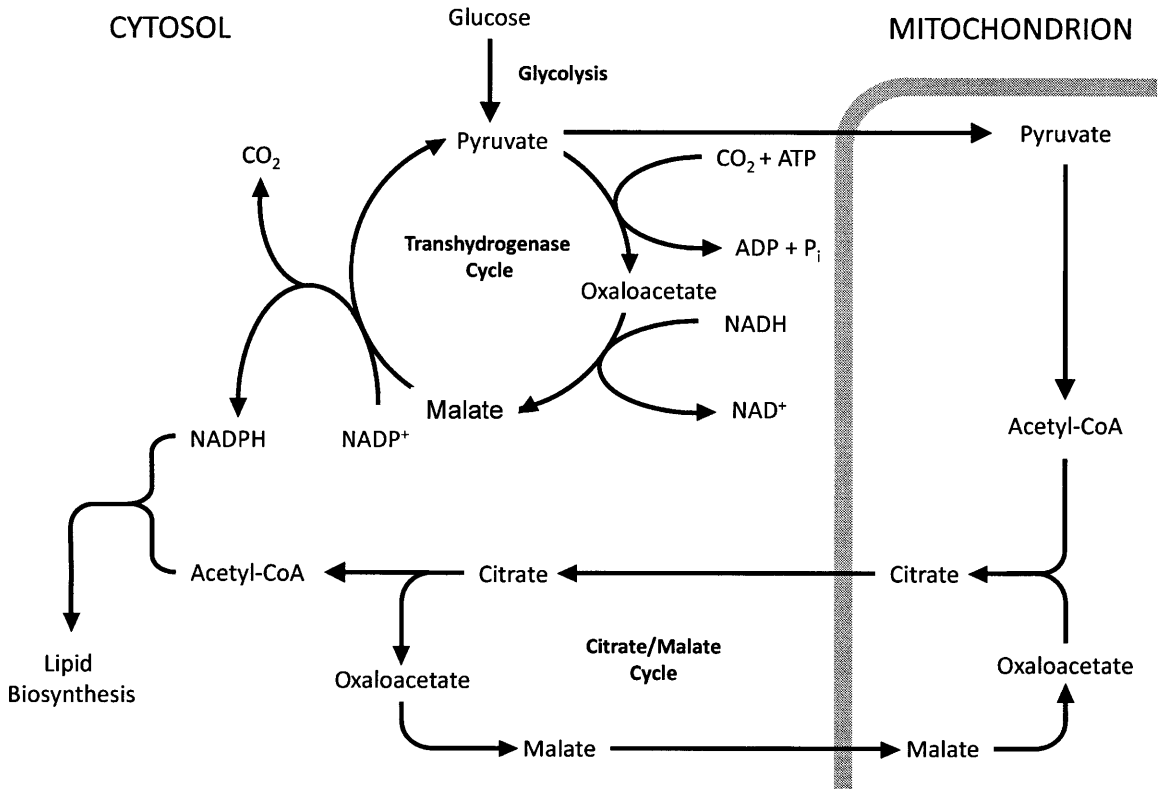

Figure 8-8: The transhydrogenase cycle and other key pathways in lipid accumulation. As part of the transhydrogenase cycle, cytosolic malic enzyme converts malate to pyruvate, generating NADPH in the process that can be used in lipid biosynthesis. This figure is adapted from Ratledge and Wynn 2002 [119].

pathway for NADPH production (from 60% to as much as 90% depending on the specific organism) [21]. They also note that the flux through malic enzyme as an alternative for NADPH was found to be low in all yeasts investigated. These findings coupled with the results of our flux analysis suggest that if NADPH is limiting lipid synthesis, the currently dormant transhydrogenase cycle might prove to be a fruitful target for metabolic engineering to increase lipid flux.

To understand the impact of increased aeration on *Yarrowia* metabolism, and especially on fatty acid synthesis, we can compare and contrast experiments L1 and H1. Since measured fluxes are treated not as fixed values but as measurements alongside metabolite mass isotopomers, estimated fluxes are not guaranteed to perfectly reflect the originally measured values. However, with respect to our lipid synthesis fluxes, we see our original trend not only holds true but is actually accentuated with the consideration of our isotopic labeling experiments. Instead of the originally calculated 40% increase, we now observe a doubling in lipid production at greater aeration. Both extracellular and isotopic data agree that high aeration confers a significant boost to fatty acid synthesis.

Cells under high aeration did not demonstrate an increased need for NADPH because of high fatty acid production. Amino acid synthesis in these conditions is much reduced, freeing up NADPH and allowing it to be utilized elsewhere. For this reason, we see that the pentose phosphate pathway fluxes in L1 and H1 remain almost constant. It appears that cells in L1 are growing in number (mass per cell remains constant) while cells in L1 are growing in size (cell count per volume remains constant).

We also notice a relatively large glycerol output flux in the linear, low-aeration experiment: 26% of incoming glucose exits the system as glycerol in low aeration, but this drops 16% with high aeration. No sizeable amounts of extracellular glycerol were measured, so it is most likely that this represents carbon exiting the network as carbohydrate storage (e.g., trehalose or glycogen). Because the model lacks a free G6P or F6P output flux, these exiting carbohydrates are assigned to glycerol output. *Y. lipolytica* is a naturally dimorphic yeast forming yeast cells, pseudohyphae, and

septate hyphae under different conditions [16]. Vega and Domínguez have compared the compositions of yeast and hyphal cells and have shown that hyphal walls exhibit a higher content of aminosugars [150]. Other studies have indicated that starvation conditions induce hyphal growth of *Y. lipolytica* [66]. These findings suggest increasing aeration could improve growth conditions (by reducing the fraction of hypoxic cells in the culture) which in turn would lead to less hyphae and more yeast-like cells and lower amounts of cell-wall carbohydrates.

We have observed that lipid production in a fed-batch bioreactor culture of *Y. lipolytica* is highest during the late growth and early stationary phase. We have shown that lipid production can be significantly increased by increasing aeration. Cell metabolism was studied during different growth stages and aeration rates to better understand the reasons behind this upshift in production. We saw that the pentose phosphate pathway was very active at times of high amino acid and fatty acid production and served as the primary source of NADPH for these anabolic activities. Meanwhile, the transhydrogenase cycle was negligibly active and played only a minor role in cofactor synthesis. Cells in a higly aerated environment diverted less resources to carbohydrate and amino acid synthesis, indicating that (1) yeast-like cell morphology is dominant in these conditions and (2) cells were growing in size (due to triacylglyceride accumulation) instead of in number.

Chapter 9

Recommendations for Future Research

9.1 Dynamic Metabolic Flux Analysis

Current methods of flux analysis allow for flux estimation in systems where either isotopic labeling or metabolism is at a non-steady state, but not both. In the former case, NMFA is used to generate fluxes, as discussed in this thesis. In the latter case, metabolism changes at a significantly slower rate than isotopic labeling, such that

$$\max\left(\mathbf{V}^{-1}\cdot\mathbf{A}\cdot\mathbf{C}\right) \ll t_{\text{exp}} \ll \min\left(\mathbf{R}\cdot\left|\frac{d\mathbf{R}}{dt}\right|^{-1}\right) \tag{9.1}$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{V} & 0 \\ 0 & \mathbf{C} \end{bmatrix} \tag{9.2}$$

where $\mathbf{V}$ is a diagonal matrix of network fluxes, $\mathbf{A}$ is an adjacency matrix in which a nonzero entry $a(i,j)$ indicates that the ith flux leads to the jth metabolite pool, $\mathbf{C}$ is a diagonal matrix of metabolite concentrations, t_{exp} is the duration of the MFA experiment, $\mathbf{R}$ is a diagonal matrix comprising fluxes and concentrations, and t is time.

The left-hand term of Equation 9.1 represents the maximum of the cell's metabo-

lite pools' residence times. This term serves as an approximate measure of the time required for the network to arrive at an isotopic steady state. The right-hand term represents the minimum time scale at which cellular metabolism (i.e., fluxes and concentrations) change significantly. If a window can be found between these two bounds, and if MFA can be feasibly conducted at a time scale within that window, steady-state flux analysis will serve to accurately measure fluxes even at a metabolic non-steady state. Such an assumption was successfully used previously in studies of 1,3-propanediol production by *E. coli* and hepatic lipoapoptosis [12, 105].

The conditions of Equation 9.1 do not hold true for many interesting, metabolically dynamic systems, leaving them currently unacessible by flux analysis. For instance, rapid cellular responses to sudden changes in environment (such as a substrate bolus or a temperature shift) cannot be captured using current modeling tools. New methods for metabolically dynamic flux analysis (DMFA) must be developed and applied.

Some introductory work has been done regarding DMFA. Experimental GC/MS data has been used to calculated fluxes in the linear threonine synthesis pathway [18]. A theoretical study of a metabolically dynamic small example network showed that ^{13}C labeling data could lead to more precise parameter determination if the kinetic rate laws of the different fluxes were already known [154]. However, beyond these efforts, very little has been done in this area. If robust, systematic experimental and computational means could be developed for the quantification of fluxes in a metabolically non-steady network, the impact could be enormous. Methods not requiring a prior knowledge of kinetic rate laws would be especially powerful.

9.2 MetranCL Development

MetranCL is a computational framework for MFA and NMFA simulation and estimation and is discussed in Chapter 2 and Appendix B. MetranCL is a useful and flexible research tool; however, there are many potential improvements that could be added to create an even more powerful version of this software:

- **User accesibility:** MetranCL currently runs in (and requires users to own) Matlab. Users also must conduct all MetranCL operations within Matlab's command-line environment. These requirements pose a relatively high barrier and will diminish the adoption of this software package in the metabolic engineering community and beyond. A simple, approachable, and intuitive graphical user interface (GUI) needs to be developed. Care should be taken in the design and implementation so that activities such as network creation and metabolite definition remain straightforward and efficient. Additionally, MetranCL should be built to function as a stand-alone program so that users are not forced to purchase Matlab.

- **Language:** MetranCL is written completely in Matlab code. Software written in the Matlab language are relatively simple to write and understand, especially when they deal heavily with matrix mathematics (as is the case in flux analysis). Unfortunately, one of the primary disadvantages of coding in Matlab is that such programs typically run much more slowly than lower-level languages. Because flux estimation and confidence interval calculation can still behave as a bottleneck in flux analysis, especially when working with large networks and/or NMFA, rewriting the most computationally intensive portions of MetranCL in a faster language (such as C++) could bring large benefits.

- **Parallelization:** The current implementation of MetranCL has some very basic parallelization capabilities. Most notably, confidence interval calculations can be parallelized parameter by parameter across different processors of a cluster, since each calculation is independent of the others. However, the current software requires a precisely configured cluster and the parallel calculations are somewhat laborious to set up. A more polished implementation of parallelization that is adaptable to a broader range of computing clusters would make MetranCL a significantly more appealing and powerful product.

9.3 Network Sensitivity Analysis

While significant efforts have been made to understand the sensitivity of flux estimation to experimental and modeling design parameters such as measurements and tracers [32, 98], researches have devoted less study and time to understanding the effect of model reaction networks on fluxes. Network sensitivity can be difficult to study due to the discrete nature of the problem. The extremely large sizes of true metabolic networks also complicate analysis. For example, a genome-scale model of *S. cerevisiae* has been constructed that contains 1149 reactions [43]. This large set of potential reactions poses several problems as we construct networks to model isotopic labeling:

- Since our measurement set is limited by practical concerns, if too many of these reactions are included in the MFA network model, the system will be underdetermined and precisionless (and therefore useless) flux estimates will be obtained.
- If too few reactions are used, the system will be greatly overdetermined and a fit will be impossible.
- If we mistakenly omit key reactions from the the model and replace them with others that in actuality are insignificant in metabolism, we could create a network within which a precise but inaccurate flux distribution can be estimated, returning a "false positive".

Methods for analyzing networks in the context of specific flux estimations could prove very helpful in combating some of these potential issues. Such tools might also be able analyze an MFA experiment for a given reaction network and then search a database of additional reactions and make suggestions of network changes that would lead to a better fit of simulated and measured data.

Appendix A

Block-Decoupled EMUs and Matrices

A.1 State Matrices

The state matrix $\mathbf{X}_n$ is defined in Chapter 2 and represents elementary metabolite unit labeling within the nth block of a decomposed network. Rows correspond to different EMUs in the block, while columns correspond to different isotopic masses. We list each of the eight $\mathbf{X}$ state matrices in Equations A.1 through A.8 for the simple example network described in Figures 2-1 and 2-2. The notation $D_{12,M+1}$ refers to the mass isotopomer fraction M+1 for the EMU composed of atoms 1 and 2 of the metabolite D.

$$\mathbf{X}_1 = \begin{bmatrix} B_{2,M+0} & B_{2,M+1} \\ C_{2,M+0} & C_{2,M+1} \\ D_{1,M+0} & D_{1,M+1} \end{bmatrix} \tag{A.1}$$

$$\mathbf{X}_2 = \begin{bmatrix} E_{1,M+0} & E_{1,M+1} \end{bmatrix} \tag{A.2}$$

$$\mathbf{X}_3 = \begin{bmatrix} B_{1,M+0} & B_{1,M+1} \\ C_{3,M+0} & C_{3,M+1} \\ D_{2,M+0} & D_{2,M+1} \end{bmatrix} \tag{A.3}$$

$$\mathbf{X}_4 = \begin{bmatrix} C_{1,M+0} & B_{1,M+1} \\ F_{1,M+0} & C_{1,M+1} \end{bmatrix} \tag{A.4}$$

$$\mathbf{X}_5 = \begin{bmatrix} B_{12,M+0} & B_{12,M+1} & B_{12,M+2} \\ C_{23,M+0} & C_{23,M+1} & C_{23,M+2} \\ D_{12,M+0} & D_{12,M+1} & D_{12,M+2} \end{bmatrix} \tag{A.5}$$

$$\mathbf{X}_6 = \begin{bmatrix} E_{12,M+0} & E_{12,M+1} & E_{12,M+2} \end{bmatrix} \tag{A.6}$$

$$\mathbf{X}_7 = \begin{bmatrix} C_{234,M+0} & C_{234,M+1} & C_{234,M+2} & C_{234,M+3} \\ D_{123,M+0} & D_{123,M+1} & D_{123,M+2} & D_{123,M+3} \end{bmatrix} \tag{A.7}$$

$$\mathbf{X}_8 = \begin{bmatrix} C_{1234,M+0} & C_{1234,M+1} & C_{1234,M+2} & C_{1234,M+3} & C_{1234,M+4} \end{bmatrix} \tag{A.8}$$

The state matrix $\mathbf{Y}_n$ is defined in Chapter 2 and represents elementary metabolite unit labeling of inputs to the nth block of a decomposed network. Rows correspond to different EMUs in the block, while columns correspond to different isotopic masses. We list each of the eight $\mathbf{Y}$ state matrices in Equations A.9 through A.16. The notation $B_2E_{1,M+2}$ refers to the mass isotopomer fraction M+2 for the composite EMU formed from the B_2 and E_1 EMUs. The composite MID is the convolution (or Cauchy product) of the MIDs of the component EMUs.

$$\mathbf{Y}_1 = \begin{bmatrix} A_{2,M+0} & A_{2,M+1} \end{bmatrix} \tag{A.9}$$

$$\mathbf{Y}_2 = \begin{bmatrix} D_{1,M+0} & D_{1,M+1} \end{bmatrix} \tag{A.10}$$

$$\mathbf{Y}_3 = \begin{bmatrix} A_{1,M+0} & A_{1,M+1} \\ E_{1,M+0} & E_{1,M+1} \end{bmatrix} \tag{A.11}$$

$$\mathbf{Y}_4 = \begin{bmatrix} B_{1,M+0} & B_{1,M+1} \end{bmatrix} \tag{A.12}$$

$$\mathbf{Y}_5 = \begin{bmatrix} B_2E_{1,M+0} & B_2E_{1,M+1} & B_2E_{1,M+2} \end{bmatrix} \tag{A.13}$$

$$\mathbf{Y}_6 = \begin{bmatrix} D_{12,M+0} & D_{12,M+1} & D_{12,M+2} \end{bmatrix} \tag{A.14}$$

$$\mathbf{Y}_7 = \begin{bmatrix} B_2E_{12,M+0} & B_2E_{12,M+1} & B_2E_{12,M+2} & B_2E_{12,M+3} \\ B_{12}J_{1,M+0} & B_{12}J_{1,M+1} & B_{12}J_{1,M+2} & B_{12}J_{1,M+3} \end{bmatrix} \tag{A.15}$$

$$\mathbf{Y}_8 = \begin{bmatrix} D_{123}F_{1,M+0} & D_{123}F_{1,M+1} & D_{123}F_{1,M+2} & D_{123}F_{1,M+3} & D_{123}F_{1,M+4} \\ D_{123}F_{1,M+0} & D_{123}F_{1,M+1} & D_{123}F_{1,M+2} & D_{123}F_{1,M+3} & D_{123}F_{1,M+4} \end{bmatrix} \tag{A.16}$$

A.2 System Matrices

The system matrix $\mathbf{A}_n$ is mathematically defined in Equation 2.2. The elements of this matrix are linear combinations of the fluxes within a particular block, and the matrix as a whole describes the stoichiometric relationships between the elementary metabolite units of that block. In Equations A.17 through A.24 below, we list each of the eight $\mathbf{A}$ system matrices for the simple example network.

$$\mathbf{A}_1 = \begin{bmatrix} -v_2 - v_{5f} & 0 & v_{5b} \\ v_2 & -v_{3f} & v_{3b} \\ v_{5f} & v_{3f} & -v_{3b} - v_4 - v_{5b} \end{bmatrix} \tag{A.17}$$

$$\mathbf{A}_2 = \begin{bmatrix} -v_2 \end{bmatrix} \tag{A.18}$$

$$\mathbf{A}_3 = \begin{bmatrix} -v_2 - v_{5f} & 0 & v_{5b} \\ 0 & -v_{3f} & v_{3b} \\ v_{5f} & v_{3f} & -v_{3b} - v_4 - v_{5b} \end{bmatrix} \tag{A.19}$$

$$\mathbf{A}_4 = \begin{bmatrix} -v_{3f} & v_{3b} \\ v_{3f} & -v_6 \end{bmatrix} \tag{A.20}$$

$$\mathbf{A}_5 = \begin{bmatrix} -v_2 - v_{5f} & 0 & v_{5b} \\ 0 & -v_{3f} & v_{3b} \\ v_{5f} & v_{3f} & -v_{3b} - v_4 - v_{5b} \end{bmatrix} \tag{A.21}$$

$$\mathbf{A}_6 = \begin{bmatrix} -v_2 \end{bmatrix} \tag{A.22}$$

$$\mathbf{A}_7 = \begin{bmatrix} -v_{3f} & v_{3b} \\ v_{3f} & -v_{3b} - v_4 - v_{5b} \end{bmatrix} \tag{A.23}$$

$$\mathbf{A}_8 = \begin{bmatrix} -v_{3f} \end{bmatrix} \tag{A.24}$$

The system matrix $\mathbf{B}_n$ is defined in Equation 2.3. Like $\mathbf{A}_n$, the elements are linear combinations of fluxes; however, in this case, the matrix describes fluxes into the block (instead of fluxes within the block). In Equations A.25 through A.32, we list each of the eight $\mathbf{B}$ system matrices for the simple example network.

$$\mathbf{B}_1 = \begin{bmatrix} v_1 \\ 0 \\ 0 \end{bmatrix} \tag{A.25}$$

$$\mathbf{B}_2 = \begin{bmatrix} v_4 \end{bmatrix} \tag{A.26}$$

$$\mathbf{B}_3 = \begin{bmatrix} v_1 & 0 \\ 0 & v_2 \\ 0 & 0 \end{bmatrix} \tag{A.27}$$

$$\mathbf{B}_4 = \begin{bmatrix} v_2 \\ 0 \\ 0 \end{bmatrix} \tag{A.28}$$

$$\mathbf{B}_5 = \begin{bmatrix} 0 \\ v_2 \\ 0 \end{bmatrix} \tag{A.29}$$

$$\mathbf{B}_6 = \begin{bmatrix} v_4 \end{bmatrix} \tag{A.30}$$

$$\mathbf{B}_7 = \begin{bmatrix} v_2 & 0 \\ 0 & v_{5f} \end{bmatrix} \tag{A.31}$$

$$\mathbf{B}_8 = \begin{bmatrix} v_2 & v_{3b} \end{bmatrix} \tag{A.32}$$

The system matrix $\mathbf{C}_n$ is explained in Chapter 2 in Equation 2.3. The diagonal is composed of the metabolite concentrations that correspond to the EMUs in $\mathbf{X}_n$. In Equations A.33 through A.40, we list each of the eight $\mathbf{C}$ system matrices for the simple example network. The notation c_{B} refers to the concentration of metabolite B.

$$\mathbf{C}_1 = \begin{bmatrix} c_{\mathrm{B}} & 0 & 0 \\ 0 & c_{\mathrm{C}} & 0 \\ 0 & 0 & c_{\mathrm{D}} \end{bmatrix} \tag{A.33}$$

$$\mathbf{C}_2 = \begin{bmatrix} c_{\mathrm{E}} \end{bmatrix} \tag{A.34}$$

$$\mathbf{C}_3 = \begin{bmatrix} c_B & 0 & 0 \\ 0 & c_C & 0 \\ 0 & 0 & c_D \end{bmatrix} \tag{A.35}$$

$$\mathbf{C}_4 = \begin{bmatrix} c_C & 0 \\ 0 & c_F \end{bmatrix} \tag{A.36}$$

$$\mathbf{C}_5 = \begin{bmatrix} c_B & 0 & 0 \\ 0 & c_C & 0 \\ 0 & 0 & c_D \end{bmatrix} \tag{A.37}$$

$$\mathbf{C}_6 = \begin{bmatrix} c_E \end{bmatrix} \tag{A.38}$$

$$\mathbf{C}_7 = \begin{bmatrix} c_C & 0 \\ 0 & c_D \end{bmatrix} \tag{A.39}$$

$$\mathbf{C}_8 = \begin{bmatrix} c_C \end{bmatrix} \tag{A.40}$$

A.3 *E. coli* Block Decoupling

The large *E. coli* network described in Tables 3.1 and 3.2 can be greatly simplified via EMU decomposition and block decoupling. We processed the network accordingly, and found that the overall system of equations could be reduced by 94% and that the largest single subsystem could be reduced by 85% (Figure 3.3. For the sake of completeness and illustration of our method, we have included the decoupled blocks of EMUs for the network in the following Tables A.1, A.2, and A.3.

Block 1

AKG_1, AKG_2, AKG_3, AKG_4, AKG_5, Ac_1, Ac_2, $AcCoA_1$, $AcCoA_2$, Asp_1, Asp_2, Asp_3, Asp_4, $CO2_1$, Cit_1, Cit_2, Cit_3, Cit_4, Cit_5, Cit_6, Cys_1, Cys_2, Cys_3, $DHAP_1$, $DHAP_2$, $DHAP_3$, $E4P_2$, $E4P_3$, $E4P_4$, $F6P_4$, $F6P_5$, $F6P_6$, FBP_1, FBP_2, FBP_3, FBP_4, FBP_5, FBP_6, Fum_3, Fum_4, $G6P_4$, $G6P_5$, $G6P_6$, GAP_1, GAP_2, GAP_3, Gln_1, Gln_2, Gln_3, Gln_4, Gln_5, Glu_1, Glu_2, Glu_3, Glu_4, Glu_5, Gly_1, Gly_2, $Glyc3P_1$, $Glyc3P_2$, $Glyc3P_3$, $ICit_1$, $ICit_2$, $ICit_3$, $ICit_4$, $ICit_5$, $ICit_6$, $KDPG_4$, $KDPG_5$, $KDPG_6$, LL-DAP_7, $MEETHF_1$, Mal_1, Mal_2, Mal_3, Mal_4, OAA_1, OAA_2, OAA_3, OAA_4, PEP_1, PEP_2, PEP_3, $3PG_1$, $3PG_2$, $3PG_3$, $6PG_4$, $6PG_5$, $6PG_6$, Pyr_1, Pyr_2, Pyr_3, $R5P_3$, $R5P_4$, $R5P_5$, $Ru5P_3$, $Ru5P_4$, $Ru5P_5$, $S7P_5$, $S7P_6$, $S7P_7$, Ser_1, Ser_2, Ser_3, Suc_3, Suc_4, $SucCoA_1$, $SucCoA_2$, $SucCoA_3$, $SucCoA_4$, Thr_1, Thr_2, Thr_3, Thr_4, $X5P_3$, $X5P_4$, $X5P_5$

Block 2

$E4P_1$, $EC2_1$, $EC2_2$, $EC3_1$, $EC3_2$, $EC3_3$, $F6P_1$, $F6P_2$, $F6P_3$, $G6P_1$, $G6P_2$, $G6P_3$, $PG6_2$, $PG6_3$, $R5P_1$, $R5P_2$, $Ru5P_1$, $Ru5P_2$, $S7P_1$, $S7P_2$, $S7P_3$, $S7P_4$, $X5P_1$, $X5P_2$

Block 3	**Block 4**	**Block 5**	**Block 6**	**Block 7**
$KDPG_1$	$KDPG_3$	$KDPG_3$	$METHF_1$	$6PG_1$

Block 8

AKG_{45}, Ac_{12}, $AcCoA_{12}$, Asp_{12}, Asp_{34}, Cit_{45}, $DHAP_{12}$, $E4P_{23}$, $F6P_{45}$, FBP_{23}, FBP_{45}, Fum_{34}, $G6P_{45}$, GAP_{12}, Gln_{45}, Glu_{45}, $Glyc3P_{12}$, $ICit_{45}$, $KDPG_{45}$, Mal_{12}, Mal_{34}, OAA_{12}, OAA_{34}, PEP_{12}, $3PG_{12}$, $6PG_{45}$, $R5P_{34}$, $Ru5P_{34}$, $S7P_{56}$, Suc_{34}, $SucCoA_{12}$, $SucCoA_{34}$, Thr_{34}, $X5P_{34}$

Block 9

Asp_{23}, $DHAP_{23}$, $E4P_{34}$, $F6P_{56}$, FBP_{12}, FBP_{56}, Fum_{23}, $G6P_{56}$, GAP_{23}, $Glyc3P_{23}$, $KDPG_{56}$, Mal_{23}, OAA_{23}, PEP_{23}, $3PG_{23}$, $6PG_{56}$, $R5P_{45}$, $Ru5P_{45}$, $S7P_{67}$, Suc_{23}, $SucCoA_{23}$, $X5P_{45}$

Table A.1: Block decoupling in the large *E. coli* network for blocks 1 through 9. The notation AKG_{45} refers to the EMU consisting of AKG atoms 4 and 5.

Block 10

$EC2_{12}$, $EC3_{12}$, $F6P_{12}$, $G6P_{12}$, $R5P_{12}$, $Ru5P_{12}$, $S7P_{12}$, $S7P_{34}$, $X5P_{12}$

Block 11

AKG_{34}, Cit_{34}, Gln_{34}, Glu_{34}, $ICit_{34}$

Block 12

AKG_{23}, Cit_{23}, Gln_{23}, Glu_{23}, $ICit_{23}$

Block 13

$EC3_{23}$, $F6P_{23}$, $G6P_{23}$, $S7P_{23}$

Block 14

Gly_{12}, Ser_{12}

Block 15

$6PG_{23}$

Block 16

Ala_{23}

Block 17

Cys_{23}

Block 18

$KDPG_{23}$

Block 19

Pyr_{23}

Block 20

Ser_{23}

Block 21

Phe_{12}

Block 22

Thr_{12}

Block 23

Tyr_{12}

Block 24

Asp_{123}, Asp_{234}, $DHAP_{123}$, $E4P_{234}$, $F6P_{456}$, FBP_{123}, FBP_{456}, Fum_{234}, $G6P_{456}$, GAP_{123}, $Glyc3P_{123}$, $KDPG_{456}$, Mal_{123}, Mal_{234}, OAA_{123}, OAA_{234}, PEP_{123}, $3PG_{123}$, $6PG_{456}$, $R5P_{345}$, $Ru5P_{345}$, $S7P_{567}$, Suc_{234}, $SucCoA_{123}$, $SucCoA_{234}$, $X5P_{345}$

Block 25

AKG_{234}, Cit_{234}, Gln_{234}, Glu_{234}, $ICit_{234}$

Block 26

AKG_{345}, Cit_{345}, Gln_{345}, Glu_{345}, $ICit_{345}$

Block 27

$EC3_{123}$, $F6P_{123}$, $G6P_{123}$, $S7P_{123}$

Block 28

Ala_{123}

Block 29

Cys_{123}

Table A.2: Block decoupling in the large *E. coli* network for blocks 10 through 29. The notation AKG_{45} refers to the EMU consisting of AKG atoms 4 and 5.

Block 30	Block 31	Block 32	Block 33	Block 34
$KDPG_{123}$	$6PG_{123}$	Pyr_{123}	Ser_{123}	Thr_{234}

Block 35

$E4P_{1234}$, $F6P_{3456}$, $G6P_{3456}$, $6PG_{3456}$, $R5P_{2345}$, $Ru5P_{2345}$, $S7P_{4567}$, $X5P_{2345}$

Block 36

Asp_{1234}, Fum_{1234}, Mal_{1234}, OAA_{1234}, Suc_{1234}, $SucCoA_{1234}$

Block 37

AKG_{2345}, Cit_{2345}, Gln_{2345}, Glu_{2345}, $ICit_{2345}$

Block 38	Block 39	Block 40
Met_{2345}	Thr_{1234}	Val_{2345}

Block 41

AKG_{12345}, Cit_{12345}, Gln_{12345}, Glu_{12345}, $ICit_{12345}$

Block 42	Block 43	Block 44	Block 45
Ile_{23456}	Leu_{23456}	Met_{12345}	Val_{12345}

Block 46	Block 47	Block 48
Cit_{123456}, $ICit_{123456}$	$Phe_{23456789}$	$Phe_{123456789}$

Table A.3: Block decoupling in the large *E. coli* network for blocks 30 through 48. The notation AKG_{45} refers to the EMU consisting of AKG atoms 4 and 5.

Appendix B

MetranCL Documentation

B.1 Introduction

MetranCL is a powerful computational tool for MFA and NMFA simulation, estimation, and continuation. MetranCL requires Matlab (version 7.4.0 or newer) as well as Matlab's statistics and optimization toolboxes. A directory tree of the necessary files and folders is found in Figure B-1. This appendix first introduces each of the user-defined classes introduced by MetranCL. We next introduce and explain selected MetranCL functions. Finally, Matlab code is provided for creating an example network.

B.2 Classes

`atom` represents an atom within a metabolite and contains the following fields:

- `id` (`string`) is the atom's identification tag. It should be unique from all other atom `id`s for a particular metabolite, and is usually composed of numerical characters.
- `type` (`string`) is the atom's element symbol (e.g., `'C'`, `'H'`, or `'O'`).

`data` represents a generic piece of data and contains the following fields:

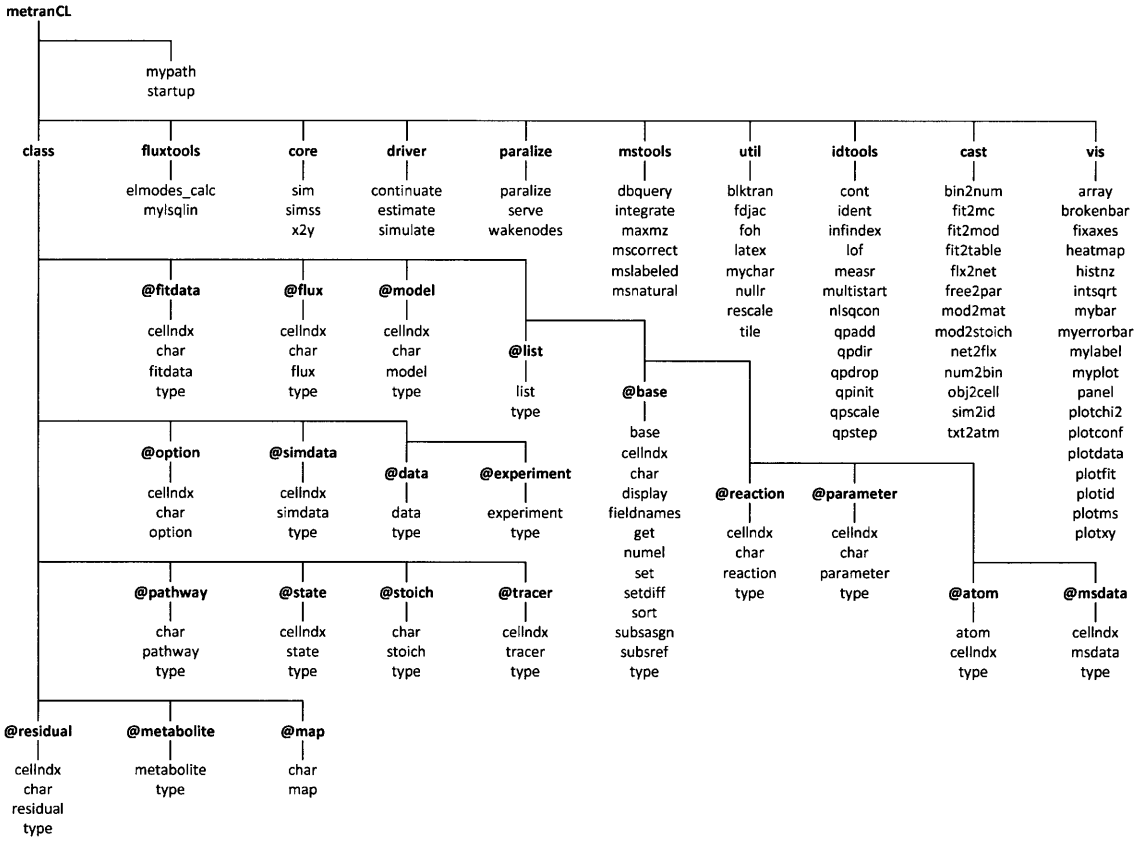

Figure B-1: Directory tree of the files and folders comprising MetranCL. Folders are written in bold font and files are written in normal font. All files are Matlab m-files (ending with a .m suffix). Lines connecting folders to lower files or folders indicate that the lower object is contained within the upper folder.

- `id` (`string`) is the data item's identification tag.
- `on` (`logical`) is `true` of `false` if the data is to be used or ignored, respectively.
- `val` (`double`) is the data value.
- `std` (`double`) is the standard error of the data.

`experiment` represents the design and measurements of a flux analysis experiment and contains the following fields:

- `id` (`string`) is the experiment's identification tag.
- `on` (`logical`) is `true` of `false` if the experiment is to be used or ignored, respectively.
- `tracer` (`tracer`) is a description of the tracer(s) used in the experiment.
- `data_flx` (`data`) is the set of all measured fluxes in the experiment.
- `data_cxn` (`data`) is the set of all measured metabolite concentrations in the experiment.
- `data_ms` (`msdata`) is the set of all mass spectrometry measurements in the experiment.
- `notes` (`string`) is a list of notes or reminders concerning the experiment.

`fitdata` holds the results for flux and concentration estimation and continuation and contains the following fields:

- `alf` (`double`) is the α-level of the continuation. The confidence level is equal to one minus the α-level.
- `dof` (`double`) is the degrees of freedom in the experiment (the number of measurements minus the number of parameters).

- `Echi2` (`double`) is a two-element array containing the lower and upper bounds of the expected sum of squared residuals.
- `chi2` (`double`) is the estimation's sum of squared residuals.
- `par` (`parameter`) contains estimated values and confidence intervals for fluxes and concentrations.
- `res` (`residual`) contains residual values for each measurement used in the estimation.

`flux` represents a unidirectional flux within a reaction in a metabolic network and contains the following fields:

- `id` (`string`) is the flux's identification tag.
- `rxn` (`string`) is the identification tag of the reaction to which the flux belongs.
- `dir` (`string`) is `'f'` or `'b'` if the flux is in the forwards or backwards directions, respectively.
- `sub` (`stoich`) describes the identity and stoichiometry of the flux's substrates.
- `prod` (`stoich`) describes the identity and stoichiometry of the flux's products.
- `fix` (`logical`) is `true` of `false` if the flux is fixed or free, respectively.
- `lb` (`double`) is the user-defined lower bound of the flux.
- `ub` (`double`) is the user-defined upper bound of the flux.
- `val` (`double`) is the user-defined value of the flux.

`list` represents a simple tag and value pair and contains the following fields:

- `id` (`string`) is the identification tag of the list item.
- `it` (any) is the value of the list item.

`map` maps information from one set of coordinates to a new set. It is usually used with atoms and contains the following fields:

- `id` (`string`) is the identification tags for the original coordinates.
- `tr` (`string`) is the identification tags for the new coordinates.

`metabolite` represents a metabolite within the reaction network and contains the following fields:

- `id` (`string`) is the metabolite's identification tag.
- `atoms` (`atom`) is an array containing the metabolite's atoms.
- `sym` (`list`) is a list item of all symmetric atoms within the metabolite. The list item's value is a map that maps each symmetric atom to its twin.
- `eqv` (`list`) is a list of all equivalent atoms within the metabolite. Each list item's value is a space-delimited string of atom identification tags within an equivalent group.

`model` holds all of the information necessary for a flux analyis experiment (i.e., the model network, the experimental set-up, and all measurement data) and contains the following fields:

- `options` (`option`) is a group of user-defined settings that govern flux analysis computation.
- `mets` (`metabolite`) defines all metabolites participating in the metabolic network.
- `states` (`state`) defines all metabolite states in the metabolic network.
- `rates` (`reaction`) defines all of the reactions in the metabolic network.
- `expts` (`experiment`) holds all measurement information (concentrations, fluxes, and/or labeling).

- `notes` (`string`) is a list of notes or reminders concerning the model.

`msdata` identifies and quantifies a particular mass spectrometry measurement and contains the following fields:

- `id` (`string`) is the measurement's identification tag.
- `on` (`logical`) is `true` of `false` if the measurement is to be used or ignored, respectively.
- `state` (`string`) is the identification tag of the metabolite state from which the measured fragment originates.
- `atoms` (`string`) is a space-delimited string of identification tags for atoms both within the metabolite fragment and defined in the model's reaction network (i.e., potentially labeled atoms).
- `more` (`string`) is the chemical formula for atoms within the metabolite fragment but not defined in the model's reaction network (i.e., unlabeled atoms). Note that atoms added by derivatization methods should be included in this formula.
- `norm` (`logical`) is the mass spectrometry measurement normalization factor.
- `time` (`double`) is the measurement time. For multiple time points, this is a vector.
- `val` (`double`) is the measurement value. For multiple time points, this is a matrix whose rows correspond to masses and columns to times.
- `std` (`double`) is the standard error of the measurement. For multiple time points, this is a matrix whose rows correspond to masses and columns to times.

`option` holds a list of user-defined settings that govern flux simulation, estimation, and continuation and contains the following fields:

- `hpc_on` (`logical`) is `true` if parallelization is to be implemented on a cluster for flux estimation or continuation.
- `int_maxstep` (`double`) is the maximum integration stepsize.
- `int_timeout` (`double`) is the maximum time (in seconds) that will be spent on a single forward simulation during a flux estimation.
- `int_reltol` (`double`) is the relative tolerance used to control numerical error when simulating nonstationary labeling measurements.
- `int_senstol` (`double`) is the relative tolerance used to control numerical error when simulating the sensitivities of nonstationary labeling measurements.
- `int_tspan` (`double`) is a vector of time points to be simulated. It is only applicable in the nonstationary case. If the vector has only two elements and the first is zero, then the simulated points will be chosen automatically to give a smooth profile reflecting the stiffness of the system. If the vector has three or more elements, then measurements will only be simulated for those specified time points.2
- `int_ms_norm` (`logical`) is `true` if mass spectrometry data is to be normalized.
- `int_path_type` (`string`) is `'em'` for elementary modes and `'ep'` for extreme pathways when conducting pathway analysis.
- `sim_na` (`logical`) is `true` if natural abundances are to be included in the simulation.
- `sim_sens` (`logical`) is `true` if measurement sensitivities are to be simulated.
- `sim_ss` (`logical`) is `true` if only stationary measurements are to be simulated. In this case, time points specified within the measurement data and the remainder of the options are ignored.
- `sim_tunit` (`string`) is the units for any measurement time points (e.g., `'h'`).

- `fit_nudge` (`double`) is set to a value between 0 and 1 to randomly nudge the initial guess in a flux estimation.
- `fit_omit` (`double`) is a vector of time points to omit from the experimental data.
- `fit_renorm` (`logical`) is `true` if mass spectrometry measurements are to be fit to the normalization factor (specified in the measurement data).
- `fit_reltol` (`double`) is the relative tolerance used in the flux estimation.
- `fit_tau` (`double`) sets the λ parameter in the Levenberg-Marquardt optimization scheme used in flux estimation.
- `cont_alpha` (`double`) is the α-level used for parameter continuation.
- `cont_steps` (`double`) is the desired number of steps in each parameter continuation.
- `cont_reltol` (`double`) is the allowable predictor error in the χ^2 value during parameter continuation.

`parameter` holds data for a specific flux or concentration parameter generated during estimation and continuation and contains the following fields:

- `id` (`string`) is the parameter's identification tag.
- `free` (`logical`) is `true` or `false` if the parameter is free or fixed, respectively.
- `type` (`string`) is `'flux'` or `'pool'` if the parameter represents a metabolic flux or metabolite concentration, respectively.
- `val` (`double`) is the parameter's estimated value.
- `std` (`double`) is the parameter's standard error calculated from the inverse of the Hessian.

- `lb` (`double`) is the lower bound of the parameter's confidence interval calculated from continuation.
- `ub` (`double`) is the upper bound of the parameter's confidence interval calculated from continuation.
- `cor` (`double`) is the fit's correlation vector for the parameter relative to all other parameters in the system.
- `cov` (`double`) is the fit's covariance vector for the parameter relative to all other parameters in the system.
- `cont` (`logical`) is `true` if confidence intervals are to be calculated for this parameter.
- `chi2s` (`double`) is a vector of sums of squared residuals calculated during parameter continuation.
- `vals` (`double`) is a vector of parameter values calculated during continuation.
- `unit` (`string`) is the parameter value's unit of measurement.

`pathway` is an elementary mode or extreme pathway resulting from metabolic network analysis and contains the following fields:

- `id` (`string`) is the identification tag of the mode/pathway.
- `flx` (`logical`) is a vector corresponding to network reactions in which `true` indicates that a reaction is included in the mode/pathway.
- `sub` (`stoich`) describes the identity and stoichiometry of the substrates of the mode/pathway.
- `prod` (`stoich`) describes the identity and stoichiometry of the products of the mode/pathway.

`reaction` represents a chemical reaction within a metabolic network and contains the following fields:

- `id` (`string`) is the reaction's identification tag.
- `on` (`logical`) is `true` or `false` if the reaction is to be used or ignored.
- `flx` (`flux`) contains the flux and chemical properties of the reaction. A reversible reaction will have an array of two fluxes (forward and backward) while an irreversible reaction will have just one (forward).
- `path` (`string`) is the pathway to which the reaction belongs.
- `unit` (`string`) is the reaction flux's unit of measurement.

`residual` holds a parameter's residual and associated information resulting from a flux estimation and contains the following fields:

- `expt` (`string`) is the identification tag of the experiment to which the residual belongs.
- `type` (`string`) is `'flux'`, `'pool'`, or `'ms'` for flux, concentration, or mass spectrometry measurements, respectively.
- `id` (`string`) is the identification tag of the measurement from which the residual is calculated.
- `peak` (`string`) is the relative mass of the measurement from which the residual is calculated (e.g., `'M+0'` or `'M+1'`), in the case of mass spectrometry measurements.
- `time` (`double`) is the time of the measurement from which the residual is calculated.
- `val` (`double`) is the value of the residual.

- `std` (`double`) is the standard error of the measurement from which the residual is calculated.
- `cont` (`double`) is a vector with the relative contributions of all the flux and concentration parameters to the residual.

`simdata` holds simulation information for a labeling measurement and contains the following fields:

- `expt` (`string`) is the identification tag of the experiment of the simulated measurement.
- `id` (`string`) is the identification tag of the simulated measurement.
- `time` (`double`) is the measurement time point.
- `val` (`double`) is the value of the simulated measurement.
- `sens` (`double`) is the simulated sensitivity matrix.
- `tunit` (`string`) is the unit of the measurement time point.

`state` represents a metabolite state within the reaction network. While the `metabolite` class describes the properties of a general chemical compound, the `state` class describes the properties of a specific manifestation of that metabolite in a certain cellular compartment. It contains the following fields:

- `id` (`string`) is the metabolite state's identification tag.
- `val` (`double`) is the value of the metabolite state's concentration.
- `lb` (`double`) is the user-defined lower bound of the metabolite state.
- `ub` (`double`) is the user-defined upper bound of the metabolite state.
- `met` (`string`) is the identification tag of the metabolite in the metabolite state.

- `phase` (`string`) is the identification tag of the phase (i.e., compartment) to which the metabolite state belongs.
- `fix` (`logical`) is `true` or `false` if the metabolite state's concentration is fixed or free, respectively.
- `type` (`string`) is `'src'`, `'snk'`, or `'int'` if the metabolite state is a source, sink, or internal, respectively.
- `bal` (`logical`) should be set to `false` if the metabolite state appears to be balanced in the network (i.e., it has both inputs and outputs) but is actually not balanced. One common example is CO_2.
- `unit` (`string`) is the unit of measurement for the concentration of the metabolite state.

`stoich` is a stoichiometric component of a reaction and contains the following fields:

- `id` (`string`) is the identification tag of metabolite state participating in the reaction.
- `val` (`double`) is the stoichiometric value of the metabolite state in the reaction.
- `map` (`map`) is an atomic transition map for the metabolite state in the reaction.

`tracer` represents an isotopic tracer used in a labeling experiment and contains the following fields:

- `id` (`string`) is the tracer's identification tag.
- `state` (`string`) is the tracer's metabolite state.
- `atoms` (`list`) describes which atoms in the tracer are labeled, where the list item's identification tag is a space-delimited string of atom identification tags, and the list item's value is a two-element vector in which the first element is

the fractional impurity of the tracer and the second element is the fractional purity.

- `frac` (`double`) is the fraction of the metabolite state in the isotopic form described by the tracer.

B.3 Initialization Functions

`reaction` Construction of reaction class.

- `rxn = reaction(s)` creates a reaction object `rxn` from `s`. `s` can be a string, reaction object, cell array of strings, struct, or empty. If `s` is a string or cell array, each row or cell should contain a reaction of the form `'a A + b B + ... -> c C + d D + ...'` A double-sided arrow (`'<->'`) can be used to designate a reversible reaction. Any term in the equation can be followed by a pair of matched parentheses containing the atom mapping for that term. If `s` is numeric, the reaction counter is reset to `s`.

- `rxn = reaction(s,p1,v1,...)` uses the list of (`p1`,`v1`) property name and value pairs to set the corresponding property values of `rxn`.

- `rxn = reaction` creates a default reaction object.

`msdata` Construction of msdata class.

- `msd = msdata(s)` creates a msdata object `msd` from `s`. `s` can be a string, cell array of strings, msdata object, struct, or empty. If `s` is a string or cell array, each row or cell should be of the form `'A: B @ C1 C2 C3'` where `A` is the name of the measurement, `B` is the measured metabolite state and `C1`, `C2`, `C3` are the atoms of `B` that are contained in the fragment.

- `msd = msdata(s,p1,v1,...)` uses the list of (`p1`,`v1`) property name and value pairs to set the corresponding property values of `msd`.

- `msd = msdata` creates a default msdata object.

`tracer` Construction of tracer class.

- `tr = tracer(s)` creates a tracer object `tr` from `s`. `s` can be a string, cell array of strings, tracer object, struct, or empty. If `s` is a string or cell array, each row or cell should be of the form `'A: B @ C1 C2, C3 C4'` where `A` is the tracer name, `B` is the identification tag of the metabolite state, and `C1 C2, C3 C4` specifies groups of atoms that have the same labeling.
- `tr = tracer(s,p1,v1)` uses the list of (`p1`,`v1`) property name and value pairs to set the corresponding property values of `tr`.
- `tr = tracer` creates a default tracer object.

`experiment` Construction of experiment class.

- `xpt = experiment(tr)` creates an experiment object `xpt` from the tracer object `tr`. If `tr` is already an experiment object, `tr` is returned in `xpt`. If `tr` is empty, an empty experiment is returned.
- `xpt = experiment(tr,p1,v1,...)` uses the list of (`p1`,`v1`) property name and value pairs to set the corresponding property values of `xpt`.
- `xpt = experiment` creates a default experiment object.

`model` Construction of model class

- `mod = model(s)` creates a model object `mod` from a reaction object or struct `s`. If `s` is already a model object, its metabolites and metabolite states are checked for consistency. If `s` is empty, an empty model is returned.
- `mod = model(s,p1,v1)` uses the list of (`p1`,`v1`) property name and value pairs to set the corresponding property values of `mod`.
- `mod = model` creates a default model object.

B.4 Driver Functions

`simulate` Model simulation.

- `sim = simulate(mod)` simulates labeling measurements in the model object `mod` and returns the results in the simdata object `sim`.
- `[sim,par] = simulate(mod)` also returns the parameter structure `par`.

`estimate` Parameter estimation.

- `fit = estimate(mod)` estimates free parameters in the model object `mod` and returns the results in the fitdata object `fit`.
- `[fit,par,mnt] = estimate(mod)` also returns the parameter structure `par` and the measurement structure `mnt`.
- `fit = estimate(mod,n)` performs `n` estimations using a multistart method and returns the best result.

`continuate` Continuation analysis of parameters near fitted estimates.

- `fit = continuate(fit,mod)` performs continuation starting from the optimal parameter estimates in the fitdata object `fit` and returns the updated fitdata object.
- `[fit,bestfit] = continuate(fit,mod)` also returns the best fit encountered during the continuation.

B.5 Utility Functions

`startup` Set preferences and update path.

- `startup` sets figure, axes, line, and text object defaults. The function also initializes the random number generator to a random state and updates the path to include those directories specified by the `mypath` function. `startup` should be called at the beginning of every new MetranCL session.

`mod2stoich` Computation of feasible fluxes.

- `flx = mod2stoich(mod)` computes the nearest feasible solution (in the least-squares sense) to the flux distribution specified in the model object `mod` and returns the result in the vector `flx`.
- `[flx,par] = mod2stoich(mod)` also returns the parameter structure `par` for the network.

`fit2mod` Update model with optimal parameter estimates.

- `mod = fit2mod(mod0,fit)` replaces the flux and pool size parameters in the model object `mod0` with the optimal values contained in the fitdata object `fit` and returns an updated model `mod`.

`integrate` Integrate GC/MS raw data.

- `integrate(dir,lib)` searches raw GC/MS data found in the directory `dir` for the metabolite peaks specified in the library structure `lib` and then integrates and determines mass isotopomer distributions. All results are stored in `dir` in the files `data.csv`.

B.6 Modeling *E. coli* Metabolism

We have included an example script written for Matlab that will generate the large model of *E. coli* metabolism that we used in Chapter 3. The script inputs flux, concentration, and MS measurements into the model, estimates fluxes and concentrations, and lastly calculates 95% confidence intervals for each parameter.

```
r = reaction({

% Glycolysis
'G6P (abcdef) <-> F6P (abcdef)'
'F6P (abcdef) -> FBP (abcdef)'
'FBP (abcdef) <-> DHAP (cba) + GAP (def)'
'DHAP (abc) <-> GAP (abc)'
'GAP (abc) <-> PG3 (abc)'
'PG3 (abc) <-> PEP (abc)'
'PEP (abc) -> Pyr (abc)'

% Pentose phosphate pathway
'G6P (abcdef) -> PG6 (abcdef)'
'PG6 (abcdef) -> Ru5P (bcdef) + CO2 (a)'
'Ru5P (abcde) <-> X5P (abcde)'
'Ru5P (abcde) <-> R5P (abcde)'
'X5P (abcde) <-> GAP (cde) + EC2 (ab)'
'F6P (abcdef) <-> E4P (cdef) + EC2 (ab)'
'S7P (abcdefg) <-> R5P (cdefg) + EC2 (ab)'
'F6P (abcdef) <-> GAP (def) + EC3 (abc)'
'S7P (abcdefg) <-> E4P (defg) + EC3 (abc)'

% Entner-Doudoroff pathway
'PG6 (abcdef) -> KDPG (abcdef)'
'KDPG (abcdef) -> Pyr (abc) + GAP (def)'

% TCA cycle
'Pyr (abc) -> AcCoA (bc) + CO2 (a)'
'OAA (abcd) + AcCoA (ef) -> Cit (dcbfea)'
'Cit (abcdef) <-> ICit (abcdef)'
'ICit (abcdef) <-> AKG (abcde) + CO2 (f)'
'AKG (abcde) -> SucCoA (bcde) + CO2 (a)'
'SucCoA (abcd) <-> Suc (abcd)'
'Suc (abcd) <-> Fum (abcd)'
'Fum (abcd) <-> Mal (abcd)'
'Mal (abcd) <-> OAA (abcd)'

% Amphibolic reactions
'Mal (abcd) -> Pyr (abc) + CO2 (d)'
'PEP (abc) + CO2 (d) <-> OAA (abcd)'

% Acetic acid formation
'AcCoA (ab) <-> Ac (ab)'

% PDO biosynthesis
'DHAP (abc) <-> Glyc3P (abc)'
'Glyc3P (abc) -> Glyc (abc)'
'Glyc (abc) -> HPA (abc)'
'HPA (abc) -> PDO (abc)'

% Amino acid biosynthesis
'AKG (abcde) -> Glu (abcde)'
'Glu (abcde) -> Gln (abcde)'
'Glu (abcde) -> Pro (abcde)'
```

```
['Glu (abcde) + CO2 (f) + Gln (ghijk) + Asp (lmno) + AcCoA (pq) -> Arg ',...
    '(abcdef) + AKG (ghijk) + Fum (lmno) + Ac (pq)']
'OAA (abcd) + Glu (efghi) -> Asp (abcd) + AKG (efghi)'
'Asp (abcd) -> Asn (abcd)'
'Pyr (abc) + Glu (defgh) -> Ala (abc) + AKG (defgh)'
'PG3 (abc) + Glu (defgh) -> Ser (abc) + AKG (defgh)'
'Ser (abc) <-> Gly (ab) + MEETHF (c)'
'Gly (ab) <-> CO2 (a) + MEETHF (b)'
'Thr (abcd) -> Gly (ab) + AcCoA (cd)'
'Ser (abc) + AcCoA (de) -> Cys (abc) + Ac (de)'
['Asp (abcd) + Pyr (efg) + Glu (hijkl) + SucCoA (mnop) -> LLDAP ',...
    '(abcdgfe) + AKG (hijkl) + Suc (mnop)']
'LLDAP (abcdefg) -> Lys (abcdef) + CO2 (g)'
'Asp (abcd) -> Thr (abcd)'
['Asp (abcd) + METHF (e) + Cys (fgh) + SucCoA (ijkl) -> Met (abcde) + ',...
    'Pyr (fgh) + Suc (ijkl)']
'Pyr (abc) + Pyr (def) + Glu (ghijk) -> Val (abcef) + CO2 (d) + AKG (ghijk)'
['AcCoA (ab) + Pyr (cde) + Pyr (fgh) + Glu (ijklm) -> Leu (abdghe) + ',...
    'CO2 (c) + CO2 (f) + AKG (ijklm)']
'Thr (abcd) + Pyr (efg) + Glu (hijkl) -> Ile (abfcdg) + CO2 (e) + AKG (hijkl)'
['PEP (abc) + PEP (def) + E4P (ghij) + Glu (klmno) -> Phe (abcefghij) + ',...
    'CO2 (d) + AKG (klmno)']
['PEP (abc) + PEP (def) + E4P (ghij) + Glu (klmno) -> Tyr (abcefghij) + ',...
    'CO2 (d) + AKG (klmno)']
['Ser (abc) + R5P (defgh) + PEP (ijk) + E4P (lmno) + PEP (pqr) + Gln ',...
    '(stuvw) -> Trp (abcedklmnoj) + CO2 (i) + GAP (fgh) + Pyr (pqr) + ',...
    'Glu (stuvw)']
['R5P (abcde) + FTHF (f) + Gln (ghijk) + Asp (lmno) -> His (edcbaf) + ',...
    'AKG (ghijk) + Fum (lmno)']

% One carbon metabolism
'MEETHF (a) -> METHF (a)'
'MEETHF (a) -> FTHF (a)'

% Transport
'Gluc.pre (abcdef) -> G6P (abcdef)'
'Gluc.ext (abcdef) -> G6P (abcdef)'
'Cit.ext (abcdef) -> Cit (abcdef)'
'Glyc.ext (abc) + Force.ext <-> Glyc (abc) + Force'
'PDO (abc) -> PDO.ext (abc)'
'Ac (ab) -> Ac.ext (ab)'
'CO2 (a) -> CO2.ext (a)'

% Biomass formation
['0.488 Ala + 0.281 Arg + 0.229 Asn + 0.229 Asp + 0.087 Cys + 0.250 Glu + ',...
    0.250 Gln + 0.582 Gly + 0.090 His + 0.276 Ile + 0.428 Leu + 0.326 Lys + ',...
    0.146 Met + 0.176 Phe + 0.210 Pro + 0.205 Ser + 0.241 Thr + 0.054 Trp + ',...
    0.131 Tyr + 0.402 Val + 0.205 G6P + 0.071 F6P + 0.754 R5P + 0.129 GAP + ',...
    0.619 PG3 + 0.051 PEP + 0.083 Pyr + 2.510 AcCoA + 0.087 AKG + 0.340 OAA + ',...
    0.443 MEETHF -> 39.68 Biomass']

% Forcing flux
'Force -> Force.ext'
```

```
})

t = tracer({
'[1-13C]Gluc: Gluc.ext @ 1'
'[U-13C]Gluc: Gluc.ext @ 1 2 3 4 5 6'
})
t.frac = [0.75,0.25];
t.atoms(1).it = [0.01;0.99];
t.atoms(2).it = [0.01;0.99];

x = experiment(t);

f = data('R61.f R68.f R64.f R66.f R67.f R62.f R65.f');
f.val = [92.6,2.0,129.2,177.5,0.9,0.3,0.3];
f.std = f.val*0.05;
x.data_flx = f;

c = data('AKG Ala Asp Cit Glu Gly Ile Leu Mal Met Phe Pyr Ser Suc Thr Tyr Val');
c.val = [0.8,21,14,37,138,1.0,2.2,13,2.8,1.0,1.2,16,3,0.5,4.2,3.6,7.7]*1e-4;
c.std = f.val*0.15;
x.data_cxn = c;

d = msdata({
'AKG346: AKG @ 1 2 3 4 5'
'Ala232: Ala @ 2 3'
'Ala260: Ala @ 1 2 3'
'Asp302: Asp @ 1 2'
'Asp376: Asp @ 1 2'
'Asp390: Asp @ 2 3 4'
'Asp418: Asp @ 1 2 3 4'
'Cit459: Cit @ 1 2 3 4 5 6'
'Glu330: Glu @ 2 3 4 5'
'Glu432: Glu @ 1 2 3 4 5'
'Gly218: Gly @ 2'
'Gly246: Gly @ 1 2'
'Ile200: Ile @ 2 3 4 5 6'
'Ile274: Ile @ 2 3 4 5 6'
'Leu274: Leu @ 2 3 4 5 6'
'Mal419: Mal @ 1 2 3 4'
'Met218: Met @ 2 3 4 5'
'Met292: Met @ 2 3 4 5'
'Met320: Met @ 1 2 3 4 5'
'Phe234: Phe @ 2 3 4 5 6 7 8 9'
'Phe302: Phe @ 1 2'
'Phe308: Phe @ 2 3 4 5 6 7 8 9'
'Phe336: Phe @ 1 2 3 4 5 6 7 8 9'
'Pyr174: Pyr @ 1 2 3'
'Ser288: Ser @ 2 3'
'Ser302: Ser @ 1 2'
'Ser362: Ser @ 2 3'
'Ser390: Ser @ 1 2 3'
'Suc289: Suc @ 1 2 3 4'
'Thr376: Thr @ 2 3 4'
'Thr404: Thr @ 1 2 3 4'
```

```
'Tyr302: Tyr @ 1 2'
'Val260: Val @ 2 3 4 5'
'Val288: Val @ 1 2 3 4 5'
});

d{'AKG346'}.more = 'c9h28o5n1si2';
d{'Ala232'}.more = 'c8h26onsi2';
d{'Ala260'}.more = 'c8h26o2nsi2';
d{'Asp302'}.more = 'c12h32o2n1si2';
d{'Asp376'}.more = 'c14h38o3nsi3';
d{'Asp390'}.more = 'c14h40o3n1si3';
d{'Asp418'}.more = 'c14h40o4n1si3';
d{'Cit459'}.more = 'c14h39o6si3';
d{'Glu330'}.more = 'c12h36o2n1si2';
d{'Glu432'}.more = 'c14h42o4n1si3';
d{'Gly218'}.more = 'c8h24onsi2';
d{'Gly246'}.more = 'c8h24o2nsi2';
d{'Ile200'}.more = 'c6h26nsi';
d{'Ile274'}.more = 'c8h32onsi2';
d{'Leu274'}.more = 'c8h32onsi2';
d{'Mal419'}.more = 'c14h39o5si3';
d{'Met218'}.more = 'c6h24nsis';
d{'Met292'}.more = 'c8h30nosi2s';
d{'Met320'}.more = 'c8h30no2si2s';
d{'Phe234'}.more = 'c6h24nsi';
d{'Phe302'}.more = 'c12h32o2nsi2';
d{'Phe308'}.more = 'c8h30onsi2';
d{'Phe336'}.more = 'c8h30o2nsi2';
d{'Pyr174'}.more = 'c3h12o3n1si';
d{'Ser288'}.more = 'c12h34onsi2';
d{'Ser302'}.more = 'c12h32o2nsi2';
d{'Ser362'}.more = 'c14h40o2nsi3';
d{'Ser390'}.more = 'c14h40o3n1si3';
d{'Suc289'}.more = 'c8h25o4si2';
d{'Thr376'}.more = 'c14h42o2nsi3';
d{'Thr404'}.more = 'c14h42o3nsi3';
d{'Tyr302'}.more = 'c12h32o2nsi2';
d{'Val260'}.more = 'c8h30onsi2';
d{'Val288'}.more = 'c8h30o2nsi2';

d{'AKG346'}.val = [[0.8605,0.8539,0.8417,0.8403,0.8116,0.7642];
                   [0.1026,0.1100,0.1191,0.1127,0.1267,0.1447];
                   [0.0369,0.0350,0.0353,0.0335,0.0345,0.0423];
                   [0.0000,0.0012,0.0039,0.0135,0.0273,0.0489];
                   [0.0000,0.0000,0.0000,0.0000,0.0000,0.0000]];
d{'Ala232'}.val = [[0.8605,0.8539,0.8417,0.8403,0.8116,0.7642];
                   [0.1026,0.1100,0.1191,0.1127,0.1267,0.1447];
                   [0.0369,0.0350,0.0353,0.0335,0.0345,0.0423];
                   [0.0000,0.0012,0.0039,0.0135,0.0273,0.0489];
                   [0.0000,0.0000,0.0000,0.0000,0.0000,0.0000]];

d{'AKG346'}.std = [[0.8605,0.8539,0.8417,0.8403,0.8116,0.7642];
                   [0.1026,0.1100,0.1191,0.1127,0.1267,0.1447];
                   [0.0369,0.0350,0.0353,0.0335,0.0345,0.0423];
```

```
                    [0.0000,0.0012,0.0039,0.0135,0.0273,0.0489];
                    [0.0000,0.0000,0.0000,0.0000,0.0000,0.0000]];
d{'Ala232'}.std = [[0.8605,0.8539,0.8417,0.8403,0.8116,0.7642];
                    [0.1026,0.1100,0.1191,0.1127,0.1267,0.1447];
                    [0.0369,0.0350,0.0353,0.0335,0.0345,0.0423];
                    [0.0000,0.0012,0.0039,0.0135,0.0273,0.0489];
                    [0.0000,0.0000,0.0000,0.0000,0.0000,0.0000]];

t = [0.00056;0.00111;0.00222;0.00417;0.00833;0.01667];

for i = 1:length(d)
    d(i).time = t;
end

x.data_ms = d;
m = model(r,'expts',x);

% Define metabolite symmetries
m.mets{'Suc'}.sym = list('flip',map('1:4 2:3 3:2 4:1'))
m.mets{'Fum'}.sym = list('flip',map('1:4 2:3 3:2 4:1'))

m.options.fit_reinit = true;
m.options.sim_ss = false;
m.options.int_reltol = 0.001;

m.states{'Glyc.ext'}.bal = false;
```

Appendix C

Rapid Sampler Design and Construction

The materials and equipment necessary to construct the rapid sampler (discussed in Chapter 5) are listed in Table C.1. The vendors and part numbers for the particular supplies we used are also listed. A basic diagram of the overall apparatus is shown in Figure C-1. The sampler can be divided into five basic modules, as indicated in Figure C-1: support, sampling, sparging, vacuum, and valve control:

- **Support:** The sampler is constructed on a mobile cart, so that it can be moved with ease. This consideration is especially important in case there are multiple reactors that might require sampling at different times. Mobility is also important when space is at a premium. We situated our larger pieces of equipment (the computer and the vacuum flask) on the lower shelf, while we built the valves and quenching tubes on the upper shelf so that they would be closer to our benchtop reactors (thereby minimizing sampling lines and corresponding delays). Two ring stands were linked together by rods and clamps on the upper shelf to hold the valves and manifolds above the quenching tubes.

- **Sampling:** Quenching tubes sat on a rack in a styrofoam cooler (lined with plastic to prevent leakage) filled with ethanol. Before sampling, dry ice was added to the cooler to bring the bath temperature down to -20°C. The design

Sample Tubes		
50-mL centrifuge tubes	VWR	21008-951
2-hole stoppers (size 6)	VWR	59582-246
Rigid tubing $^{11}/_{64}$" × $^{1}/_{4}$"	Cole-Parmer	95100-02
Straight connectors 4-6 mm	Cole-Parmer	06288-10
Straight connectors 6-8 mm	Cole-Parmer	06288-20
Sparging needles	VWR	BD405148
Valve Circuitry		
Solenoid pinch valves	ASCO Valves	SCH284A00524VDC
Valve cable glands	ASCO Valves	290416-001PG9
15A 13.8V power supply	RadioShack	22-508
Barrier strips	RadioShack	274-670
Spade and ring tongues	RadioShack	64-407
1A 250V glass fuses	RadioShack	278-1224
Inline fuse holders	RadioShack	270-1281
Hookup wire (22AWG)	RadioShack	278-1224
Valve Control		
Desktop PC	Dell	Inspiron 570
LabVIEW software	National Instruments	776671-35
Data aquisition card	National Instruments	779083-01
Connector block	National Instruments	778673-01
Shielded cable	National Instruments	778621-01
Tubing and Connectivity		
Manifolds with luer locks	Cole-Parmer	30600-42
Barbed luer adapters	Cole-Parmer	45504-80
Luer plugs	Cole-Parmer	4504-56
Silicone tubing $^{1}/_{8}$" × $^{3}/_{16}$"	Cole-Parmer	95802-04
Silicone tubing $^{3}/_{16}$" × $^{5}/_{16}$"	Cole-Parmer	95802-09
Vacuum flask	VWR	KT953760-2002
1-hole stopper (size 9)	VWR	59581-403
Vacuum tubing $^{1}/_{4}$" × $^{1}/_{2}$"	VWR	62995-059
Support Frame		
2-shelf lab cart	VWR	82020-798
Support stands	VWR	12985-068
Aluminum rods	VWR	60079-406
Clamp holders	VWR	12621-250
3-outlet extension cord	Office Depot	350605

Table C.1: List of materials used in assembling the rapid sampling apparatus. The vendor and part number for each piece of equipment is also supplied.

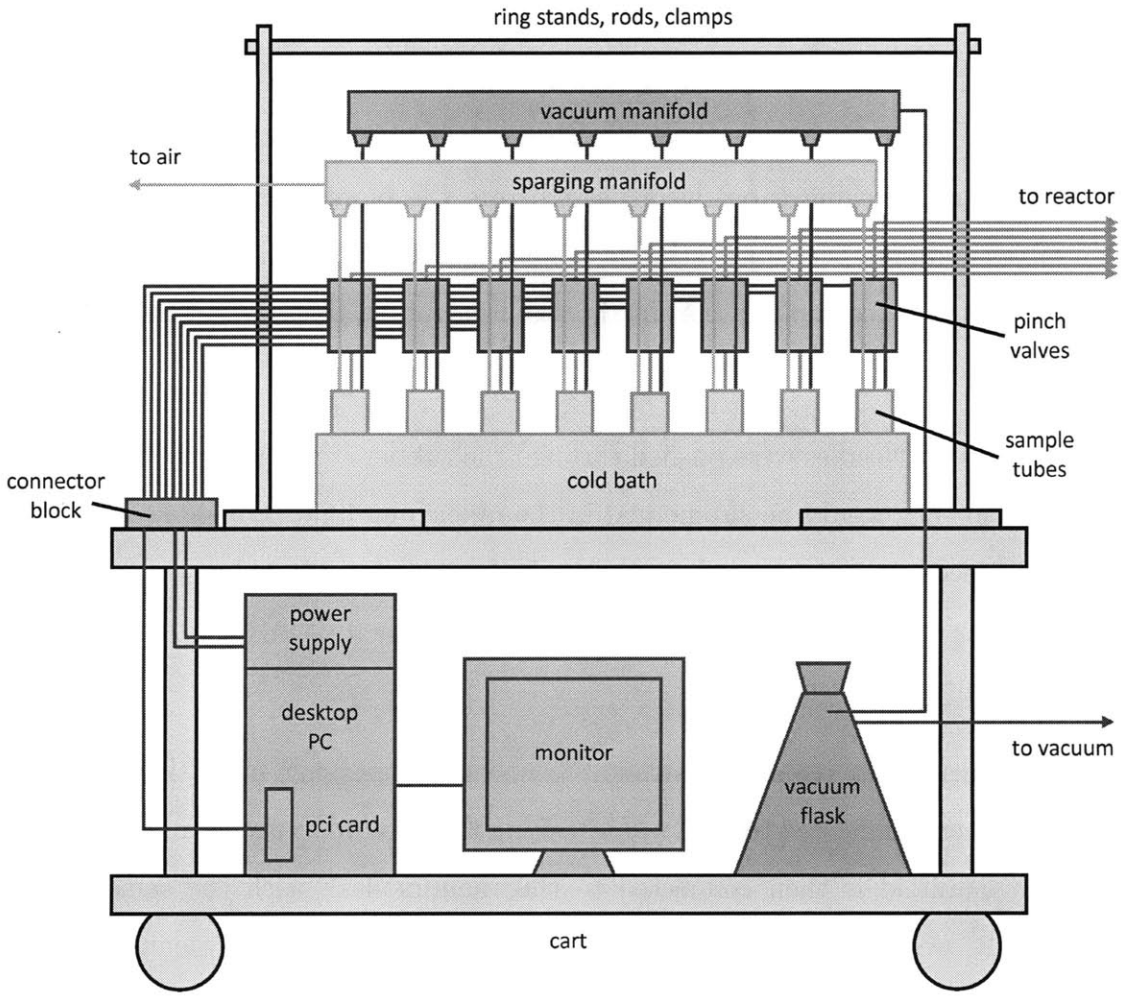

Figure C-1: Basic layout of the rapid sampling apparatus. Red indicates electrical circuitry for valve control, blue indicates air lines for sparging and mixing, purple indicates vacuum lines, and grey indicates the support framework for the sampler. The cold bath sits on the upper shelf of the cart and inside it sit the sampling tubes. A pair of ring stands are connected by rods and clamps (not shown) which hold the solenoid pinch valves and the vacuum and sparging manifolds. The connector block sits on the upper shelf and is wired to each valve. The computer rests on the lower shelf and is wired to the connector block. The valves' power supply is also wired to the connector block. The computer and power supply also must all be connected to an electrical outlet (not shown). Finally, a vacuum flask connects the vacuum manifold to the vacuum source, serving as a trap.

of the quenching tubes is depicted in Figure 5-2. Approximately two inches of rigid tubing was pushed through the each hole of a size 6 rubber stopper for every quenching tube, so that an equal amount protrudes from either size. One external line was then connected to a 4-6 mm straight connector (sampling line) and the other to a 6-8 mm straight connector (vacuum line). Sampling lines (one per quenching tube) were passed from the bioreactor through the pinch valves and onto the tubes' straight connectors. It is important that the silicone tubing used as sampling lines fit the requirements of the valves. Our particular valves were designed for 1/8" × 3/16" tubing.

- **Sparging:** Needles were pushed through the rubber stoppers of each quenching tube so they would be submerged in the quenching fluid. Outside of the tube, each needle was connected to 3/16" × 5/16" silicone tubing that ran up to the sparging manifold. The sparging manifold was then connected to a regulated air or nitrogen supply.

- **Vacuum:** The remaining straight connectors emerging from the quenching tubes were connected to 3/16" × 5/16" tubing that ran up to the vacuum manifold. The manifold is then connected to the vacuum flask with the same tubing. Finally, we used vacuum tubing to connect the flask to house vacuum.

- **Valve control:** Valves were wired to the connector block and the power supply as explained in detail in Figure C-2. The connector block was hooked to the National Instruments PCI card using the shielded cable. The desktop was loaded with LabVIEW software as well as software specific to the PCI data output card.

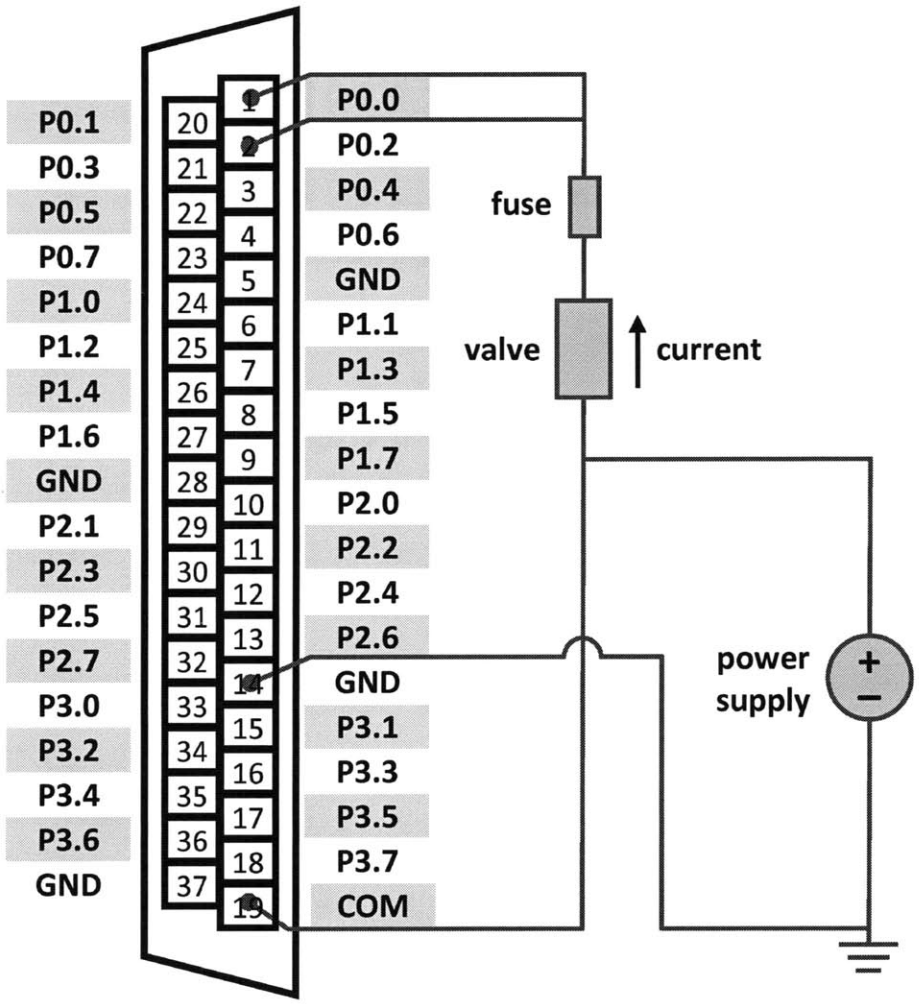

Figure C-2: The NI PCI-6517 from National Instruments can be wired to solenoid valves and a power supply and can then be programmed from LabVIEW to control these valves. An example connection is shown here. Two lines lead out from the valve because it requires twice the current that one port is able to receive. A fuse is included in the circuitry to prevent the valve and ports from being damaged by an accidental high current.

Appendix D

Tracer Optimization Code

D.1 Driver Function

`evolve` Evolve tracer mixtures with high-precision flux estimates.

```
function [pops,sels,dscs] = evolve(mod,tracers,varargin)
%EVOLVE Evolve tracer mixtures with high-precision flux estimates
%   [POPS,SELS,DSCS] = EVOLVE(MOD,TRACERS) creates an initial population of
%   mixtures of tracers from TRACERS and evaluates them in the model MOD,
%   and then evolves that population to find new tracer mixtures that
%   generate high-precision flux estimates.
%
%   [POPS,SELS,DSCS] = EVOLVE(MOD,TRACERS,'PropertyName1',PropertyValue1,
%   'PropertyName2',PropertyValue2,...) defines multiple properties of the
%   EVOLVE function. Property names include:
%   HPC_ON      If TRUE use parallel computing
%   NEST        The number of estimations for each tracer
%   NEXP        The number of simulated experiments for each tracer
%   FAST        If TRUE use std deviations, if FALSE use conf intervals
%   INITPOP     Initial population of tracer mixtures
%   NROUNDS     Number of rounds of selection
%   NPOP        Number of tracer mixtures in initial population
%   TOURNEYSIZE The size of each selection tournament
%   NTOURNEYS   The number of selection tournaments within each round
%   PMUT        The mutation frequency
%   WEIGHT      The weight assigned to each independent flux
%   RANDFREQ    The frequency of a bit being on for init random population

    % Default property values
    hpc_on = true;
    nest = 6;
    nexp = 6;
    fast = false;
```

```
initpop = [];
nrounds = 10;
npop = 100;
tourneysize = 4;
ntourneys = 25;
pmut = 0.001;
weight = [];
randfreq = 0.1;
sparse = true;

% Assign user-defined property values
while ~isempty(varargin)
    property = lower(varargin{1});
    value = varargin{2};
    eval([property,' = value;']);
    varargin(1:2) = [];
end

% Create encoder object for available tracers list
enc = encoder(tracers);

% Create argument list to be passed into fitness object
params = {'hpc_on',hpc_on,'nest',nest,'nexp',nexp,'fast',fast,...
          'weight',weight,'sparse',sparse};

% Create fitness object
fness = fitness(mod,enc,params{:});

% Initialize containers
pops = cell([1,nrounds+1]);
sels = cell([1,nrounds]);
dscs = cell([1,nrounds]);
time = zeros([1,nrounds]);

% If initial population not user-defined, create random population
len = size(encode(enc,tracers),2);
if isempty(initpop)
    pop = logical(zeros([npop,len]));
    for i = 1:npop
       pop(i,:) = rand([1,len]) < randfreq;
    end
else
    pop = initpop;
end

% Loop over each round of selection
pops{1} = pop;
for i = 1:nrounds
    tic;
    [chsomes,sel,dsc] = tournament(pop,tourneysize,ntourneys,fness);
    pop = mate(chsomes,npop);
    pop = mutation(pop,pmut);
    pops{i+1} = pop;
    sels{i} = sel;
```

```
        dscs{i} = dsc;
        time(i) = toc;
        save evolve_temp
    end
end
```

D.2 Evaluation Functions

`evalfitness` Method to calculate fitness of a given tracer mixture.

```
function score = evalfitness(obj,tracer,id)
% Method to calculate fitness of a given tracer mixture

    if isstruct(obj)
        obj.encoder = encoder(obj.encoder);
        obj = fitness(obj);
    end

    sparse = obj.sparse;

    i = 0;
    j = 0;
    nmax = 2*obj.nexp;
    fit = fitdata();
    while i < obj.nexp && (i+j) < nmax
        samplefit = tracerfit(obj,tracer);
        if ~isempty(samplefit)
            i = i+1;
            fit(i) = samplefit;
        else
            j = j+1;
        end
    end

    % Pare down the fitdata object
    fit = struct('chi2',extract(fit,'chi2'), ...
                 'par',extract(fit,'par'));

    % Pare down parameter objects within fitdata objects
    for j = 1:length(fit)
        p = fit(j).par;
        fit(j).par = struct('id',p.id, ...
                            'lb',extract(p,'lb'), ...
                            'std',extract(p,'std'), ...
                            'ub',extract(p,'ub'), ...
                            'val',extract(p,'val'));
    end

    % Get fitness score
```

```
    [score_f,stdev_f,score_s,stdev_s] = fitscore(fit,obj.weight);
    score.valfast = score_f;
    score.stdfast = stdev_f;
    score.valslow = score_s;
    score.stdslow = stdev_s;
    if ~sparse
        score.fit = fit;
        score.tracer = decode(obj.encoder,tracer);
    end
    score.chsome = tracer;
    score.id = id;

end

% ------------------------------------------------------------------------
function fit = tracerfit(obj,tracer)
% Helper function to calculate a fit for a tracer and model

    % Decode tracer and get measurements
    tracer = decode(obj.encoder,tracer);
    mod = getmnts(obj,tracer);

    % Set fitting options and estimate
    fit = estimate(mod,obj.nest);

    % A valid fit was not found; all confidence intervals are unknown
    if fit.Echi2(end) < fit.chi2
        fit = [];

    % Fast calculation only needs std, otherwise get intervals
    elseif ~obj.fast
        fit.par.cont = obj.pars;
        mod.options.fit_tau = 1e-16;
        fail = true;
        while fail && mod.options.fit_tau < 1
            try
                fit = continuate(fit,mod);
                fail = false;
            catch
                mod.options.fit_tau = mod.options.fit_tau*1000;
            end
        end
    end
end

% ------------------------------------------------------------------------
function mod = getmnts(obj,tracer)
% Helper function to introduce normal error into measurements

    % Simulate labeling and introduce normally distributed error
    mod = obj.mod;
    mod.expts.tracer = tracer;
    sim = simulate(mod);
```

```
    % If nonstationary, don't include t=0 as a measurement
    nt = size(sim.time,1);
    if nt > 1
        for i = 1:length(sim.val)
            mod.expts.data_ms(i).val = sim(i).val(:,2:end);
        end
        mod.expts.data_ms.time = sim.time(2:end,:)-repmat(sim.time(1,:),nt-1,1);
        mod.options.fit_tau = 1e-6;
    else
        mod.expts.data_ms.val = sim.val;
        mod.expts.data_ms.time = sim.time;
        mod.options.fit_tau = 1e-16;
    end

    % Iterate over MS data, create standard error, give normal error
    for i = 1:length(mod.expts.data_ms)
        ms = mod.expts.data_ms(i);
        mod.expts.data_ms(i).std = max(obj.msstd*ms.val,obj.msstdmin);
        ms = mod.expts.data_ms(i);
        mod.expts.data_ms(i).val = max(normrnd(ms.val,ms.std),0);
    end

    % Iterate over flux mnts, create standard error, give normal error
    for i = 1:length(mod.expts.data_flx)
        flx = mod.expts.data_flx(i);
        mod.expts.data_flx(i).val = mod.rates.flx{flx.id{1}}.val;
        flx = mod.expts.data_flx(i);
        mod.expts.data_flx(i).std = max(obj.flxstd*flx.val,obj.flxstdmin);
        flx = mod.expts.data_flx(i);
        mod.expts.data_flx(i).val = max(normrnd(flx.val,flx.std),0);
    end
end
```

`fitscore` Helper function that scores confidence intervals.

```
function [score_f,stdev_f,score_s,stdev_s] = fitscore(fit,weight,stdev,alf)
% Helper function that scores confidence intervals

    if ~isstruct(fit)
        score_f = NaN;
        stdev_f = NaN;
        score_s = NaN;
        stdev_s = NaN;
        return
    end

    oldfit = fit;
    fit = struct('chi2',{},'par',{});
    for i=1:length(oldfit)
        if ~isempty([oldfit(i).par.ub])
            fit(end+1) = oldfit(i);
        end
```

```
    end

    if nargin < 3
        stdev = 1;
    end
    if nargin < 4
        alf = 3;
    end

    npars = length(fit(1).par);
    if length(weight) < npars
        temp = logical(zeros(1,npars));
        temp(weight) = true;
        weight = temp;
    end
    pars = logical(weight);

    len = length(fit);
    idx = repmat(pars,[1,len]);
    weight = repmat(weight,[1,len]);
    pars = [fit.par];
    pars = pars(idx);
    weight = weight(idx);

    vabs = abs([pars.val]);
    vals = [pars.val]./vabs;

    stds = [pars.std]./vabs;
    stds(stds > alf) = alf;
    scores = exp(-2*stds/stdev^2);
    scores(isnan(scores)) = 0;
    scores = reshape(scores.*weight,[],length(fit));
    score_f = mean(sum(scores,1));
    stdev_f = std(sum(scores,1));

    lbs = [pars.lb]./vabs;
    ubs = [pars.ub]./vabs;
    lbs = min([lbs;vals]);
    ubs = max([ubs;vals]);
    lo = lbs < vals-alf;
    hi = ubs > vals+alf;
    lbs(lo) = vals(lo)-alf;
    ubs(hi) = vals(hi)+alf;
    rng = ubs-lbs;
    scores = exp(-rng/stdev^2);
    scores(isnan(scores)) = 0;
    scores = reshape(scores.*weight,[],length(fit));
    score_s = mean(sum(scores,1));
    stdev_s = std(sum(scores,1));

end
```

D.3 Selection Function

`tournament` Conducts tournament selection for a population.

```
function [chsomes,sels,dscs] = tournament(pop,trnysize,ntrnys,fness)
% Conducts tournament selection for a population

    % Error if there are more tourney slots than participants
    plen = size(pop,1);
    nsamples = ntrnys*trnysize;
    if nsamples > plen+trnysize
        error('Population too small to accomodate number of tournaments');
    end

    % Get a randomized list of participants
    [i,i] = sort(rand([1,plen]));
    if nsamples > length(i)
        j = sort(rand([1,plen]));
        i = [i,j(1:mod(plen,nsamples))];
    else
        i = i(1:nsamples);
    end

    % Conduct all tournaments
    if fness.hpc_on
        pops = permute(pop(i,:),[3 2 1]);
        ids = permute(num2str(i'),[3,2,1]);
        scores = paralize('evalfitness',[],fness,pops,ids);
        scores = permute(scores,[3,2,1]);
    else
        for j = 1:nsamples
            scores(j) = evalfitness(fness,pop(i(j),:),num2str(i(j)));
        end
    end

    % Retrieve scores for each tournament
    if fness.fast
        vals = [scores.valfast];
    else
        vals = [scores.valslow];
    end

    % Pull out best r6esult from each tournament
    vals = reshape(vals,trnysize,ntrnys);
    [i,i] = sort(vals,1);
    imax = i(end,:);
    idx = imax+(0:trnysize:nsamples-trnysize);
    sels = scores(idx);
    dscs = scores(setdiff(1:nsamples,idx));
    chsomes = reshape([sels.chsome],[],ntrnys)';
end
```

D.4 Recombination Functions

`mate` Mates chromosomes in a population.

```
function pop = mate(chsomes,nfinal)
% Mates chromosomes in a population

    ncross = 2;
    ninit = size(chsomes,1);
    nadj = nfinal+rem(nfinal,2);
    nrep = ceil(nadj/ninit);
    [k,k] = sort(rand([ninit,nrep]));
    k = k(1:nadj);
    pop = logical(zeros([nadj,size(chsomes,2)]));
    for i = 1:2:size(k,2)
        chsome1 = chsomes(k(i),:);
        chsome2 = chsomes(k(i+1),:);
        [x,y] = crossover(chsome1,chsome2,ncross);
        pop(i,:) = x;
        pop(i+1,:) = y;
    end
    pop = pop(1:nfinal,:);
end
```

`crossover` Performs a random crossover for a chromosome pair.

```
function [x,y] = crossover(a,b,n)
% Performs a random crossover for a chromosome pair

    % Validate inputs
    if length(a) ~= length(b)
        error('Chromosomes must have equal length');
    end

    % Can't have more crossovers than length of chromosome
    len = length(a);
    n = min(n,len);

    % Randomly select crossover points
    [i,i] = sort(rand([1,len]));
    pts = sort(i(1:n));
    pts = unique([pts,len]);

    [x,y] = deal(zeros(size(a)));
    i0 = 1;
    for i = 1:length(pts)
        x(i0:pts(i)) = a(i0:pts(i));
        y(i0:pts(i)) = b(i0:pts(i));
        tmp = x;
```

```
        x = y;
        y = tmp;
        i0 = pts(i)+1;
    end
end
```

`mutation` Performs random point mutations on a chromosome.

```
function x = mutation(a,p)
% Performs random point mutations on a chromosome

    x = (rand(size(a)) > p) == a;

end
```

D.5 Fitness Class

This file should be stored in the folder `metranCL/class/@fitness`.

`fitness` Constructor method for fitness object.

```
function obj = fitness(s,varargin)
% Constructor method for fitness object

    if isstruct(s) && isfield(s,'weight')
        obj = class(s,'fitness');
    else
        mod = s;
        enc = varargin{1};
        varargin = varargin(2:end);

        % Set parameters
        obj.alf = 0.05;
        obj.fast = true;
        obj.nest = 3;
        obj.nexp = 3;
        obj.msstd = 0.01;
        obj.msstdmin = 0.001;
        obj.flxstd = 0.05;
        obj.flxstdmin = 0.001;
        obj.sparse = true;

        % Initialize fields
        mod.options.hpc_on = false;
        mod.options.fit_reltol = 0.01;
        mod.options.fit_tau = 0.001;
        mod.options.fit_reinit = false;
```

```
        obj.hpc_on = false;
        obj.mod = mod;
        obj.mat = mod2mat(mod);
        obj.net = flx2net(mod.rates.flx.val',obj.mat);
        obj.weight = [];
        obj.encoder = enc;
        if rem(length(varargin),2) ~= 0
            error('Property names and values must be in pairs');
        end
        props = reshape(varargin,2,[])';

        % Parse through any property name/value pairs
        for i = 1:size(props,1)
            if ~isfield(obj,props{i,1})
                error('Invalid property name');
            else
                obj.(props{i,1}) = props{i,2};
            end
        end

        frees = [obj.mat.meta.free];
        if isempty(obj.weight)
            obj.weight = frees;
        elseif length(obj.weight) < length(frees) % || any(~islogical(obj.weight))
            temp = logical(zeros(size(frees)));
            temp(obj.weight) = true;
            obj.weight = temp;
        end
        obj.pars = logical(obj.weight);
        dir = cell2mat(mod.rates.flx.dir);
        dir = [dir(2:end)=='b',false];
        obj.netpars = dir(obj.pars);

        obj = class(obj,'fitness');

    end
end
```

D.6 Encoder Class

These files should be stored in the folder `metranCL/class/@encoder`.

`encoder` Constructor method for encoder object.

```
function obj = encoder(s)
% Constructor method for encoder object

    if isstruct(s) && isfield(s,'nbits')
        obj = class(s,'encoder');
```

```
    else

        tracers = s;

        % Set constants
        obj.nbits = 8;

        % Store information for each potential tracer
        tracers = tracersort(tracers);
        obj.tracers = tracers;
        obj.atoms = tracers.atoms.id;
        obj.states = tracers.state;
        len = length(tracers);

        % Make sure that no fields were empty
        if len ~= length(obj.atoms) || len ~= length(obj.states)
            error('Tracers must have defined atom ids and states.');
        end

        % Make sure that no duplicate tracers were given
        for i = 1:len
            if sum(strcmp(obj.atoms{i},obj.atoms) & ...
                   strcmp(obj.states{i},obj.states)) > 1
               error('Tracers must have unique atom ids and states.');
            end
        end

        % Unique list of states for natural abundance
        obj.natural = sort(unique(obj.states));

        obj = class(obj,'encoder');
    end
end
```

encode Method to encode a chromsome from a tracer mixture.

```
function chsome = encode(obj,tracer)
% Method to encode a chromosome from a multi-tracer. Tracers with natural
% abundance should be indicated with an atom id of 'na'.

    % Initialize chromosome
    chsome = logical(zeros([1,(length(obj.states)+length(obj.natural))*obj.nbits]));

    % Can't normalize if natural abundance frac isn't included in tracer object
    % Normalize tracer
    % tracer = normalize(obj,tracer);

    % Encode each tracer frac to chromosome
    zs = zeros(1,length(obj.natural));
    statefracs = zs;
    for i = 1:length(tracer)
        idx = strcmp(tracer(i).state,obj.states) & ...
```

```
            strcmp(tracer(i).atoms.id,obj.atoms);
        if isempty(idx)
            error('An input tracer was not pre-defined.');
        end

        stateidx = strmatch(obj.states{i},obj.natural);
        statefracs(stateidx) = statefracs(stateidx)+tracer(i).frac;

        gene = floor(tracer(i).frac*(2^obj.nbits-1));
        gene = dec2bin(gene,obj.nbits) == '1';
        idx = find(idx);
        range = (idx-1)*obj.nbits+1:idx*obj.nbits;
        chsome(range) = gene;
    end

    statenat = max([zs;1-statefracs]);
    for i = 1:length(obj.natural)
        gene = round(statenat(i)*(2^obj.nbits-1));
        gene = dec2bin(gene,obj.nbits) == '1';
        idx = length(tracer)+i;
        range = (idx-1)*obj.nbits+1:idx*obj.nbits;
        chsome(range) = gene;
    end

    chsome = normalize(obj,chsome);

end
```

decode Method to decode a tracer mixture from a chromosome.

```
function trcr = decode(obj,chsome)
% Method to decode a multi-tracer from a chromosome.

    chsome = normalize(obj,chsome);
    bins = reshape(chsome,obj.nbits,[])';
    fracs = bins*2.^(obj.nbits-1:-1:0)'/(2^obj.nbits-1);
    trcr = obj.tracers;
    trcr.frac = fracs(1:length(trcr))';

end
```

normalize Method to renormalize individual fractions of a tracer mixture.

```
function chsome = normalize(obj,chsome)
% Method to renormalize individual fractions of a multi-tracer.

    for i = 1:length(obj.natural)
        state = obj.natural{i};
        idx = [strmatch(state,obj.states)',length(obj.states)+i];
```

```
        bins = reshape(chsome,obj.nbits,[])';
        bins = bins(idx,:);
        if all(all(~bins))
            bins(end,:) = ~bins(end,:);
        end
        fracs = bins*2.^(obj.nbits-1:-1:0)';
        fracs = fracs/sum(fracs)*(2^obj.nbits-1);
        floored = floor(fracs);
        [j,j] = sort(fracs-floored,'descend');
        n = 2^obj.nbits-1-sum(floored);
        floored(j(1:n)) = floored(j(1:n))+1;
        for j = 1:length(idx)
            gene = dec2bin(floored(j),obj.nbits) == '1';
            range = (idx(j)-1)*obj.nbits+1:idx(j)*obj.nbits;
            chsome(range) = gene;
        end
    end

end
```

D.7 Common Files

These files `get.m`, `subsasgn.m`, and `subsref.m` should be stored in the folders `metranCL/class/@fitness` and `metranCL/class/@encoder`. The file `extract.m` should be stored in the folders `metranCL/class/@fitdata` and `metranCL/class/@parameter`.

`get` Get base object properties.

```
function varargout = get(x,varargin)
% BASE/GET Get base object properties.

    len = length(varargin);
    varargout = cell(1,len);
    for i = 1:len,
        index.type = '.';
        index.subs = varargin{i};
        varargout{i} = subsref(x,index);
    end
end
```

`subsasgn` Define index assignment for base objects.

```
function x = subsasgn(x,index,val)
```

```
%BASE/SUBSASGN Define index assignment for base objects.

% handle size increase
if strcmp(index(1).type,'()'),
    len = max(index(1).subs{1});
    if len > length(x), x(end+1:len) = feval(class(x)); end
end

% call subsasgn recursively
if length(index) > 1,
    val = subsasgn(subsref(x,index(1)),index(2:end),val);
end

% call class-specific method
x = subsa(x,index(1),val);

%-------------------------------------------------------------------------
% Class-specific subscript assignment
function x = subsa(x,index,val)

switch index.type
case '()',
    x(index.subs{:}) = val;
case '.',
    if isempty(x), return, end
    type = class(x);
    x = struct(x);
    if ischar(val),
        val = cellstr(val);
    elseif isnumeric(val) || islogical(val),
        len = length(x);
        if len > 1,
            val = num2cell(val,1:ndims(val)-1);
        else
            val = {val};
        end
    elseif isobject(val) || isstruct(val),
        val = obj2cell(val,x,index.subs);
    end
    if isfield(x,index.subs),
        [x.(index.subs)] = deal(val{:});
    else
        error('Field ''%s'' does not exist.',index.subs);
    end
    x = feval(type,x);
case '{}',
    i = cellndx(x,index.subs);
    x(i) = val;
end
```

`subsref` Define field name indexing for base objects.

```
function x = subsref(x,index)
%BASE/SUBSREF Define field name indexing for base objects.

% call class-specific subsref
x = subsr(x,index(1));

% call subsref recursively
if length(index) > 1,
    x = subsref(x,index(2:end));
end

%--------------------------------------------------------------------------
% Class-specific subscripted reference
function x = subsr(x,index)

switch index.type
case '()',
    x = x(index.subs{:});
case '.',
    cls = class(x);
    if length(x) > 0,
        t = x(1);
    else
        t = feval(cls); % get default template
    end
    t = struct(t);
    x = struct(x);
    if ischar(t.(index.subs)),
        x = {x.(index.subs)};
    else
        x = {x.(index.subs)};
        nrows = cellfun('size',x,1);
        if isempty(nrows) || all(nrows == max(nrows) | nrows == 0),
            x = [x{:}];
        end
    end
case '{}',
    i = cellndx(x,index.subs);
    x = x(i);
end
```

`extract` Extract fields from structures.

```
function t = extract(h,f)
% Extract fields from structures

t = {h.(f)};
```

Appendix E

Flux Analysis Results for *Y. lipolytica*

E.1 GC/MS Measurements

Four nonstationary metabolic flux experiments were conducted for *Y. lipolytica* as discussed in Chapter 8. Experiments L1, L2, H1, and H2 respectively refer to NMFA at low aeration and linear growth, low aeration and stationary growth, high aeration and linear growth, and high aeration and stationary growth. A combination of rapid and manual samples were taken in each case and assayed for intracellular isotopic metabolite labeling. These measurements are comprehensively listed in Tables E.1 (L1 part 1), E.2 (L1 part 2), E.3 (L2 part 1), E.4 (L2 part 2), E.5 (H1 part 1), E.6 (H1 part 2), E.7 (H2 part 1), and E.8 (H2 part 2).

E.2 Measurement Fits

Simulations using a model reaction network produced measurements *in silico* that were statistically comparable to those observed experimentally. These fits are shown in Figures E-1 (L1), E-2 (L2), E-3 (H1), and E-4 (H2). Figure 8.7 lists overall statistical parameters pertaining to each fit.

E.3 Flux Estimations

Fluxes were estimated for each reaction in the model reaction network described in Tables 8.5 and 8.6. We also calculated 95% confidence intervals for each net and exchange flux. The estimated values for each of these fluxes as well as the lower and upper bounds of the confidence intervals are shown in Tables E.9, E.10, E.11, and E.12.

Times	(h)	0.0006	0.0011	0.0022	0.0042	0.0083	0.0167	0.0667	0.2889	0.5167	1.0833	1.5167	2.0333	4.2167
Frag	Mass	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
Akg346	M+0	0.7179	0.7180	0.7174	0.7165	0.7165	0.7132	0.6833	0.5538	0.4721	0.3193	0.2765	0.2451	0.1984
Akg346	M+1	0.1892	0.1893	0.1895	0.1894	0.1898	0.1898	0.1970	0.2267	0.2428	0.2647	0.2673	0.2662	0.2612
Akg346	M+2	0.0774	0.0767	0.0776	0.0779	0.0774	0.0798	0.0942	0.1478	0.1775	0.2268	0.2383	0.2459	0.2567
Akg346	M+3	0.0131	0.0136	0.0131	0.0136	0.0138	0.0142	0.0199	0.0491	0.0711	0.1180	0.1336	0.1474	0.1679
Akg346	M+4	0.0024	0.0023	0.0024	0.0024	0.0026	0.0026	0.0049	0.0172	0.0273	0.0506	0.0590	0.0656	0.0789
Akg346	M+5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0046	0.0076	0.0163	0.0198	0.0229	0.0281
Ala232	M+0	0.7544	0.7498	0.7400	0.7413	0.7200	0.6815	0.5865	0.5186	0.5058	0.4804	0.4742	0.4694	0.4520
Ala232	M+1	0.1663	0.1678	0.1686	0.1697	0.1763	0.1861	0.2160	0.2421	0.2537	0.2602	0.2628	0.2663	0.2745
Ala232	M+2	0.0687	0.0705	0.0783	0.0759	0.0867	0.1082	0.1550	0.1869	0.1866	0.2006	0.2029	0.2047	0.2117
Ala232	M+3	0.0097	0.0104	0.0114	0.0111	0.0141	0.0192	0.0321	0.0402	0.0404	0.0444	0.0449	0.0450	0.0468
Ala232	M+4	0.0009	0.0015	0.0017	0.0019	0.0029	0.0049	0.0095	0.0116	0.0120	0.0130	0.0139	0.0134	0.0139
Ala260	M+0	0.7487	0.7418	0.7318	0.7330	0.7135	0.6715	0.5725	0.5019	0.4846	0.4613	0.4521	0.4474	0.4302
Ala260	M+1	0.1697	0.1723	0.1744	0.1741	0.1794	0.1899	0.2198	0.2432	0.2536	0.2570	0.2585	0.2636	0.2704
Ala260	M+2	0.0695	0.0705	0.0713	0.0704	0.0711	0.0751	0.0871	0.0983	0.1052	0.1075	0.1116	0.1131	0.1188
Ala260	M+3	0.0108	0.0134	0.0195	0.0191	0.0299	0.0511	0.0952	0.1243	0.1224	0.1374	0.1392	0.1385	0.1422
Ala260	M+4	0.0013	0.0019	0.0029	0.0034	0.0050	0.0094	0.0180	0.0235	0.0243	0.0260	0.0279	0.0269	0.0280
Ala260	M+5	0.0000	0.0000	0.0000	0.0000	0.0011	0.0030	0.0069	0.0086	0.0090	0.0104	0.0101	0.0101	0.0104
Asp302	M+0	0.7204	0.7195	0.7201	0.7198	0.7189	0.7183	0.7010	0.6314	0.5861	0.5018	0.4740	0.4533	0.4220
Asp302	M+1	0.1896	0.1900	0.1892	0.1897	0.1905	0.1899	0.1957	0.2217	0.2409	0.2797	0.2914	0.2996	0.3112
Asp302	M+2	0.0744	0.0749	0.0749	0.0748	0.0748	0.0757	0.0838	0.1141	0.1319	0.1625	0.1744	0.1824	0.1963
Asp302	M+3	0.0131	0.0130	0.0132	0.0132	0.0131	0.0135	0.0158	0.0254	0.0316	0.0423	0.0453	0.0487	0.0528
Asp302	M+4	0.0023	0.0024	0.0024	0.0023	0.0024	0.0024	0.0033	0.0064	0.0082	0.0115	0.0126	0.0134	0.0151
Asp376	M+0	0.6525	0.6522	0.6522	0.6521	0.6525	0.6526	0.6379	0.5740	0.5324	0.4558	0.4289	0.4096	0.3860
Asp376	M+1	0.2182	0.2167	0.2193	0.2190	0.2177	0.2178	0.2195	0.2400	0.2569	0.2869	0.2942	0.3052	0.3089
Asp376	M+2	0.1017	0.1028	0.1016	0.1018	0.1023	0.1023	0.1103	0.1362	0.1533	0.1804	0.1906	0.1977	0.2100
Asp376	M+3	0.0229	0.0230	0.0223	0.0223	0.0227	0.0224	0.0261	0.0380	0.0430	0.0565	0.0631	0.0640	0.0688
Asp376	M+4	0.0047	0.0054	0.0046	0.0048	0.0048	0.0048	0.0059	0.0106	0.0128	0.0170	0.0198	0.0198	0.0225
Asp390	M+0	0.6432	0.6426	0.6427	0.6430	0.6414	0.6417	0.6156	0.5170	0.4585	0.3583	0.3257	0.3029	0.2707
Asp390	M+1	0.2234	0.2231	0.2227	0.2228	0.2243	0.2238	0.2319	0.2660	0.2837	0.3093	0.3147	0.3182	0.3185
Asp390	M+2	0.1035	0.1044	0.1044	0.1041	0.1043	0.1039	0.1127	0.1459	0.1671	0.2038	0.2158	0.2245	0.2372
Asp390	M+3	0.0239	0.0237	0.0241	0.0239	0.0237	0.0242	0.0305	0.0519	0.0647	0.0905	0.1003	0.1072	0.1202
Asp390	M+4	0.0054	0.0056	0.0055	0.0055	0.0056	0.0056	0.0078	0.0154	0.0202	0.0285	0.0328	0.0354	0.0397
Asp390	M+5	0.0007	0.0007	0.0006	0.0008	0.0007	0.0008	0.0015	0.0036	0.0051	0.0080	0.0090	0.0098	0.0115
Asp418	M+0	0.6335	0.6327	0.6330	0.6330	0.6322	0.6305	0.6037	0.4986	0.4312	0.3222	0.2876	0.2637	0.2280
Asp418	M+1	0.2277	0.2280	0.2278	0.2274	0.2283	0.2282	0.2333	0.2524	0.2667	0.2834	0.2866	0.2860	0.2843
Asp418	M+2	0.1067	0.1069	0.1070	0.1072	0.1070	0.1077	0.1155	0.1560	0.1809	0.2139	0.2241	0.2326	0.2412
Asp418	M+3	0.0253	0.0255	0.0252	0.0252	0.0255	0.0262	0.0347	0.0624	0.0799	0.1144	0.1266	0.1353	0.1502
Asp418	M+4	0.0059	0.0060	0.0060	0.0061	0.0061	0.0062	0.0097	0.0224	0.0299	0.0469	0.0531	0.0580	0.0671
Asp418	M+5	0.0009	0.0009	0.0009	0.0010	0.0009	0.0011	0.0023	0.0063	0.0089	0.0144	0.0165	0.0184	0.0216
Cit459	M+0	0.6117	0.6113	0.6109	0.6095	0.6058	0.5969	0.5589	0.4746	0.4255	0.2863	0.2367	0.1986	0.1403
Cit459	M+1	0.2357	0.2354	0.2352	0.2358	0.2364	0.2374	0.2405	0.2445	0.2451	0.2418	0.2369	0.2316	0.2177
Cit459	M+2	0.1144	0.1149	0.1153	0.1158	0.1175	0.1219	0.1384	0.1697	0.1852	0.2224	0.2327	0.2395	0.2453
Cit459	M+3	0.0291	0.0292	0.0294	0.0295	0.0305	0.0325	0.0429	0.0682	0.0845	0.1339	0.1527	0.1682	0.1926
Cit459	M+4	0.0075	0.0075	0.0077	0.0077	0.0081	0.0091	0.0143	0.0290	0.0391	0.0712	0.0850	0.0963	0.1178
Cit459	M+5	0.0014	0.0014	0.0014	0.0014	0.0015	0.0019	0.0038	0.0099	0.0142	0.0299	0.0372	0.0435	0.0561
Cit459	M+6	0.0002	0.0002	0.0002	0.0002	0.0003	0.0004	0.0010	0.0031	0.0047	0.0106	0.0136	0.0162	0.0216
Glu330	M+0	0.7012	0.7009	0.7011	0.7005	0.7005	0.6986	0.6692	0.5466	0.4740	0.3279	0.2860	0.2580	0.2194
Glu330	M+1	0.2030	0.2030	0.2030	0.2034	0.2031	0.2037	0.2114	0.2450	0.2641	0.2952	0.2996	0.3014	0.2989
Glu330	M+2	0.0783	0.0784	0.0783	0.0784	0.0787	0.0795	0.0936	0.1453	0.1723	0.2233	0.2377	0.2475	0.2617
Glu330	M+3	0.0146	0.0147	0.0147	0.0147	0.0148	0.0151	0.0203	0.0457	0.0634	0.1049	0.1192	0.1290	0.1445
Glu330	M+4	0.0026	0.0026	0.0026	0.0026	0.0026	0.0028	0.0047	0.0139	0.0204	0.0371	0.0436	0.0484	0.0568
Glu330	M+5	0.0003	0.0003	0.0003	0.0004	0.0003	0.0004	0.0008	0.0029	0.0047	0.0094	0.0111	0.0125	0.0148
Glu432	M+0	0.6172	0.6176	0.6171	0.6169	0.6170	0.6150	0.5890	0.4780	0.4106	0.2725	0.2323	0.2056	0.1688
Glu432	M+1	0.2351	0.2349	0.2351	0.2355	0.2351	0.2353	0.2381	0.2498	0.2555	0.2594	0.2564	0.2530	0.2443
Glu432	M+2	0.1113	0.1110	0.1113	0.1111	0.1113	0.1124	0.1248	0.1702	0.1935	0.2330	0.2426	0.2480	0.2540
Glu432	M+3	0.0275	0.0274	0.0275	0.0274	0.0276	0.0280	0.0344	0.0657	0.0868	0.1359	0.1520	0.1635	0.1807
Glu432	M+4	0.0066	0.0066	0.0066	0.0066	0.0066	0.0069	0.0098	0.0252	0.0364	0.0644	0.0747	0.0823	0.0949
Glu432	M+5	0.0018	0.0019	0.0018	0.0019	0.0018	0.0019	0.0029	0.0082	0.0126	0.0249	0.0298	0.0337	0.0402
Glu432	M+6	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0007	0.0022	0.0035	0.0074	0.0090	0.0102	0.0125
Gly218	M+0	0.7681	0.7675	0.7709	0.7694	0.7691	0.7698	0.7560	0.7336	0.7260	0.7121	0.7024	0.6986	0.6733
Gly218	M+1	0.1568	0.1584	0.1585	0.1563	0.1590	0.1584	0.1689	0.1858	0.1938	0.2034	0.2122	0.2159	0.2333
Gly218	M+2	0.0671	0.0673	0.0638	0.0663	0.0645	0.0647	0.0665	0.0705	0.0694	0.0723	0.0732	0.0739	0.0779
Gly218	M+3	0.0080	0.0068	0.0068	0.0080	0.0075	0.0071	0.0087	0.0101	0.0108	0.0122	0.0122	0.0116	0.0155
Gly246	M+0	0.7573	0.7589	0.7623	0.7606	0.7619	0.7602	0.7429	0.7144	0.7098	0.6827	0.6763	0.6692	0.6398
Gly246	M+1	0.1669	0.1638	0.1651	0.1630	0.1640	0.1609	0.1659	0.1752	0.1778	0.1894	0.1901	0.1975	0.2129
Gly246	M+2	0.0695	0.0687	0.0664	0.0679	0.0661	0.0704	0.0799	0.0966	0.0964	0.1080	0.1088	0.1124	0.1218
Gly246	M+3	0.0063	0.0086	0.0062	0.0085	0.0080	0.0085	0.0110	0.0130	0.0151	0.0182	0.0204	0.0184	0.0236

Table E.1: Mass isotopomer distributions for NMFA experiment L1 (part 1). Each column represents a different time point and each row represents a different metabolite mass isotopomer. The time points are listed in the first row (in hours).

Times	(h)	0.0006	0.0011	0.0022	0.0042	0.0083	0.0167	0.0667	0.2889	0.5167	1.0833	1.5167	2.0333	4.2167
Frag	Mass	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
Ile200	M+0	0.8115	0.8102	0.8111	0.8096	0.8093	0.8107	0.7950	0.7823	0.7812	0.7656	0.7645	0.7566	0.7349
Ile200	M+1	0.1438	0.1443	0.1428	0.1441	0.1445	0.1418	0.1487	0.1545	0.1536	0.1604	0.1575	0.1612	0.1679
Ile200	M+2	0.0411	0.0417	0.0423	0.0417	0.0416	0.0430	0.0503	0.0561	0.0566	0.0631	0.0633	0.0648	0.0744
Ile200	M+3	0.0036	0.0038	0.0038	0.0046	0.0046	0.0045	0.0060	0.0072	0.0085	0.0109	0.0127	0.0148	0.0183
Ile274	M+0	0.7409	0.7332	0.7392	0.7315	0.7420	0.7298	0.7175	0.7056	0.7072	0.6970	0.6926	0.6882	0.6681
Ile274	M+1	0.1854	0.1854	0.1840	0.1877	0.1781	0.1909	0.1893	0.1993	0.1947	0.1964	0.1991	0.2016	0.2109
Ile274	M+2	0.0664	0.0726	0.0692	0.0709	0.0702	0.0718	0.0811	0.0839	0.0837	0.0868	0.0895	0.0895	0.0963
Ile274	M+3	0.0074	0.0088	0.0076	0.0099	0.0097	0.0075	0.0120	0.0112	0.0144	0.0198	0.0188	0.0208	0.0247
Leu274	M+0	0.7343	0.7326	0.7351	0.7321	0.7353	0.7330	0.7273	0.7009	0.6935	0.6796	0.6716	0.6663	0.6385
Leu274	M+1	0.1840	0.1831	0.1865	0.1849	0.1833	0.1853	0.1877	0.1992	0.2012	0.2050	0.2083	0.2080	0.2159
Leu274	M+2	0.0712	0.0733	0.0691	0.0718	0.0700	0.0709	0.0721	0.0807	0.0850	0.0864	0.0884	0.0915	0.1010
Leu274	M+3	0.0105	0.0109	0.0093	0.0112	0.0115	0.0108	0.0125	0.0174	0.0170	0.0233	0.0250	0.0269	0.0333
Leu274	M+4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0018	0.0033	0.0057	0.0067	0.0073	0.0105
Mal419	M+0	0.6354	0.6346	0.6344	0.6337	0.6325	0.6231	0.5888	0.4954	0.4480	0.3239	0.2855	0.2613	0.2255
Mal419	M+1	0.2257	0.2254	0.2260	0.2255	0.2254	0.2282	0.2339	0.2523	0.2624	0.2828	0.2858	0.2861	0.2855
Mal419	M+2	0.1066	0.1071	0.1068	0.1074	0.1072	0.1095	0.1211	0.1563	0.1726	0.2132	0.2247	0.2319	0.2407
Mal419	M+3	0.0253	0.0257	0.0258	0.0258	0.0269	0.0297	0.0398	0.0639	0.0767	0.1136	0.1268	0.1358	0.1503
Mal419	M+4	0.0060	0.0061	0.0061	0.0064	0.0067	0.0076	0.0123	0.0235	0.0290	0.0471	0.0542	0.0595	0.0683
Mal419	M+5	0.0009	0.0010	0.0009	0.0010	0.0012	0.0016	0.0032	0.0067	0.0086	0.0146	0.0171	0.0188	0.0220
Pyr174	M+0	0.8605	0.8539	0.8417	0.8403	0.8116	0.7642	0.6537	0.6465	0.6174	0.5512	0.5379	0.5234	0.4935
Pyr174	M+1	0.1026	0.1100	0.1191	0.1127	0.1267	0.1447	0.1828	0.2000	0.2154	0.2335	0.2356	0.2472	0.2595
Pyr174	M+2	0.0369	0.0350	0.0353	0.0335	0.0345	0.0423	0.0573	0.0611	0.0725	0.0820	0.0808	0.0879	0.0983
Pyr174	M+3	0.0000	0.0012	0.0039	0.0135	0.0273	0.0489	0.0988	0.0924	0.0947	0.1241	0.1318	0.1318	0.1364
Pyr174	M+4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0067	0.0000	0.0000	0.0092	0.0114	0.0096	0.0115
Ser288	M+0	0.7265	0.7245	0.7248	0.7241	0.7214	0.7141	0.6794	0.6038	0.5973	0.5832	0.5766	0.5696	0.5460
Ser288	M+1	0.1882	0.1894	0.1893	0.1888	0.1894	0.1930	0.2069	0.2490	0.2576	0.2647	0.2717	0.2753	0.2908
Ser288	M+2	0.0716	0.0717	0.0728	0.0724	0.0738	0.0769	0.0918	0.1144	0.1125	0.1175	0.1170	0.1193	0.1238
Ser288	M+3	0.0119	0.0127	0.0117	0.0129	0.0134	0.0136	0.0182	0.0267	0.0267	0.0279	0.0281	0.0295	0.0317
Ser288	M+4	0.0018	0.0018	0.0013	0.0018	0.0020	0.0023	0.0037	0.0059	0.0059	0.0062	0.0063	0.0063	0.0070
Ser302	M+0	0.7174	0.7198	0.7201	0.7200	0.7172	0.7150	0.6955	0.6538	0.6478	0.6354	0.6314	0.6248	0.6076
Ser302	M+1	0.1888	0.1861	0.1847	0.1864	0.1863	0.1868	0.1881	0.1928	0.1956	0.2013	0.2049	0.2075	0.2152
Ser302	M+2	0.0788	0.0782	0.0797	0.0781	0.0808	0.0812	0.0944	0.1219	0.1228	0.1266	0.1270	0.1296	0.1352
Ser302	M+3	0.0134	0.0136	0.0134	0.0133	0.0136	0.0141	0.0177	0.0250	0.0264	0.0283	0.0281	0.0293	0.0318
Ser302	M+4	0.0016	0.0023	0.0022	0.0022	0.0021	0.0029	0.0043	0.0065	0.0075	0.0081	0.0085	0.0086	0.0095
Ser362	M+0	0.6549	0.6526	0.6528	0.6535	0.6492	0.6463	0.6133	0.5429	0.5398	0.5245	0.5184	0.5146	0.4898
Ser362	M+1	0.2159	0.2180	0.2177	0.2184	0.2207	0.2187	0.2312	0.2644	0.2682	0.2781	0.2830	0.2842	0.2994
Ser362	M+2	0.1023	0.1021	0.1022	0.1010	0.1024	0.1052	0.1171	0.1397	0.1405	0.1421	0.1426	0.1441	0.1495
Ser362	M+3	0.0221	0.0222	0.0225	0.0221	0.0228	0.0237	0.0295	0.0407	0.0399	0.0418	0.0425	0.0436	0.0469
Ser362	M+4	0.0049	0.0050	0.0047	0.0048	0.0051	0.0055	0.0077	0.0107	0.0102	0.0113	0.0116	0.0118	0.0122
Ser390	M+0	0.6449	0.6427	0.6446	0.6425	0.6400	0.6361	0.6019	0.5329	0.5227	0.5084	0.4997	0.4901	0.4652
Ser390	M+1	0.2235	0.2232	0.2236	0.2243	0.2242	0.2243	0.2328	0.2526	0.2567	0.2622	0.2662	0.2712	0.2794
Ser390	M+2	0.1031	0.1042	0.1029	0.1034	0.1037	0.1041	0.1111	0.1311	0.1350	0.1401	0.1430	0.1460	0.1538
Ser390	M+3	0.0233	0.0241	0.0235	0.0240	0.0259	0.0279	0.0406	0.0601	0.0625	0.0643	0.0655	0.0668	0.0728
Ser390	M+4	0.0052	0.0055	0.0054	0.0056	0.0059	0.0067	0.0107	0.0180	0.0181	0.0188	0.0193	0.0195	0.0216
Suc289	M+0	0.7388	0.7381	0.7366	0.7377	0.7372	0.7289	0.6756	0.5457	0.4899	0.3350	0.2981	0.2733	0.2375
Suc289	M+1	0.1746	0.1747	0.1755	0.1753	0.1750	0.1778	0.1937	0.2353	0.2530	0.2939	0.2987	0.3012	0.3000
Suc289	M+2	0.0731	0.0738	0.0744	0.0734	0.0740	0.0779	0.1028	0.1546	0.1719	0.2238	0.2369	0.2463	0.2587
Suc289	M+3	0.0113	0.0114	0.0116	0.0117	0.0118	0.0129	0.0213	0.0462	0.0606	0.1015	0.1135	0.1220	0.1366
Suc289	M+4	0.0021	0.0020	0.0018	0.0019	0.0021	0.0024	0.0058	0.0154	0.0200	0.0360	0.0411	0.0445	0.0520
Suc289	M+5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0008	0.0029	0.0040	0.0084	0.0096	0.0105	0.0125
Thr376	M+0	0.6541	0.6494	0.6522	0.6510	0.6550	0.6565	0.6489	0.6303	0.6190	0.5788	0.5649	0.5474	0.5104
Thr376	M+1	0.2245	0.2238	0.2265	0.2250	0.2230	0.2206	0.2280	0.2340	0.2374	0.2451	0.2485	0.2567	0.2648
Thr376	M+2	0.1009	0.1054	0.1015	0.1007	0.1019	0.1028	0.1014	0.1093	0.1144	0.1288	0.1326	0.1369	0.1538
Thr376	M+3	0.0205	0.0215	0.0198	0.0216	0.0201	0.0201	0.0203	0.0254	0.0282	0.0388	0.0445	0.0483	0.0571
Thr376	M+4	0.0000	0.0000	0.0000	0.0016	0.0000	0.0000	0.0013	0.0010	0.0010	0.0085	0.0096	0.0107	0.0139
Thr404	M+0	0.6433	0.6454	0.6436	0.6416	0.6372	0.6400	0.6382	0.6199	0.6038	0.5601	0.5458	0.5258	0.4853
Thr404	M+1	0.2257	0.2253	0.2263	0.2268	0.2298	0.2282	0.2290	0.2304	0.2363	0.2430	0.2463	0.2495	0.2508
Thr404	M+2	0.1047	0.1035	0.1062	0.1044	0.1074	0.1040	0.1056	0.1139	0.1183	0.1334	0.1364	0.1426	0.1588
Thr404	M+3	0.0241	0.0232	0.0225	0.0232	0.0227	0.0242	0.0233	0.0288	0.0329	0.0470	0.0508	0.0573	0.0723
Thr404	M+4	0.0022	0.0026	0.0014	0.0040	0.0029	0.0036	0.0039	0.0070	0.0087	0.0146	0.0169	0.0209	0.0266
Val260	M+0	0.7362	0.7363	0.7365	0.7358	0.7361	0.7324	0.7013	0.6403	0.6153	0.5943	0.5850	0.5694	0.5350
Val260	M+1	0.1827	0.1820	0.1818	0.1807	0.1820	0.1816	0.1870	0.1983	0.2024	0.2060	0.2059	0.2090	0.2172
Val260	M+2	0.0705	0.0705	0.0703	0.0722	0.0707	0.0725	0.0873	0.1139	0.1240	0.1315	0.1359	0.1417	0.1559
Val260	M+3	0.0104	0.0107	0.0110	0.0109	0.0112	0.0123	0.0188	0.0337	0.0400	0.0461	0.0500	0.0541	0.0618
Val260	M+4	0.0003	0.0005	0.0003	0.0005	0.0000	0.0012	0.0056	0.0128	0.0159	0.0189	0.0202	0.0221	0.0260
Val288	M+0	0.7317	0.7332	0.7307	0.7325	0.7326	0.7277	0.6958	0.6403	0.6121	0.5901	0.5798	0.5660	0.5265
Val288	M+1	0.1852	0.1836	0.1837	0.1827	0.1831	0.1832	0.1891	0.1971	0.2042	0.2039	0.2058	0.2104	0.2153
Val288	M+2	0.0717	0.0721	0.0739	0.0725	0.0722	0.0744	0.0831	0.0996	0.1057	0.1130	0.1165	0.1182	0.1324
Val288	M+3	0.0110	0.0107	0.0113	0.0116	0.0118	0.0131	0.0235	0.0415	0.0493	0.0574	0.0602	0.0656	0.0756
Val288	M+4	0.0004	0.0003	0.0004	0.0007	0.0004	0.0016	0.0062	0.0146	0.0195	0.0224	0.0240	0.0265	0.0336
Val288	M+5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0023	0.0069	0.0092	0.0118	0.0131	0.0132	0.0164

Table E.2: Mass isotopomer distributions for NMFA experiment L1 (part 2). Each column represents a different time point and each row represents a different metabolite mass isotopomer. The time points are listed in the first row (in hours).

Times	(h)	0.0067	0.0108	0.0192	0.0358	0.1819	0.2948	0.5089	1.0278	1.4667	1.9500	2.5167	3.9167
Frag	Mass	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
Akg346	M+0	0.7162	0.7158	0.7097	0.7038	0.5959	0.5228	0.4244	0.2957	0.2472	0.2222	0.2042	0.1840
Akg346	M+1	0.1894	0.1888	0.1918	0.1918	0.2145	0.2259	0.2364	0.2487	0.2480	0.2473	0.2468	0.2447
Akg346	M+2	0.0778	0.0785	0.0805	0.0837	0.1239	0.1501	0.1836	0.2232	0.2370	0.2435	0.2473	0.2532
Akg346	M+3	0.0138	0.0141	0.0144	0.0163	0.0432	0.0632	0.0946	0.1356	0.1548	0.1643	0.1722	0.1807
Akg346	M+4	0.0025	0.0026	0.0028	0.0037	0.0159	0.0261	0.0414	0.0644	0.0751	0.0811	0.0862	0.0896
Akg346	M+5	0.0000	0.0000	0.0000	0.0005	0.0050	0.0090	0.0151	0.0245	0.0288	0.0315	0.0333	0.0363
Akg346	M+6	0.0000	0.0000	0.0000	0.0000	0.0011	0.0020	0.0035	0.0060	0.0069	0.0076	0.0081	0.0090
Ala232	M+0	0.7320	0.7061	0.6658	0.5858	0.4428	0.4326	0.4225	0.4133	0.4077	0.4079	0.4023	0.4008
Ala232	M+1	0.1740	0.1819	0.1968	0.2237	0.2785	0.2834	0.2888	0.2940	0.2950	0.2958	0.2980	0.2974
Ala232	M+2	0.0795	0.0927	0.1127	0.1496	0.2147	0.2183	0.2219	0.2246	0.2286	0.2271	0.2298	0.2303
Ala232	M+3	0.0122	0.0158	0.0197	0.0314	0.0479	0.0491	0.0500	0.0507	0.0513	0.0516	0.0524	0.0538
Ala232	M+4	0.0023	0.0035	0.0051	0.0088	0.0143	0.0147	0.0152	0.0155	0.0156	0.0158	0.0156	0.0157
Ala260	M+0	0.7256	0.6985	0.6568	0.5721	0.4219	0.4152	0.4031	0.3932	0.3874	0.3878	0.3843	0.3788
Ala260	M+1	0.1777	0.1855	0.2005	0.2280	0.2758	0.2769	0.2822	0.2841	0.2859	0.2857	0.2875	0.2887
Ala260	M+2	0.0708	0.0721	0.0770	0.0879	0.1138	0.1148	0.1169	0.1213	0.1223	0.1230	0.1225	0.1252
Ala260	M+3	0.0215	0.0355	0.0523	0.0882	0.1471	0.1514	0.1540	0.1572	0.1595	0.1589	0.1602	0.1615
Ala260	M+4	0.0038	0.0065	0.0100	0.0169	0.0293	0.0293	0.0308	0.0313	0.0315	0.0313	0.0322	0.0324
Ala260	M+5	0.0005	0.0020	0.0034	0.0065	0.0109	0.0112	0.0120	0.0116	0.0119	0.0121	0.0122	0.0122
Asp302	M+0	0.7191	0.7153	0.7038	0.6741	0.5545	0.5296	0.5022	0.4713	0.4582	0.4525	0.4453	0.4394
Asp302	M+1	0.1887	0.1896	0.1918	0.2000	0.2424	0.2515	0.2612	0.2717	0.2765	0.2781	0.2811	0.2831
Asp302	M+2	0.0761	0.0785	0.0846	0.0998	0.1531	0.1639	0.1763	0.1905	0.1965	0.1992	0.2024	0.2047
Asp302	M+3	0.0135	0.0138	0.0160	0.0205	0.0373	0.0409	0.0446	0.0492	0.0507	0.0516	0.0525	0.0534
Asp302	M+4	0.0025	0.0026	0.0034	0.0050	0.0109	0.0118	0.0131	0.0144	0.0153	0.0154	0.0156	0.0162
Asp376	M+0	0.6537	0.6501	0.6395	0.6125	0.5053	0.4834	0.4581	0.4273	0.4150	0.4083	0.4058	0.4001
Asp376	M+1	0.2163	0.2171	0.2197	0.2249	0.2553	0.2608	0.2660	0.2760	0.2818	0.2828	0.2861	0.2885
Asp376	M+2	0.1023	0.1046	0.1097	0.1235	0.1714	0.1799	0.1943	0.2058	0.2102	0.2136	0.2136	0.2165
Asp376	M+3	0.0225	0.0236	0.0255	0.0312	0.0500	0.0555	0.0599	0.0649	0.0676	0.0683	0.0688	0.0687
Asp376	M+4	0.0053	0.0047	0.0056	0.0078	0.0160	0.0180	0.0189	0.0222	0.0218	0.0237	0.0227	0.0225
Asp390	M+0	0.6419	0.6359	0.6156	0.5722	0.4080	0.3746	0.3403	0.2992	0.2843	0.2767	0.2706	0.2642
Asp390	M+1	0.2241	0.2237	0.2295	0.2397	0.2795	0.2869	0.2955	0.3007	0.3016	0.3038	0.3047	0.3041
Asp390	M+2	0.1032	0.1073	0.1145	0.1298	0.1832	0.1958	0.2068	0.2235	0.2298	0.2321	0.2336	0.2353
Asp390	M+3	0.0245	0.0266	0.0310	0.0434	0.0909	0.0987	0.1093	0.1219	0.1265	0.1284	0.1307	0.1341
Asp390	M+4	0.0057	0.0061	0.0082	0.0121	0.0284	0.0326	0.0353	0.0398	0.0424	0.0428	0.0444	0.0455
Asp390	M+5	0.0007	0.0004	0.0012	0.0026	0.0086	0.0096	0.0108	0.0127	0.0132	0.0133	0.0135	0.0142
Asp418	M+0	0.6300	0.6250	0.6070	0.5592	0.3850	0.3495	0.3114	0.2683	0.2523	0.2431	0.2367	0.2316
Asp418	M+1	0.2285	0.2279	0.2329	0.2433	0.2721	0.2782	0.2819	0.2842	0.2846	0.2840	0.2830	0.2834
Asp418	M+2	0.1066	0.1076	0.1087	0.1168	0.1591	0.1696	0.1830	0.1989	0.2038	0.2066	0.2091	0.2096
Asp418	M+3	0.0273	0.0304	0.0375	0.0555	0.1135	0.1237	0.1344	0.1474	0.1520	0.1563	0.1584	0.1610
Asp418	M+4	0.0065	0.0076	0.0104	0.0175	0.0478	0.0538	0.0609	0.0691	0.0735	0.0749	0.0768	0.0776
Asp418	M+5	0.0012	0.0015	0.0029	0.0059	0.0168	0.0186	0.0210	0.0233	0.0249	0.0255	0.0262	0.0268
Asp418	M+6	0.0000	0.0000	0.0005	0.0013	0.0044	0.0052	0.0059	0.0068	0.0073	0.0075	0.0078	0.0077
Cit459	M+0	0.6117	0.6108	0.6102	0.6082	0.5934	0.5808	0.5597	0.5106	0.4813	0.4478	0.4261	0.3667
Cit459	M+1	0.2353	0.2356	0.2355	0.2359	0.2363	0.2362	0.2358	0.2342	0.2317	0.2294	0.2272	0.2218
Cit459	M+2	0.1146	0.1150	0.1152	0.1157	0.1201	0.1237	0.1295	0.1418	0.1489	0.1566	0.1613	0.1750
Cit459	M+3	0.0292	0.0294	0.0296	0.0302	0.0349	0.0390	0.0460	0.0627	0.0732	0.0851	0.0929	0.1141
Cit459	M+4	0.0076	0.0076	0.0077	0.0080	0.0109	0.0137	0.0183	0.0301	0.0377	0.0463	0.0523	0.0679
Cit459	M+5	0.0014	0.0014	0.0015	0.0016	0.0032	0.0046	0.0072	0.0135	0.0176	0.0224	0.0258	0.0348
Cit459	M+6	0.0003	0.0003	0.0003	0.0003	0.0009	0.0015	0.0025	0.0051	0.0068	0.0087	0.0102	0.0139
Glu330	M+0	0.7017	0.7002	0.6970	0.6872	0.5819	0.5137	0.4202	0.3027	0.2588	0.2364	0.2199	0.2047
Glu330	M+1	0.2023	0.2030	0.2038	0.2063	0.2282	0.2407	0.2558	0.2698	0.2724	0.2733	0.2739	0.2737
Glu330	M+2	0.0782	0.0787	0.0803	0.0847	0.1284	0.1554	0.1916	0.2359	0.2518	0.2596	0.2648	0.2707
Glu330	M+3	0.0147	0.0150	0.0155	0.0174	0.0423	0.0602	0.0863	0.1224	0.1377	0.1456	0.1518	0.1578
Glu330	M+4	0.0026	0.0027	0.0029	0.0037	0.0149	0.0230	0.0351	0.0521	0.0595	0.0636	0.0668	0.0696
Glu330	M+5	0.0004	0.0004	0.0004	0.0006	0.0034	0.0056	0.0088	0.0135	0.0155	0.0167	0.0176	0.0184
Glu432	M+0	0.6102	0.6076	0.6069	0.5963	0.5020	0.4403	0.3553	0.2470	0.2065	0.1857	0.1697	0.1558
Glu432	M+1	0.2314	0.2305	0.2318	0.2321	0.2394	0.2416	0.2425	0.2376	0.2340	0.2312	0.2281	0.2251
Glu432	M+2	0.1092	0.1090	0.1108	0.1145	0.1482	0.1687	0.1948	0.2244	0.2338	0.2390	0.2414	0.2448
Glu432	M+3	0.0269	0.0270	0.0279	0.0301	0.0596	0.0800	0.1095	0.1496	0.1655	0.1745	0.1806	0.1868
Glu432	M+4	0.0065	0.0066	0.0070	0.0080	0.0239	0.0357	0.0532	0.0785	0.0890	0.0955	0.1000	0.1041
Glu432	M+5	0.0102	0.0125	0.0101	0.0123	0.0180	0.0229	0.0306	0.0431	0.0488	0.0509	0.0550	0.0571
Glu432	M+6	0.0035	0.0043	0.0035	0.0042	0.0059	0.0074	0.0098	0.0139	0.0158	0.0164	0.0178	0.0186
Gly218	M+0	0.7825	0.7864	0.7791	0.7736	0.7622	0.7734	0.7470	0.7484	0.7526	0.7395	0.7390	0.7378
Gly218	M+1	0.1521	0.1555	0.1581	0.1641	0.1730	0.1642	0.1851	0.1860	0.1834	0.1900	0.1955	0.2003
Gly218	M+2	0.0653	0.0580	0.0628	0.0623	0.0648	0.0624	0.0678	0.0657	0.0640	0.0635	0.0655	0.0619
Gly218	M+3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0071	0.0000	0.0000
Gly246	M+0	0.7746	0.7741	0.7682	0.7701	0.7508	0.7456	0.7345	0.7227	0.7427	0.7244	0.7168	0.7124
Gly246	M+1	0.1567	0.1659	0.1678	0.1627	0.1718	0.1728	0.1768	0.1824	0.1690	0.1746	0.1833	0.1818
Gly246	M+2	0.0687	0.0600	0.0639	0.0672	0.0757	0.0769	0.0848	0.0891	0.0860	0.0923	0.0918	0.1031
Gly246	M+3	0.0000	0.0000	0.0000	0.0000	0.0017	0.0047	0.0039	0.0057	0.0023	0.0087	0.0081	0.0027

Table E.3: Mass isotopomer distributions for NMFA experiment L2 (part 1). Each column represents a different time point and each row represents a different metabolite mass isotopomer. The time points are listed in the first row (in hours).

Times	(h)	0.0067	0.0108	0.0192	0.0358	0.1819	0.2948	0.5089	1.0278	1.4667	1.9500	2.5167	3.9167
Frag	Mass	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
Ile200	M+0	0.8083	0.8108	0.8059	0.8053	0.7797	0.7775	0.7670	0.7660	0.7549	0.7533	0.7547	0.7437
Ile200	M+1	0.1475	0.1460	0.1490	0.1466	0.1550	0.1567	0.1622	0.1602	0.1659	0.1635	0.1653	0.1661
Ile200	M+2	0.0431	0.0406	0.0428	0.0441	0.0550	0.0575	0.0602	0.0609	0.0659	0.0668	0.0638	0.0674
Ile200	M+3	0.0007	0.0015	0.0008	0.0040	0.0071	0.0079	0.0100	0.0125	0.0126	0.0136	0.0148	0.0162
Ile274	M+0	0.7337	0.7403	0.7364	0.7209	0.7065	0.6983	0.6898	0.6898	0.6969	0.6906	0.6953	0.6769
Ile274	M+1	0.1878	0.1888	0.1971	0.1902	0.1975	0.2032	0.2080	0.2024	0.1997	0.1974	0.1959	0.2025
Ile274	M+2	0.0673	0.0674	0.0665	0.0765	0.0839	0.0810	0.0845	0.0889	0.0900	0.0913	0.0883	0.0925
Ile274	M+3	0.0066	0.0035	0.0000	0.0125	0.0122	0.0174	0.0177	0.0189	0.0134	0.0207	0.0205	0.0264
Leu274	M+0	0.7328	0.7323	0.7348	0.7320	0.7057	0.6973	0.6818	0.6792	0.6727	0.6712	0.6683	0.6650
Leu274	M+1	0.1867	0.1870	0.1855	0.1870	0.2012	0.2052	0.2060	0.2062	0.2093	0.2074	0.2121	0.2115
Leu274	M+2	0.0715	0.0714	0.0708	0.0702	0.0790	0.0783	0.0858	0.0863	0.0860	0.0898	0.0886	0.0906
Leu274	M+3	0.0089	0.0094	0.0089	0.0109	0.0141	0.0182	0.0199	0.0218	0.0260	0.0244	0.0252	0.0264
Leu274	M+4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0011	0.0037	0.0062	0.0060	0.0072	0.0059	0.0065
Mal419	M+0	0.6258	0.6165	0.5889	0.5492	0.4174	0.3824	0.3375	0.2846	0.2644	0.2521	0.2434	0.2369
Mal419	M+1	0.2272	0.2302	0.2364	0.2436	0.2654	0.2705	0.2766	0.2823	0.2819	0.2824	0.2819	0.2814
Mal419	M+2	0.1077	0.1079	0.1127	0.1211	0.1541	0.1660	0.1802	0.1966	0.2044	0.2082	0.2093	0.2125
Mal419	M+3	0.0298	0.0340	0.0441	0.0584	0.1009	0.1109	0.1251	0.1406	0.1472	0.1520	0.1556	0.1582
Mal419	M+4	0.0076	0.0088	0.0130	0.0193	0.0424	0.0482	0.0553	0.0653	0.0698	0.0718	0.0748	0.0752
Mal419	M+5	0.0016	0.0022	0.0041	0.0064	0.0149	0.0162	0.0186	0.0223	0.0238	0.0244	0.0255	0.0262
Mal419	M+6	0.0002	0.0004	0.0008	0.0018	0.0041	0.0047	0.0056	0.0067	0.0071	0.0075	0.0077	0.0078
Pyr174	M+0	0.8081	0.7955	0.7351	0.6635	0.5067	0.4922	0.4762	0.4568	0.4487	0.4461	0.4470	0.4310
Pyr174	M+1	0.1325	0.1338	0.1652	0.1938	0.2578	0.2672	0.2705	0.2764	0.2842	0.2798	0.2851	0.2881
Pyr174	M+2	0.0407	0.0401	0.0485	0.0546	0.0823	0.0818	0.0909	0.0957	0.0965	0.0996	0.0947	0.1033
Pyr174	M+3	0.0187	0.0306	0.0512	0.0864	0.1431	0.1509	0.1512	0.1594	0.1593	0.1602	0.1639	0.1602
Pyr174	M+4	0.0000	0.0000	0.0000	0.0017	0.0095	0.0079	0.0104	0.0110	0.0105	0.0130	0.0094	0.0135
Ser288	M+0	0.7237	0.7240	0.7164	0.7156	0.6568	0.6351	0.6098	0.5917	0.5783	0.5759	0.5634	0.5474
Ser288	M+1	0.1909	0.1889	0.1930	0.1928	0.2227	0.2375	0.2499	0.2647	0.2743	0.2747	0.2830	0.2921
Ser288	M+2	0.0726	0.0726	0.0765	0.0766	0.0966	0.1003	0.1099	0.1122	0.1144	0.1149	0.1183	0.1240
Ser288	M+3	0.0120	0.0126	0.0129	0.0130	0.0195	0.0223	0.0250	0.0252	0.0272	0.0288	0.0291	0.0303
Ser302	M+0	0.7142	0.7178	0.7144	0.7085	0.6814	0.6687	0.6550	0.6486	0.6420	0.6395	0.6357	0.6290
Ser302	M+1	0.1857	0.1862	0.1855	0.1857	0.1924	0.1909	0.1951	0.1960	0.1984	0.1983	0.2023	0.2052
Ser302	M+2	0.0833	0.0802	0.0858	0.0858	0.1015	0.1088	0.1153	0.1182	0.1236	0.1236	0.1244	0.1297
Ser302	M+3	0.0155	0.0143	0.0143	0.0173	0.0210	0.0253	0.0275	0.0285	0.0285	0.0301	0.0287	0.0278
Ser302	M+4	0.0014	0.0014	0.0000	0.0027	0.0037	0.0063	0.0070	0.0082	0.0074	0.0082	0.0090	0.0084
Ser362	M+0	0.6540	0.6529	0.6517	0.6463	0.5943	0.5742	0.5521	0.5349	0.5230	0.5193	0.5100	0.4928
Ser362	M+1	0.2185	0.2177	0.2171	0.2213	0.2443	0.2525	0.2657	0.2749	0.2816	0.2855	0.2910	0.3012
Ser362	M+2	0.1006	0.1020	0.1044	0.1039	0.1212	0.1293	0.1332	0.1385	0.1419	0.1413	0.1435	0.1492
Ser362	M+3	0.0223	0.0228	0.0224	0.0231	0.0318	0.0351	0.0380	0.0398	0.0423	0.0419	0.0436	0.0446
Ser362	M+4	0.0046	0.0046	0.0044	0.0052	0.0082	0.0087	0.0098	0.0103	0.0112	0.0108	0.0109	0.0115
Ser390	M+0	0.6445	0.6421	0.6409	0.6358	0.5813	0.5568	0.5385	0.5173	0.5047	0.5014	0.4888	0.4730
Ser390	M+1	0.2215	0.2249	0.2247	0.2247	0.2434	0.2514	0.2547	0.2663	0.2709	0.2700	0.2786	0.2864
Ser390	M+2	0.1038	0.1039	0.1030	0.1055	0.1158	0.1230	0.1308	0.1352	0.1409	0.1409	0.1440	0.1466
Ser390	M+3	0.0246	0.0235	0.0257	0.0277	0.0447	0.0515	0.0561	0.0597	0.0611	0.0641	0.0657	0.0689
Ser390	M+4	0.0056	0.0056	0.0057	0.0058	0.0118	0.0136	0.0157	0.0170	0.0181	0.0184	0.0181	0.0199
Suc289	M+0	0.7369	0.7355	0.7258	0.7078	0.6149	0.5508	0.4715	0.3740	0.3558	0.3100	0.3294	0.2860
Suc289	M+1	0.1774	0.1761	0.1827	0.1850	0.2106	0.2275	0.2470	0.2606	0.2619	0.2700	0.2626	0.2688
Suc289	M+2	0.0748	0.0753	0.0787	0.0861	0.1243	0.1483	0.1767	0.2118	0.2216	0.2362	0.2281	0.2453
Suc289	M+3	0.0106	0.0121	0.0124	0.0171	0.0379	0.0524	0.0742	0.1044	0.1094	0.1215	0.1212	0.1314
Suc289	M+4	0.0004	0.0009	0.0004	0.0039	0.0116	0.0185	0.0259	0.0403	0.0431	0.0500	0.0480	0.0565
Suc289	M+5	0.0000	0.0000	0.0000	0.0000	0.0007	0.0025	0.0046	0.0086	0.0082	0.0110	0.0102	0.0115
Thr376	M+0	0.6556	0.6613	0.6678	0.6455	0.6269	0.6011	0.5851	0.5510	0.5418	0.5350	0.5346	0.5121
Thr376	M+0	0.2245	0.2229	0.2178	0.2286	0.2312	0.2343	0.2425	0.2506	0.2571	0.2520	0.2525	0.2565
Thr376	M+0	0.1025	0.1002	0.1004	0.1038	0.1092	0.1240	0.1245	0.1351	0.1437	0.1441	0.1429	0.1509
Thr376	M+0	0.0174	0.0155	0.0139	0.0210	0.0290	0.0369	0.0414	0.0500	0.0558	0.0580	0.0602	0.0663
Thr376	M+0	0.0000	0.0000	0.0000	0.0011	0.0037	0.0038	0.0065	0.0134	0.0017	0.0109	0.0097	0.0142
Thr404	M+0	0.6410	0.6385	0.6469	0.6401	0.6081	0.5867	0.5641	0.5375	0.5251	0.5117	0.5070	0.4941
Thr404	M+0	0.2253	0.2315	0.2331	0.2255	0.2359	0.2406	0.2442	0.2441	0.2465	0.2482	0.2499	0.2473
Thr404	M+0	0.1089	0.1049	0.0995	0.1086	0.1099	0.1176	0.1205	0.1328	0.1366	0.1386	0.1412	0.1430
Thr404	M+0	0.0239	0.0242	0.0206	0.0243	0.0374	0.0429	0.0519	0.0576	0.0652	0.0682	0.0693	0.0748
Thr404	M+0	0.0009	0.0008	0.0000	0.0014	0.0087	0.0122	0.0185	0.0220	0.0254	0.0278	0.0275	0.0306
Thr404	M+0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0008	0.0060	0.0013	0.0055	0.0051	0.0102
Val260	M+0	0.7371	0.7334	0.7321	0.7246	0.6659	0.6324	0.6033	0.5864	0.5777	0.5706	0.5668	0.5521
Val260	M+1	0.1816	0.1849	0.1835	0.1856	0.1922	0.2041	0.2035	0.2088	0.2083	0.2084	0.2103	0.2110
Val260	M+2	0.0716	0.0710	0.0739	0.0746	0.1028	0.1142	0.1282	0.1313	0.1381	0.1400	0.1394	0.1489
Val260	M+3	0.0096	0.0106	0.0105	0.0135	0.0295	0.0353	0.0456	0.0504	0.0536	0.0545	0.0571	0.0586
Val260	M+4	0.0000	0.0000	0.0000	0.0017	0.0095	0.0140	0.0169	0.0200	0.0197	0.0222	0.0229	0.0251
Val288	M+0	0.7297	0.7344	0.7303	0.7227	0.6629	0.6361	0.6082	0.5887	0.5702	0.5704	0.5619	0.5503
Val288	M+1	0.1917	0.1849	0.1852	0.1884	0.1964	0.1992	0.2028	0.2033	0.2107	0.2113	0.2120	0.2138
Val288	M+2	0.0693	0.0706	0.0730	0.0740	0.0911	0.0986	0.1068	0.1150	0.1191	0.1185	0.1201	0.1245
Val288	M+3	0.0093	0.0101	0.0115	0.0141	0.0325	0.0424	0.0514	0.0570	0.0619	0.0595	0.0641	0.0664
Val288	M+4	0.0000	0.0000	0.0000	0.0008	0.0123	0.0160	0.0218	0.0241	0.0258	0.0266	0.0283	0.0294
Val288	M+5	0.0000	0.0000	0.0000	0.0000	0.0048	0.0078	0.0089	0.0115	0.0123	0.0131	0.0136	0.0149

Table E.4: Mass isotopomer distributions for NMFA experiment L2 (part 2). Each column represents a different time point and each row represents a different metabolite mass isotopomer. The time points are listed in the first row (in hours).

Times	(h)	0.0025	0.0044	0.0169	0.0336	0.0669	0.1697	0.2603	0.4867	1.0153	1.4956	2.0231	3.0222
Frag	Mass	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
Akg346	M+0	0.7148	0.7141	0.7124	0.7025	0.6845	0.6035	0.5562	0.4644	0.3500	0.3101	0.2859	0.2615
Akg346	M+1	0.1900	0.1906	0.1907	0.1936	0.1970	0.2170	0.2281	0.2500	0.2709	0.2763	0.2762	0.2786
Akg346	M+2	0.0786	0.0785	0.0797	0.0843	0.0935	0.1274	0.1456	0.1778	0.2136	0.2256	0.2330	0.2394
Akg346	M+3	0.0136	0.0137	0.0141	0.0156	0.0194	0.0368	0.0478	0.0712	0.1062	0.1186	0.1286	0.1368
Akg346	M+4	0.0026	0.0026	0.0027	0.0034	0.0047	0.0117	0.0167	0.0268	0.0423	0.0490	0.0534	0.0581
Akg346	M+5	0.0003	0.0003	0.0003	0.0004	0.0008	0.0029	0.0045	0.0079	0.0137	0.0160	0.0180	0.0200
Ala232	M+0	0.7436	0.7428	0.6990	0.6458	0.6138	0.5748	0.5672	0.5488	0.5406	0.5240	0.5295	0.5133
Ala232	M+1	0.1696	0.1703	0.1811	0.1968	0.2098	0.2229	0.2259	0.2347	0.2393	0.2465	0.2432	0.2511
Ala232	M+2	0.0743	0.0745	0.0992	0.1273	0.1432	0.1612	0.1642	0.1716	0.1757	0.1803	0.1787	0.1835
Ala232	M+3	0.0113	0.0107	0.0169	0.0241	0.0262	0.0318	0.0332	0.0349	0.0349	0.0377	0.0383	0.0391
Ala232	M+4	0.0012	0.0017	0.0038	0.0061	0.0069	0.0092	0.0095	0.0100	0.0094	0.0109	0.0104	0.0118
Ala260	M+0	0.7385	0.7345	0.6903	0.6375	0.6021	0.5613	0.5477	0.5297	0.5219	0.5022	0.5047	0.4896
Ala260	M+1	0.1734	0.1756	0.1872	0.2001	0.2123	0.2270	0.2319	0.2390	0.2449	0.2474	0.2472	0.2515
Ala260	M+2	0.0696	0.0683	0.0710	0.0782	0.0826	0.0896	0.0917	0.0953	0.0985	0.1037	0.1036	0.1063
Ala260	M+3	0.0165	0.0184	0.0430	0.0678	0.0843	0.0974	0.1022	0.1087	0.1090	0.1164	0.1146	0.1204
Ala260	M+4	0.0021	0.0031	0.0070	0.0123	0.0142	0.0179	0.0196	0.0202	0.0197	0.0220	0.0219	0.0233
Ala260	M+5	0.0000	0.0000	0.0015	0.0041	0.0045	0.0069	0.0069	0.0070	0.0060	0.0083	0.0079	0.0084
Asp302	M+0	0.7198	0.7191	0.7174	0.7102	0.6957	0.6480	0.6237	0.5815	0.5229	0.5032	0.4895	0.4779
Asp302	M+1	0.1881	0.1887	0.1889	0.1913	0.1961	0.2122	0.2231	0.2431	0.2703	0.2796	0.2851	0.2891
Asp302	M+2	0.0756	0.0758	0.0771	0.0804	0.0871	0.1087	0.1183	0.1327	0.1537	0.1606	0.1661	0.1715
Asp302	M+3	0.0137	0.0137	0.0137	0.0147	0.0169	0.0242	0.0268	0.0324	0.0395	0.0421	0.0438	0.0454
Asp302	M+4	0.0025	0.0025	0.0026	0.0031	0.0037	0.0060	0.0068	0.0087	0.0112	0.0120	0.0126	0.0132
Asp376	M+0	0.6528	0.6529	0.6524	0.6466	0.6313	0.5913	0.5677	0.5302	0.4771	0.4550	0.4451	0.4362
Asp376	M+1	0.2162	0.2168	0.2154	0.2174	0.2222	0.2339	0.2430	0.2570	0.2793	0.2882	0.2922	0.2933
Asp376	M+2	0.1030	0.1012	0.1033	0.1058	0.1123	0.1296	0.1387	0.1533	0.1716	0.1781	0.1835	0.1865
Asp376	M+3	0.0230	0.0237	0.0236	0.0246	0.0273	0.0349	0.0379	0.0445	0.0538	0.0577	0.0573	0.0606
Asp376	M+4	0.0048	0.0053	0.0053	0.0057	0.0069	0.0092	0.0114	0.0129	0.0162	0.0176	0.0186	0.0197
Asp390	M+0	0.6443	0.6425	0.6419	0.6321	0.6076	0.5417	0.5094	0.4509	0.3843	0.3623	0.3466	0.3329
Asp390	M+1	0.2228	0.2231	0.2238	0.2255	0.2332	0.2560	0.2667	0.2869	0.3057	0.3119	0.3135	0.3144
Asp390	M+2	0.1031	0.1044	0.1035	0.1073	0.1162	0.1379	0.1484	0.1686	0.1917	0.1991	0.2051	0.2114
Asp390	M+3	0.0236	0.0238	0.0243	0.0273	0.0325	0.0470	0.0549	0.0665	0.0838	0.0889	0.0949	0.0990
Asp390	M+4	0.0056	0.0055	0.0057	0.0067	0.0087	0.0136	0.0162	0.0208	0.0266	0.0289	0.0299	0.0316
Asp390	M+5	0.0006	0.0008	0.0008	0.0012	0.0018	0.0032	0.0041	0.0055	0.0069	0.0077	0.0085	0.0090
Asp418	M+0	0.6332	0.6326	0.6294	0.6193	0.5969	0.5236	0.4885	0.4240	0.3504	0.3225	0.3054	0.2900
Asp418	M+1	0.2273	0.2280	0.2285	0.2301	0.2338	0.2489	0.2564	0.2716	0.2856	0.2897	0.2898	0.2934
Asp418	M+2	0.1068	0.1068	0.1073	0.1103	0.1170	0.1431	0.1563	0.1783	0.2012	0.2104	0.2164	0.2175
Asp418	M+3	0.0254	0.0255	0.0265	0.0303	0.0373	0.0574	0.0662	0.0827	0.1042	0.1131	0.1192	0.1244
Asp418	M+4	0.0061	0.0060	0.0065	0.0079	0.0112	0.0196	0.0236	0.0312	0.0415	0.0454	0.0488	0.0522
Asp418	M+5	0.0010	0.0010	0.0012	0.0016	0.0028	0.0056	0.0068	0.0091	0.0128	0.0140	0.0153	0.0164
Cit459	M+0	0.6115	0.6118	0.6071	0.6018	0.5931	0.5676	0.5489	0.5020	0.4103	0.3558	0.3118	0.2564
Cit459	M+1	0.2353	0.2350	0.2357	0.2360	0.2368	0.2389	0.2399	0.2427	0.2458	0.2463	0.2461	0.2446
Cit459	M+2	0.1146	0.1146	0.1170	0.1195	0.1231	0.1327	0.1394	0.1553	0.1833	0.1989	0.2107	0.2240
Cit459	M+3	0.0293	0.0293	0.0304	0.0319	0.0341	0.0416	0.0473	0.0619	0.0920	0.1105	0.1260	0.1456
Cit459	M+4	0.0076	0.0076	0.0080	0.0088	0.0099	0.0140	0.0173	0.0256	0.0438	0.0551	0.0653	0.0785
Cit459	M+5	0.0014	0.0014	0.0015	0.0018	0.0023	0.0039	0.0053	0.0089	0.0170	0.0227	0.0275	0.0342
Cit459	M+6	0.0003	0.0003	0.0003	0.0004	0.0005	0.0011	0.0015	0.0028	0.0058	0.0079	0.0096	0.0123
Glu330	M+0	0.7019	0.7018	0.6998	0.6907	0.6698	0.5911	0.5427	0.4552	0.3575	0.3263	0.3046	0.2865
Glu330	M+1	0.2026	0.2027	0.2028	0.2050	0.2107	0.2327	0.2466	0.2727	0.2971	0.3032	0.3058	0.3063
Glu330	M+2	0.0780	0.0779	0.0792	0.0838	0.0936	0.1274	0.1464	0.1782	0.2117	0.2220	0.2297	0.2368
Glu330	M+3	0.0145	0.0146	0.0150	0.0166	0.0203	0.0358	0.0463	0.0662	0.0922	0.1015	0.1084	0.1144
Glu330	M+4	0.0026	0.0026	0.0028	0.0033	0.0047	0.0104	0.0143	0.0216	0.0319	0.0361	0.0390	0.0425
Glu330	M+5	0.0004	0.0004	0.0004	0.0005	0.0008	0.0021	0.0031	0.0050	0.0078	0.0089	0.0099	0.0107
Glu432	M+0	0.6157	0.6151	0.6134	0.6046	0.5873	0.5139	0.4689	0.3882	0.2954	0.2622	0.2429	0.2239
Glu432	M+1	0.2332	0.2335	0.2337	0.2340	0.2367	0.2450	0.2503	0.2596	0.2665	0.2641	0.2646	0.2617
Glu432	M+2	0.1106	0.1108	0.1115	0.1157	0.1246	0.1536	0.1700	0.1971	0.2244	0.2304	0.2371	0.2395
Glu432	M+3	0.0272	0.0273	0.0277	0.0296	0.0344	0.0538	0.0666	0.0917	0.1238	0.1347	0.1437	0.1504
Glu432	M+4	0.0066	0.0066	0.0068	0.0076	0.0098	0.0191	0.0255	0.0380	0.0558	0.0623	0.0677	0.0723
Glu432	M+5	0.0045	0.0045	0.0047	0.0057	0.0050	0.0100	0.0128	0.0175	0.0238	0.0311	0.0302	0.0349
Glu432	M+6	0.0014	0.0014	0.0015	0.0018	0.0015	0.0031	0.0040	0.0054	0.0072	0.0100	0.0094	0.0114
Gly218	M+0	0.7738	0.7709	0.7726	0.7705	0.7723	0.7450	0.7433	0.7296	0.7337	0.7001	0.6951	0.6863
Gly218	M+1	0.1577	0.1575	0.1580	0.1614	0.1662	0.1823	0.1845	0.1937	0.2004	0.2173	0.2250	0.2258
Gly218	M+2	0.0624	0.0656	0.0668	0.0636	0.0615	0.0657	0.0646	0.0694	0.0659	0.0729	0.0688	0.0756
Gly218	M+3	0.0062	0.0061	0.0026	0.0045	0.0000	0.0069	0.0076	0.0073	0.0000	0.0097	0.0111	0.0123
Gly246	M+0	0.7618	0.7608	0.7608	0.7625	0.7697	0.7334	0.7230	0.7064	0.7046	0.6817	0.6657	0.6553
Gly246	M+1	0.1604	0.1637	0.1677	0.1624	0.1621	0.1720	0.1755	0.1832	0.1921	0.1965	0.2092	0.2108
Gly246	M+2	0.0702	0.0690	0.0664	0.0693	0.0682	0.0843	0.0888	0.0970	0.1033	0.1052	0.1085	0.1134
Gly246	M+3	0.0076	0.0064	0.0051	0.0058	0.0000	0.0103	0.0127	0.0134	0.0000	0.0166	0.0166	0.0198

Table E.5: Mass isotopomer distributions for NMFA experiment H1 (part 1). Each column represents a different time point and each row represents a different metabolite mass isotopomer. The time points are listed in the first row (in hours).

Times	(h)	0.0025	0.0044	0.0169	0.0336	0.0669	0.1697	0.2603	0.4867	1.0153	1.4956	2.0231	3.0222
Frag	Mass	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
Ile200	M+0	0.8089	0.8101	0.8028	0.7887	0.7858	0.7743	0.7637	0.7543	0.7430	0.7239	0.7285	0.7023
Ile200	M+1	0.1462	0.1456	0.1470	0.1515	0.1538	0.1578	0.1602	0.1645	0.1701	0.1734	0.1709	0.1795
Ile200	M+2	0.0415	0.0418	0.0454	0.0515	0.0558	0.0588	0.0650	0.0679	0.0741	0.0781	0.0777	0.0894
Ile200	M+3	0.0034	0.0025	0.0038	0.0052	0.0046	0.0090	0.0111	0.0133	0.0128	0.0191	0.0192	0.0234
Ile274	M+0	0.7305	0.7287	0.7293	0.7178	0.7150	0.7003	0.6943	0.6866	0.6765	0.6602	0.6624	0.6355
Ile274	M+1	0.1834	0.1866	0.1864	0.1906	0.1893	0.1967	0.1965	0.2031	0.2070	0.2053	0.2027	0.2116
Ile274	M+2	0.0669	0.0660	0.0714	0.0746	0.0850	0.0821	0.0888	0.0888	0.0882	0.1016	0.1047	0.1106
Ile274	M+3	0.0192	0.0187	0.0129	0.0170	0.0107	0.0209	0.0203	0.0215	0.0283	0.0280	0.0291	0.0327
Ile274	M+4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0048	0.0011	0.0095
Leu274	M+0	0.7375	0.7347	0.7391	0.7251	0.7188	0.6724	0.6534	0.6433	0.6403	0.6142	0.6270	0.5990
Leu274	M+1	0.1818	0.1851	0.1822	0.1899	0.1969	0.2112	0.2195	0.2155	0.2176	0.2216	0.2156	0.2261
Leu274	M+2	0.0703	0.0697	0.0696	0.0741	0.0732	0.0900	0.0947	0.0990	0.1039	0.1100	0.1073	0.1148
Leu274	M+3	0.0103	0.0104	0.0091	0.0108	0.0112	0.0223	0.0265	0.0331	0.0298	0.0379	0.0368	0.0426
Leu274	M+4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0041	0.0058	0.0091	0.0083	0.0134	0.0122	0.0143
Mal419	M+0	0.6326	0.6319	0.6249	0.6114	0.5938	0.5317	0.5010	0.4432	0.3574	0.3267	0.3029	0.2870
Mal419	M+1	0.2260	0.2268	0.2282	0.2302	0.2328	0.2465	0.2531	0.2653	0.2816	0.2871	0.2901	0.2917
Mal419	M+2	0.1083	0.1077	0.1088	0.1127	0.1189	0.1400	0.1514	0.1704	0.1982	0.2072	0.2140	0.2167
Mal419	M+3	0.0255	0.0259	0.0288	0.0336	0.0389	0.0554	0.0629	0.0792	0.1036	0.1129	0.1209	0.1272
Mal419	M+4	0.0062	0.0062	0.0073	0.0093	0.0116	0.0192	0.0229	0.0300	0.0419	0.0465	0.0506	0.0539
Mal419	M+5	0.0011	0.0011	0.0016	0.0022	0.0030	0.0055	0.0066	0.0089	0.0130	0.0145	0.0160	0.0173
Pyr174	M+0	0.8332	0.8318	0.7910	0.7302	0.7162	0.6432	0.6321	0.6008	0.5934	0.5558	0.5699	0.5473
Pyr174	M+1	0.1187	0.1161	0.1336	0.1612	0.1721	0.1880	0.2033	0.2125	0.2307	0.2297	0.2271	0.2301
Pyr174	M+2	0.0423	0.0402	0.0411	0.0498	0.0535	0.0628	0.0668	0.0766	0.0781	0.0855	0.0799	0.0852
Pyr174	M+3	0.0058	0.0118	0.0343	0.0573	0.0582	0.0967	0.0938	0.1015	0.0977	0.1164	0.1138	0.1222
Pyr174	M+4	0.0000	0.0000	0.0000	0.0015	0.0000	0.0077	0.0040	0.0073	0.0000	0.0099	0.0087	0.0112
Ser288	M+0	0.7231	0.7251	0.7227	0.7152	0.7076	0.6875	0.6677	0.6556	0.6299	0.6045	0.6016	0.5783
Ser288	M+1	0.1897	0.1881	0.1894	0.1933	0.1978	0.2091	0.2203	0.2284	0.2437	0.2581	0.2606	0.2759
Ser288	M+2	0.0726	0.0724	0.0732	0.0756	0.0780	0.0838	0.0898	0.0922	0.0998	0.1060	0.1062	0.1117
Ser288	M+3	0.0125	0.0123	0.0125	0.0133	0.0143	0.0162	0.0183	0.0196	0.0220	0.0251	0.0252	0.0272
Ser302	M+0	0.7139	0.7145	0.7130	0.7100	0.7051	0.6966	0.6846	0.6795	0.6570	0.6456	0.6460	0.6318
Ser302	M+1	0.1854	0.1853	0.1853	0.1851	0.1856	0.1876	0.1901	0.1924	0.1935	0.2015	0.2029	0.2068
Ser302	M+2	0.0831	0.0823	0.0841	0.0854	0.0901	0.0925	0.0987	0.1006	0.1128	0.1154	0.1146	0.1225
Ser302	M+3	0.0148	0.0150	0.0151	0.0162	0.0160	0.0190	0.0212	0.0211	0.0273	0.0279	0.0270	0.0287
Ser302	M+4	0.0029	0.0029	0.0025	0.0032	0.0032	0.0044	0.0051	0.0060	0.0087	0.0090	0.0082	0.0090
Ser362	M+0	0.6546	0.6534	0.6539	0.6471	0.6405	0.6215	0.6046	0.5914	0.5689	0.5494	0.5439	0.5206
Ser362	M+1	0.2171	0.2166	0.2168	0.2197	0.2222	0.2330	0.2408	0.2501	0.2618	0.2702	0.2742	0.2866
Ser362	M+2	0.1006	0.1017	0.1008	0.1036	0.1060	0.1102	0.1159	0.1181	0.1257	0.1316	0.1320	0.1381
Ser362	M+3	0.0221	0.0224	0.0226	0.0234	0.0249	0.0277	0.0301	0.0310	0.0340	0.0368	0.0383	0.0417
Ser362	M+4	0.0050	0.0051	0.0053	0.0053	0.0057	0.0065	0.0074	0.0079	0.0081	0.0100	0.0098	0.0106
Ser390	M+0	0.6444	0.6442	0.6420	0.6367	0.6316	0.6118	0.5913	0.5796	0.5544	0.5281	0.5264	0.4997
Ser390	M+1	0.2225	0.2226	0.2239	0.2259	0.2261	0.2333	0.2402	0.2454	0.2551	0.2638	0.2632	0.2732
Ser390	M+2	0.1035	0.1036	0.1032	0.1035	0.1053	0.1107	0.1168	0.1193	0.1268	0.1352	0.1359	0.1436
Ser390	M+3	0.0234	0.0237	0.0244	0.0263	0.0288	0.0336	0.0390	0.0416	0.0474	0.0533	0.0544	0.0601
Ser390	M+4	0.0056	0.0054	0.0058	0.0065	0.0068	0.0084	0.0103	0.0109	0.0131	0.0150	0.0154	0.0176
Suc289	M+0	0.7384	0.7383	0.7314	0.7116	0.6823	0.5997	0.5536	0.4736	0.3813	0.3588	0.3368	0.3151
Suc289	M+1	0.1745	0.1745	0.1760	0.1829	0.1909	0.2179	0.2343	0.2608	0.2893	0.2936	0.2963	0.2997
Suc289	M+2	0.0738	0.0737	0.0779	0.0866	0.1012	0.1335	0.1493	0.1774	0.2080	0.2132	0.2214	0.2302
Suc289	M+3	0.0114	0.0114	0.0126	0.0155	0.0205	0.0359	0.0453	0.0629	0.0859	0.0929	0.1004	0.1063
Suc289	M+4	0.0019	0.0020	0.0020	0.0034	0.0051	0.0112	0.0145	0.0205	0.0294	0.0325	0.0361	0.0381
Suc289	M+5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0018	0.0027	0.0043	0.0062	0.0075	0.0077	0.0088
Thr376	M+0	0.6710	0.6686	0.6734	0.6680	0.6752	0.6400	0.6228	0.6048	0.5660	0.5362	0.5330	0.5131
Thr376	M+1	0.2144	0.2219	0.2193	0.2280	0.2225	0.2309	0.2380	0.2423	0.2629	0.2571	0.2636	0.2708
Thr376	M+2	0.1022	0.0970	0.0976	0.0964	0.1023	0.1080	0.1097	0.1213	0.1334	0.1430	0.1444	0.1481
Thr376	M+3	0.0124	0.0125	0.0097	0.0075	0.0000	0.0211	0.0295	0.0317	0.0377	0.0513	0.0537	0.0553
Thr376	M+4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0124	0.0052	0.0127
Thr404	M+0	0.6110	0.6252	0.6146	0.6239	0.6109	0.6141	0.5994	0.5693	0.5363	0.5100	0.5109	0.4868
Thr404	M+1	0.2056	0.2134	0.2219	0.2278	0.2210	0.2323	0.2399	0.2369	0.2484	0.2542	0.2472	0.2576
Thr404	M+2	0.1505	0.1312	0.1414	0.1294	0.1449	0.1246	0.1223	0.1391	0.1537	0.1496	0.1572	0.1565
Thr404	M+3	0.0315	0.0302	0.0221	0.0190	0.0231	0.0269	0.0348	0.0444	0.0472	0.0618	0.0633	0.0690
Thr404	M+4	0.0013	0.0000	0.0000	0.0000	0.0000	0.0021	0.0036	0.0103	0.0144	0.0202	0.0187	0.0260
Val260	M+0	0.7383	0.7392	0.7283	0.7049	0.6843	0.6248	0.5937	0.5673	0.5546	0.5280	0.5338	0.5103
Val260	M+1	0.1804	0.1795	0.1846	0.1875	0.1920	0.2031	0.2092	0.2158	0.2198	0.2229	0.2219	0.2277
Val260	M+2	0.0704	0.0697	0.0752	0.0868	0.0978	0.1231	0.1371	0.1477	0.1532	0.1626	0.1590	0.1682
Val260	M+3	0.0107	0.0106	0.0116	0.0170	0.0208	0.0345	0.0421	0.0470	0.0510	0.0579	0.0573	0.0627
Val260	M+4	0.0002	0.0010	0.0003	0.0037	0.0050	0.0125	0.0157	0.0188	0.0197	0.0242	0.0237	0.0258
Val288	M+0	0.7354	0.7328	0.7283	0.7065	0.6841	0.6227	0.5945	0.5660	0.5440	0.5268	0.5303	0.5040
Val288	M+1	0.1818	0.1831	0.1863	0.1897	0.1944	0.2041	0.2106	0.2139	0.2213	0.2207	0.2193	0.2271
Val288	M+2	0.0717	0.0722	0.0723	0.0799	0.0863	0.1046	0.1099	0.1201	0.1243	0.1315	0.1319	0.1379
Val288	M+3	0.0108	0.0116	0.0131	0.0203	0.0293	0.0468	0.0562	0.0653	0.0719	0.0743	0.0749	0.0818
Val288	M+4	0.0004	0.0004	0.0000	0.0036	0.0054	0.0144	0.0190	0.0229	0.0267	0.0296	0.0280	0.0306
Val288	M+5	0.0000	0.0000	0.0000	0.0000	0.0006	0.0074	0.0098	0.0113	0.0119	0.0153	0.0141	0.0161

Table E.6: Mass isotopomer distributions for NMFA experiment H1 (part 2). Each column represents a different time point and each row represents a different metabolite mass isotopomer. The time points are listed in the first row (in hours).

Times	(h)	0.0008	0.0014	0.0025	0.0086	0.0169	0.1917	0.2911	0.5350	1.0097	1.5250	2.0161	3.0164	4.0019
Frag	Mass	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
Akg346	M+0	0.7174	0.7162	0.7174	0.7132	0.7035	0.5998	0.5370	0.4371	0.3492	0.3067	0.2867	0.2684	0.2576
Akg346	M+1	0.1885	0.1889	0.1884	0.1892	0.1926	0.2182	0.2315	0.2538	0.2679	0.2727	0.2740	0.2745	0.2756
Akg346	M+2	0.0779	0.0779	0.0782	0.0797	0.0837	0.1243	0.1473	0.1813	0.2114	0.2231	0.2296	0.2366	0.2381
Akg346	M+3	0.0136	0.0138	0.0133	0.0144	0.0159	0.0394	0.0557	0.0830	0.1104	0.1228	0.1297	0.1377	0.1403
Akg346	M+4	0.0025	0.0026	0.0023	0.0028	0.0034	0.0133	0.0205	0.0326	0.0440	0.0515	0.0554	0.0584	0.0605
Akg346	M+5	0.0002	0.0003	0.0000	0.0004	0.0005	0.0037	0.0062	0.0104	0.0144	0.0179	0.0192	0.0202	0.0218
Ala232	M+0	0.7471	0.7547	0.7548	0.7043	0.6522	0.5385	0.5255	0.5226	0.5063	0.4952	0.4924	0.5010	0.4877
Ala232	M+1	0.1678	0.1654	0.1684	0.1813	0.1986	0.2435	0.2487	0.2571	0.2580	0.2619	0.2645	0.2653	0.2648
Ala232	M+2	0.0728	0.0689	0.0676	0.0947	0.1206	0.1703	0.1769	0.1753	0.1877	0.1884	0.1904	0.1896	0.1919
Ala232	M+3	0.0106	0.0096	0.0092	0.0161	0.0226	0.0359	0.0372	0.0374	0.0387	0.0411	0.0403	0.0371	0.0422
Ala232	M+4	0.0016	0.0014	0.0000	0.0036	0.0058	0.0106	0.0109	0.0076	0.0094	0.0121	0.0116	0.0070	0.0120
Ala260	M+0	0.7377	0.7455	0.7470	0.6969	0.6422	0.5155	0.5060	0.4991	0.4873	0.4739	0.4707	0.4682	0.4656
Ala260	M+1	0.1737	0.1716	0.1734	0.1843	0.2042	0.2473	0.2515	0.2575	0.2674	0.2622	0.2650	0.2737	0.2666
Ala260	M+2	0.0704	0.0699	0.0683	0.0739	0.0782	0.0987	0.0992	0.1008	0.1053	0.1070	0.1075	0.1117	0.1083
Ala260	M+3	0.0159	0.0112	0.0110	0.0366	0.0600	0.1095	0.1137	0.1162	0.1154	0.1233	0.1246	0.1235	0.1257
Ala260	M+4	0.0024	0.0018	0.0003	0.0063	0.0112	0.0214	0.0218	0.0213	0.0204	0.0236	0.0233	0.0204	0.0247
Ala260	M+5	0.0000	0.0000	0.0000	0.0020	0.0041	0.0075	0.0076	0.0051	0.0043	0.0093	0.0088	0.0026	0.0085
Asp302	M+0	0.7205	0.7196	0.7170	0.7108	0.6988	0.6213	0.5980	0.5621	0.5264	0.5106	0.5033	0.4955	0.4919
Asp302	M+1	0.1881	0.1870	0.1903	0.1884	0.1923	0.2224	0.2316	0.2482	0.2636	0.2690	0.2726	0.2749	0.2755
Asp302	M+2	0.0750	0.0761	0.0760	0.0804	0.0853	0.1185	0.1293	0.1436	0.1564	0.1623	0.1659	0.1706	0.1709
Asp302	M+3	0.0137	0.0139	0.0139	0.0158	0.0183	0.0278	0.0309	0.0351	0.0410	0.0422	0.0424	0.0449	0.0453
Asp302	M+4	0.0025	0.0029	0.0028	0.0036	0.0041	0.0079	0.0083	0.0097	0.0106	0.0128	0.0128	0.0126	0.0134
Asp376	M+0	0.6530	0.6470	0.6513	0.6486	0.6349	0.5601	0.5430	0.5138	0.4827	0.4637	0.4569	0.4435	0.4477
Asp376	M+1	0.2171	0.2197	0.2226	0.2180	0.2229	0.2445	0.2525	0.2663	0.2726	0.2810	0.2795	0.2905	0.2814
Asp376	M+2	0.1010	0.1034	0.0991	0.1023	0.1104	0.1431	0.1496	0.1622	0.1738	0.1795	0.1839	0.1933	0.1881
Asp376	M+3	0.0242	0.0248	0.0247	0.0262	0.0259	0.0408	0.0435	0.0476	0.0549	0.0572	0.0603	0.0592	0.0611
Asp376	M+4	0.0048	0.0051	0.0024	0.0049	0.0058	0.0114	0.0110	0.0101	0.0160	0.0170	0.0189	0.0134	0.0199
Asp390	M+0	0.6436	0.6458	0.6460	0.6380	0.6159	0.5066	0.4740	0.4298	0.3853	0.3636	0.3525	0.3426	0.3387
Asp390	M+1	0.2242	0.2216	0.2225	0.2239	0.2288	0.2618	0.2758	0.2880	0.3022	0.3032	0.3059	0.3145	0.3068
Asp390	M+2	0.1022	0.1029	0.1032	0.1069	0.1139	0.1502	0.1589	0.1756	0.1906	0.1993	0.2020	0.2058	0.2092
Asp390	M+3	0.0239	0.0242	0.0237	0.0252	0.0315	0.0596	0.0656	0.0775	0.0868	0.0933	0.0977	0.0988	0.1018
Asp390	M+4	0.0053	0.0053	0.0046	0.0061	0.0084	0.0173	0.0196	0.0227	0.0283	0.0305	0.0316	0.0304	0.0329
Asp390	M+5	0.0008	0.0003	0.0000	0.0000	0.0015	0.0046	0.0057	0.0063	0.0068	0.0085	0.0092	0.0078	0.0092
Asp418	M+0	0.6345	0.6343	0.6371	0.6256	0.6058	0.4846	0.4518	0.4028	0.3560	0.3285	0.3158	0.3052	0.2991
Asp418	M+1	0.2264	0.2261	0.2267	0.2286	0.2324	0.2600	0.2660	0.2749	0.2861	0.2888	0.2911	0.2939	0.2924
Asp418	M+2	0.1064	0.1076	0.1046	0.1078	0.1112	0.1430	0.1537	0.1727	0.1868	0.1950	0.1965	0.1981	0.2004
Asp418	M+3	0.0255	0.0249	0.0251	0.0291	0.0373	0.0743	0.0841	0.0969	0.1088	0.1176	0.1232	0.1280	0.1297
Asp418	M+4	0.0060	0.0062	0.0059	0.0072	0.0103	0.0266	0.0314	0.0373	0.0441	0.0487	0.0508	0.0538	0.0545
Asp418	M+5	0.0010	0.0009	0.0006	0.0013	0.0026	0.0087	0.0101	0.0117	0.0146	0.0157	0.0167	0.0171	0.0179
Cit459	M+0	0.6118	0.6113	0.6103	0.6088	0.6065	0.5852	0.5691	0.5364	0.4790	0.4348	0.3917	0.3410	0.3068
Cit459	M+1	0.2350	0.2348	0.2352	0.2353	0.2355	0.2375	0.2388	0.2406	0.2426	0.2427	0.2434	0.2434	0.2427
Cit459	M+2	0.1149	0.1150	0.1154	0.1161	0.1171	0.1242	0.1290	0.1390	0.1556	0.1674	0.1796	0.1927	0.2016
Cit459	M+3	0.0291	0.0294	0.0296	0.0299	0.0306	0.0370	0.0421	0.0527	0.0714	0.0866	0.1011	0.1186	0.1310
Cit459	M+4	0.0076	0.0076	0.0077	0.0078	0.0082	0.0117	0.0146	0.0206	0.0323	0.0418	0.0510	0.0623	0.0704
Cit459	M+5	0.0014	0.0014	0.0015	0.0015	0.0017	0.0033	0.0046	0.0075	0.0130	0.0177	0.0224	0.0280	0.0323
Cit459	M+6	0.0003	0.0003	0.0003	0.0003	0.0004	0.0009	0.0014	0.0024	0.0044	0.0063	0.0080	0.0100	0.0117
Glu330	M+0	0.7025	0.7016	0.7031	0.6985	0.6912	0.5901	0.5302	0.4423	0.3623	0.3232	0.3053	0.2903	0.2808
Glu330	M+1	0.2021	0.2025	0.2017	0.2035	0.2055	0.2322	0.2475	0.2699	0.2880	0.2936	0.2961	0.2980	0.2980
Glu330	M+2	0.0777	0.0782	0.0777	0.0794	0.0829	0.1258	0.1497	0.1830	0.2124	0.2264	0.2332	0.2393	0.2422
Glu330	M+3	0.0146	0.0147	0.0145	0.0152	0.0165	0.0371	0.0506	0.0717	0.0925	0.1046	0.1100	0.1142	0.1182
Glu330	M+4	0.0026	0.0026	0.0026	0.0029	0.0033	0.0117	0.0170	0.0255	0.0342	0.0395	0.0421	0.0440	0.0460
Glu330	M+5	0.0004	0.0004	0.0004	0.0005	0.0006	0.0025	0.0039	0.0060	0.0083	0.0100	0.0106	0.0112	0.0116
Glu432	M+0	0.6160	0.6121	0.6153	0.6045	0.5926	0.5036	0.4528	0.3726	0.2982	0.2594	0.2418	0.2290	0.2176
Glu432	M+1	0.2320	0.2308	0.2301	0.2287	0.2273	0.2403	0.2488	0.2577	0.2620	0.2596	0.2591	0.2608	0.2566
Glu432	M+2	0.1088	0.1086	0.1076	0.1087	0.1108	0.1460	0.1669	0.1936	0.2166	0.2262	0.2311	0.2356	0.2362
Glu432	M+3	0.0266	0.0266	0.0262	0.0268	0.0283	0.0539	0.0713	0.0965	0.1223	0.1366	0.1428	0.1484	0.1517
Glu432	M+4	0.0064	0.0064	0.0063	0.0067	0.0073	0.0197	0.0284	0.0417	0.0555	0.0643	0.0681	0.0712	0.0738
Glu432	M+5	0.0067	0.0100	0.0095	0.0157	0.0215	0.0238	0.0214	0.0259	0.0313	0.0369	0.0389	0.0379	0.0433
Glu432	M+6	0.0022	0.0034	0.0032	0.0055	0.0075	0.0081	0.0070	0.0082	0.0098	0.0117	0.0124	0.0119	0.0142
Gly218	M+0	0.7781	0.7749	0.8170	0.7791	0.7721	0.7596	0.7576	0.8178	0.7862	0.7218	0.7437	0.8981	0.7238
Gly218	M+1	0.1613	0.1561	0.1514	0.1594	0.1618	0.1753	0.1818	0.1822	0.2138	0.2011	0.1935	0.1019	0.2119
Gly218	M+2	0.0592	0.0657	0.0316	0.0615	0.0660	0.0636	0.0606	0.0000	0.0000	0.0688	0.0628	0.0000	0.0610
Gly218	M+3	0.0014	0.0033	0.0000	0.0000	0.0000	0.0016	0.0000	0.0000	0.0000	0.0083	0.0000	0.0000	0.0033
Gly246	M+0	0.7668	0.7567	0.7811	0.7659	0.7636	0.7586	0.7385	0.8155	0.8043	0.7046	0.7180	0.9291	0.6956
Gly246	M+1	0.1676	0.1714	0.1671	0.1665	0.1735	0.1655	0.1799	0.1744	0.1693	0.1909	0.1923	0.0709	0.1992
Gly246	M+2	0.0636	0.0705	0.0518	0.0617	0.0629	0.0759	0.0775	0.0101	0.0264	0.0923	0.0866	0.0000	0.1015
Gly246	M+3	0.0019	0.0014	0.0000	0.0059	0.0000	0.0000	0.0041	0.0000	0.0000	0.0122	0.0031	0.0000	0.0037

Table E.7: Mass isotopomer distributions for NMFA experiment H2 (part 1). Each column represents a different time point and each row represents a different metabolite mass isotopomer. The time points are listed in the first row (in hours).

Times	(h)	0.0008	0.0014	0.0025	0.0086	0.0169	0.1917	0.2911	0.5350	1.0097	1.5250	2.0161	3.0164	4.0019
Frag	Mass	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
Ile200	M+0	0.8008	0.8079	0.8163	0.8011	0.7984	0.7738	0.7657	0.7595	0.7626	0.7392	0.7392	0.7535	0.7270
Ile200	M+1	0.1522	0.1447	0.1460	0.1463	0.1485	0.1590	0.1612	0.1772	0.1682	0.1677	0.1672	0.1792	0.1728
Ile200	M+2	0.0440	0.0435	0.0377	0.0463	0.0471	0.0593	0.0625	0.0612	0.0668	0.0709	0.0751	0.0673	0.0766
Ile200	M+3	0.0030	0.0039	0.0000	0.0062	0.0061	0.0078	0.0106	0.0020	0.0024	0.0160	0.0166	0.0000	0.0189
Ile274	M+0	0.7098	0.7158	0.7388	0.7130	0.7086	0.6922	0.6915	0.7155	0.7164	0.6536	0.6604	0.7363	0.6556
Ile274	M+1	0.1788	0.1836	0.2001	0.1884	0.1921	0.1941	0.1913	0.2090	0.2204	0.2033	0.2083	0.2056	0.2096
Ile274	M+2	0.0670	0.0706	0.0612	0.0714	0.0797	0.0867	0.0902	0.0703	0.0632	0.0921	0.1011	0.0580	0.0998
Ile274	M+3	0.0444	0.0290	0.0000	0.0266	0.0196	0.0270	0.0270	0.0053	0.0000	0.0434	0.0302	0.0000	0.0301
Ile274	M+4	0.0000	0.0010	0.0000	0.0006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0076	0.0000	0.0000	0.0049
Leu274	M+0	0.7285	0.7313	0.7447	0.7311	0.7268	0.6960	0.6792	0.6747	0.6704	0.6440	0.6438	0.6588	0.6306
Leu274	M+1	0.1888	0.1853	0.1849	0.1852	0.1870	0.2040	0.2123	0.2201	0.2201	0.2151	0.2160	0.2276	0.2204
Leu274	M+2	0.0727	0.0733	0.0655	0.0711	0.0728	0.0794	0.0858	0.0889	0.0862	0.0978	0.1015	0.1011	0.1003
Leu274	M+3	0.0100	0.0097	0.0048	0.0114	0.0119	0.0177	0.0203	0.0146	0.0234	0.0325	0.0309	0.0125	0.0356
Leu274	M+4	0.0000	0.0004	0.0000	0.0011	0.0015	0.0030	0.0025	0.0017	0.0000	0.0096	0.0079	0.0000	0.0110
Mal419	M+0	0.6338	0.6341	0.6346	0.6155	0.5925	0.4941	0.4579	0.4051	0.3548	0.3262	0.3127	0.3033	0.2967
Mal419	M+1	0.2252	0.2254	0.2256	0.2299	0.2356	0.2570	0.2646	0.2781	0.2854	0.2882	0.2906	0.2906	0.2916
Mal419	M+2	0.1074	0.1076	0.1071	0.1107	0.1143	0.1411	0.1531	0.1689	0.1852	0.1953	0.1979	0.2013	0.2023
Mal419	M+3	0.0258	0.0256	0.0253	0.0325	0.0412	0.0716	0.0811	0.0949	0.1104	0.1190	0.1239	0.1277	0.1297
Mal419	M+4	0.0064	0.0061	0.0062	0.0086	0.0120	0.0254	0.0302	0.0366	0.0446	0.0494	0.0519	0.0539	0.0550
Mal419	M+5	0.0010	0.0010	0.0009	0.0021	0.0034	0.0082	0.0096	0.0122	0.0146	0.0162	0.0169	0.0175	0.0183
Pyr174	M+0	0.8190	0.8401	0.8460	0.7791	0.7164	0.5886	0.5754	0.5536	0.5627	0.5422	0.5366	0.5328	0.5246
Pyr174	M+1	0.1201	0.1129	0.1158	0.1366	0.1629	0.2223	0.2288	0.2437	0.2515	0.2385	0.2391	0.2697	0.2452
Pyr174	M+2	0.0467	0.0420	0.0382	0.0442	0.0501	0.0730	0.0759	0.0821	0.0713	0.0801	0.0884	0.0822	0.0891
Pyr174	M+3	0.0131	0.0050	0.0000	0.0393	0.0638	0.1081	0.1110	0.1206	0.1145	0.1244	0.1267	0.1153	0.1256
Pyr174	M+4	0.0000	0.0000	0.0000	0.0008	0.0055	0.0074	0.0076	0.0000	0.0000	0.0113	0.0081	0.0000	0.0127
Ser288	M+0	0.7269	0.7235	0.7263	0.7205	0.7187	0.6831	0.6745	0.6533	0.6432	0.6180	0.6126	0.6056	0.5870
Ser288	M+1	0.1878	0.1904	0.1891	0.1914	0.1918	0.2101	0.2164	0.2293	0.2343	0.2501	0.2546	0.2643	0.2700
Ser288	M+2	0.0717	0.0719	0.0717	0.0729	0.0739	0.0850	0.0879	0.0958	0.0992	0.1022	0.1043	0.1051	0.1109
Ser288	M+3	0.0119	0.0121	0.0117	0.0123	0.0129	0.0176	0.0176	0.0195	0.0207	0.0231	0.0235	0.0233	0.0257
Ser302	M+0	0.7108	0.7174	0.7182	0.7080	0.7053	0.6925	0.6835	0.6751	0.6657	0.6491	0.6489	0.6497	0.6392
Ser302	M+1	0.1881	0.1846	0.1843	0.1883	0.1858	0.1870	0.1890	0.1925	0.1982	0.1998	0.2006	0.2115	0.2014
Ser302	M+2	0.0836	0.0798	0.0834	0.0832	0.0871	0.0948	0.1003	0.1084	0.1099	0.1116	0.1142	0.1142	0.1195
Ser302	M+3	0.0151	0.0152	0.0136	0.0173	0.0183	0.0208	0.0224	0.0230	0.0254	0.0294	0.0273	0.0246	0.0317
Ser302	M+4	0.0025	0.0030	0.0004	0.0032	0.0035	0.0047	0.0048	0.0010	0.0008	0.0087	0.0084	0.0000	0.0078
Ser362	M+0	0.6537	0.6545	0.6559	0.6516	0.6480	0.6218	0.6107	0.5888	0.5777	0.5597	0.5562	0.5467	0.5312
Ser362	M+1	0.2175	0.2169	0.2145	0.2177	0.2186	0.2320	0.2372	0.2543	0.2565	0.2673	0.2671	0.2764	0.2826
Ser362	M+2	0.1005	0.1003	0.1019	0.1021	0.1033	0.1115	0.1151	0.1199	0.1224	0.1265	0.1281	0.1318	0.1344
Ser362	M+3	0.0225	0.0224	0.0228	0.0225	0.0237	0.0275	0.0282	0.0303	0.0351	0.0358	0.0362	0.0374	0.0393
Ser362	M+4	0.0056	0.0050	0.0049	0.0053	0.0054	0.0064	0.0073	0.0067	0.0084	0.0088	0.0100	0.0077	0.0102
Ser390	M+0	0.6471	0.6439	0.6483	0.6439	0.6411	0.6125	0.5991	0.5790	0.5626	0.5423	0.5372	0.5336	0.5134
Ser390	M+1	0.2195	0.2230	0.2220	0.2225	0.2238	0.2320	0.2378	0.2531	0.2523	0.2588	0.2606	0.2659	0.2704
Ser390	M+2	0.1050	0.1038	0.1018	0.1032	0.1041	0.1112	0.1143	0.1181	0.1251	0.1301	0.1313	0.1316	0.1393
Ser390	M+3	0.0235	0.0233	0.0231	0.0244	0.0247	0.0335	0.0372	0.0392	0.0474	0.0512	0.0534	0.0559	0.0568
Ser390	M+4	0.0049	0.0054	0.0049	0.0055	0.0057	0.0089	0.0096	0.0106	0.0103	0.0141	0.0141	0.0122	0.0156
Suc289	M+0	0.7395	0.7376	0.7377	0.7301	0.7156	0.6105	0.5486	0.4688	0.4056	0.3576	0.3449	0.3368	0.3215
Suc289	M+1	0.1753	0.1746	0.1752	0.1763	0.1814	0.2156	0.2352	0.2629	0.2796	0.2881	0.2897	0.2991	0.2924
Suc289	M+2	0.0729	0.0743	0.0749	0.0779	0.0839	0.1245	0.1509	0.1842	0.2068	0.2160	0.2205	0.2317	0.2284
Suc289	M+3	0.0109	0.0115	0.0115	0.0128	0.0151	0.0358	0.0474	0.0642	0.0801	0.0947	0.0991	0.1026	0.1077
Suc289	M+4	0.0015	0.0020	0.0007	0.0025	0.0036	0.0115	0.0152	0.0188	0.0266	0.0342	0.0367	0.0299	0.0397
Suc289	M+5	0.0000	0.0000	0.0000	0.0000	0.0004	0.0020	0.0027	0.0011	0.0013	0.0077	0.0080	0.0000	0.0090
Thr376	M+0	0.6922	0.6689	0.6984	0.6519	0.6641	0.6308	0.6196	0.6611	0.6392	0.5589	0.5620	0.6657	0.5387
Thr376	M+1	0.2151	0.2291	0.2278	0.2253	0.2206	0.2336	0.2433	0.2296	0.2747	0.2548	0.2528	0.2539	0.2596
Thr376	M+2	0.0927	0.0959	0.0739	0.1042	0.1002	0.1085	0.1130	0.1092	0.0861	0.1359	0.1409	0.0804	0.1422
Thr376	M+3	0.0000	0.0061	0.0000	0.0186	0.0152	0.0260	0.0241	0.0000	0.0000	0.0435	0.0443	0.0000	0.0467
Thr376	M+4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0011	0.0000	0.0000	0.0000	0.0068	0.0000	0.0000	0.0127
Thr404	M+0	0.2388	0.6100	0.6500	0.5985	0.6072	0.5733	0.5526	0.6200	0.6136	0.4766	0.5017	0.5676	0.4830
Thr404	M+1	0.0869	0.2127	0.2373	0.2077	0.2158	0.2184	0.2272	0.2384	0.2473	0.2278	0.2320	0.2546	0.2365
Thr404	M+2	0.4465	0.1366	0.1128	0.1472	0.1333	0.1503	0.1634	0.1143	0.1228	0.1939	0.1780	0.1240	0.1825
Thr404	M+3	0.1487	0.0341	0.0000	0.0411	0.0359	0.0463	0.0510	0.0273	0.0163	0.0770	0.0707	0.0537	0.0731
Thr404	M+4	0.0665	0.0065	0.0000	0.0054	0.0078	0.0117	0.0058	0.0000	0.0000	0.0247	0.0176	0.0000	0.0229
Thr404	M+5	0.0126	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0020
Val260	M+0	0.7341	0.7353	0.7279	0.7277	0.7213	0.6519	0.6223	0.5788	0.5712	0.5469	0.5424	0.5343	0.5325
Val260	M+1	0.1816	0.1808	0.1857	0.1820	0.1858	0.2006	0.2053	0.2214	0.2218	0.2226	0.2249	0.2314	0.2231
Val260	M+2	0.0730	0.0711	0.0763	0.0761	0.0766	0.1089	0.1214	0.1345	0.1307	0.1513	0.1522	0.1491	0.1582
Val260	M+3	0.0114	0.0115	0.0100	0.0125	0.0138	0.0293	0.0360	0.0491	0.0618	0.0538	0.0556	0.0663	0.0595
Val260	M+4	0.0000	0.0013	0.0000	0.0017	0.0025	0.0086	0.0140	0.0161	0.0144	0.0206	0.0217	0.0189	0.0227
Val288	M+0	0.7313	0.7317	0.7430	0.7263	0.7191	0.6492	0.6200	0.5945	0.5797	0.5461	0.5423	0.5444	0.5290
Val288	M+1	0.1849	0.1852	0.1803	0.1838	0.1875	0.2014	0.2095	0.2231	0.2272	0.2207	0.2223	0.2349	0.2245
Val288	M+2	0.0735	0.0714	0.0704	0.0758	0.0754	0.0974	0.1036	0.1156	0.1199	0.1246	0.1292	0.1412	0.1303
Val288	M+3	0.0103	0.0112	0.0064	0.0129	0.0151	0.0361	0.0459	0.0509	0.0584	0.0683	0.0668	0.0643	0.0725
Val288	M+4	0.0000	0.0005	0.0000	0.0012	0.0027	0.0113	0.0144	0.0145	0.0114	0.0256	0.0270	0.0151	0.0292
Val288	M+5	0.0000	0.0000	0.0000	0.0000	0.0002	0.0046	0.0065	0.0014	0.0034	0.0126	0.0119	0.0000	0.0133

Table E.8: Mass isotopomer distributions for NMFA experiment H2 (part 2). Each column represents a different time point and each row represents a different metabolite mass isotopomer. The time points are listed in the first row (in hours).

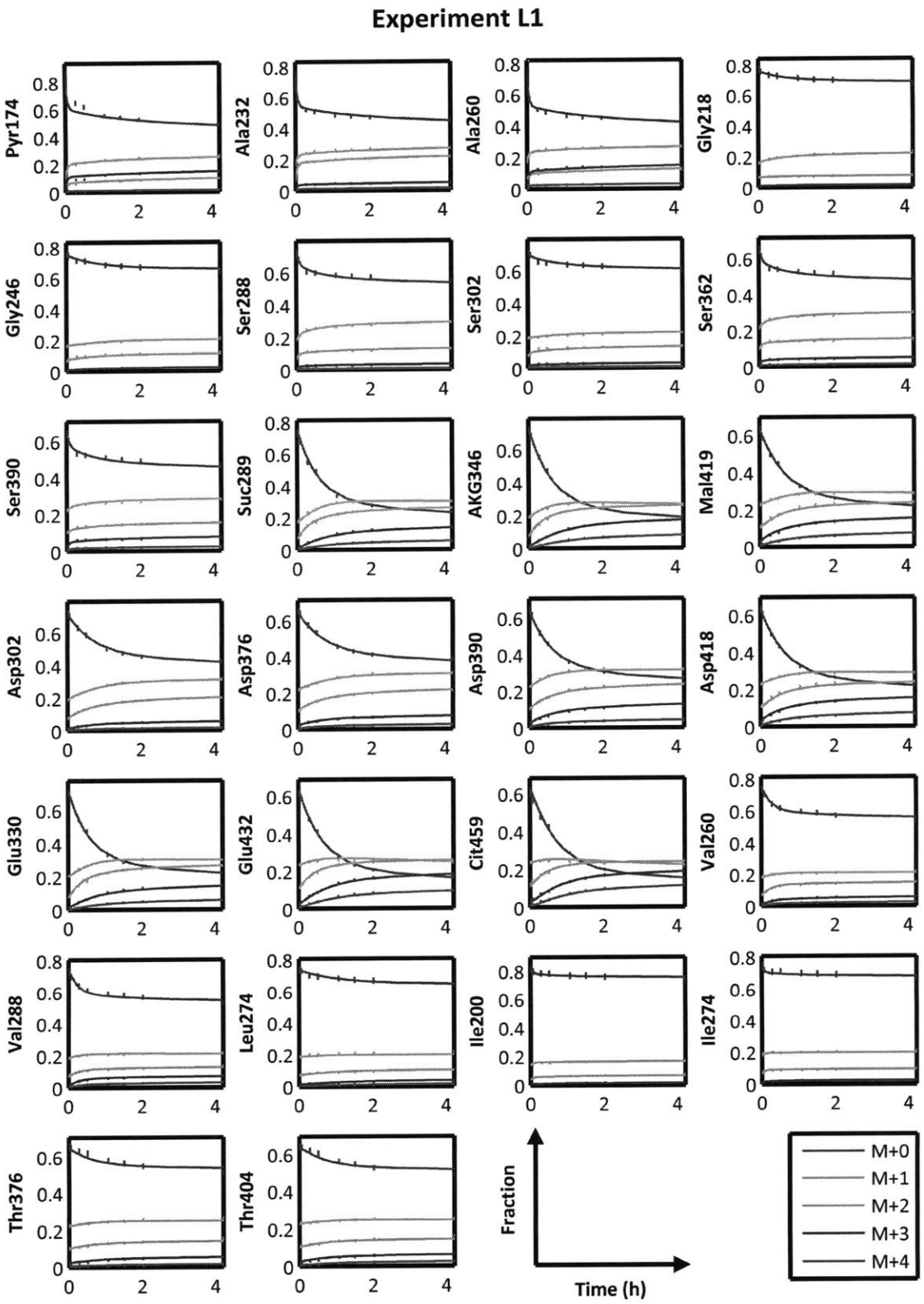

Figure E-1: Simulated isotopic labeling for experiment L1 (low aeration, linear phase) fitted to GC/MS measurements of organic and amino acid fragments.

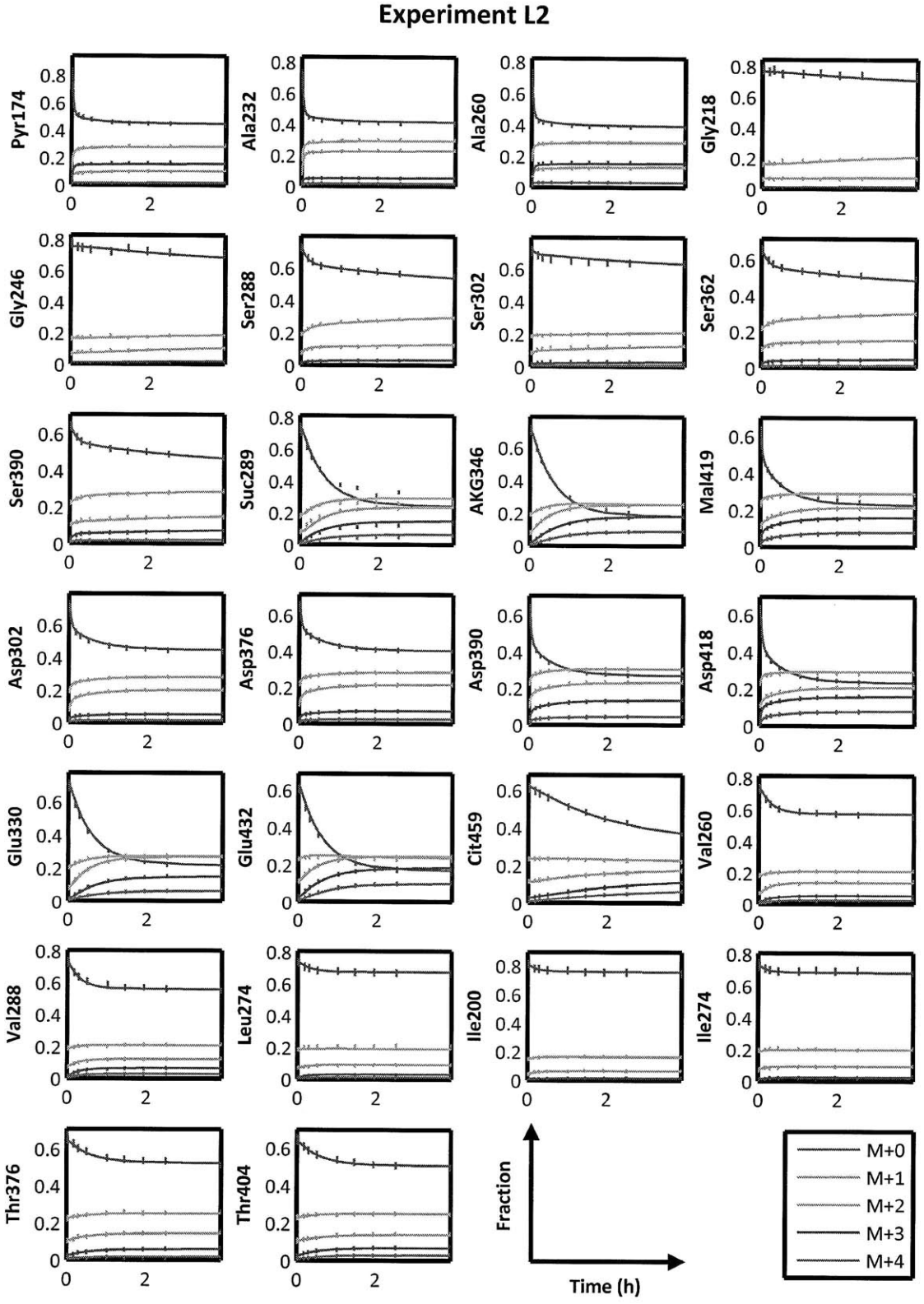

Figure E-2: Simulated isotopic labeling for experiment L2 (low aeration, stationary phase) fitted to GC/MS measurements of organic and amino acid fragments.

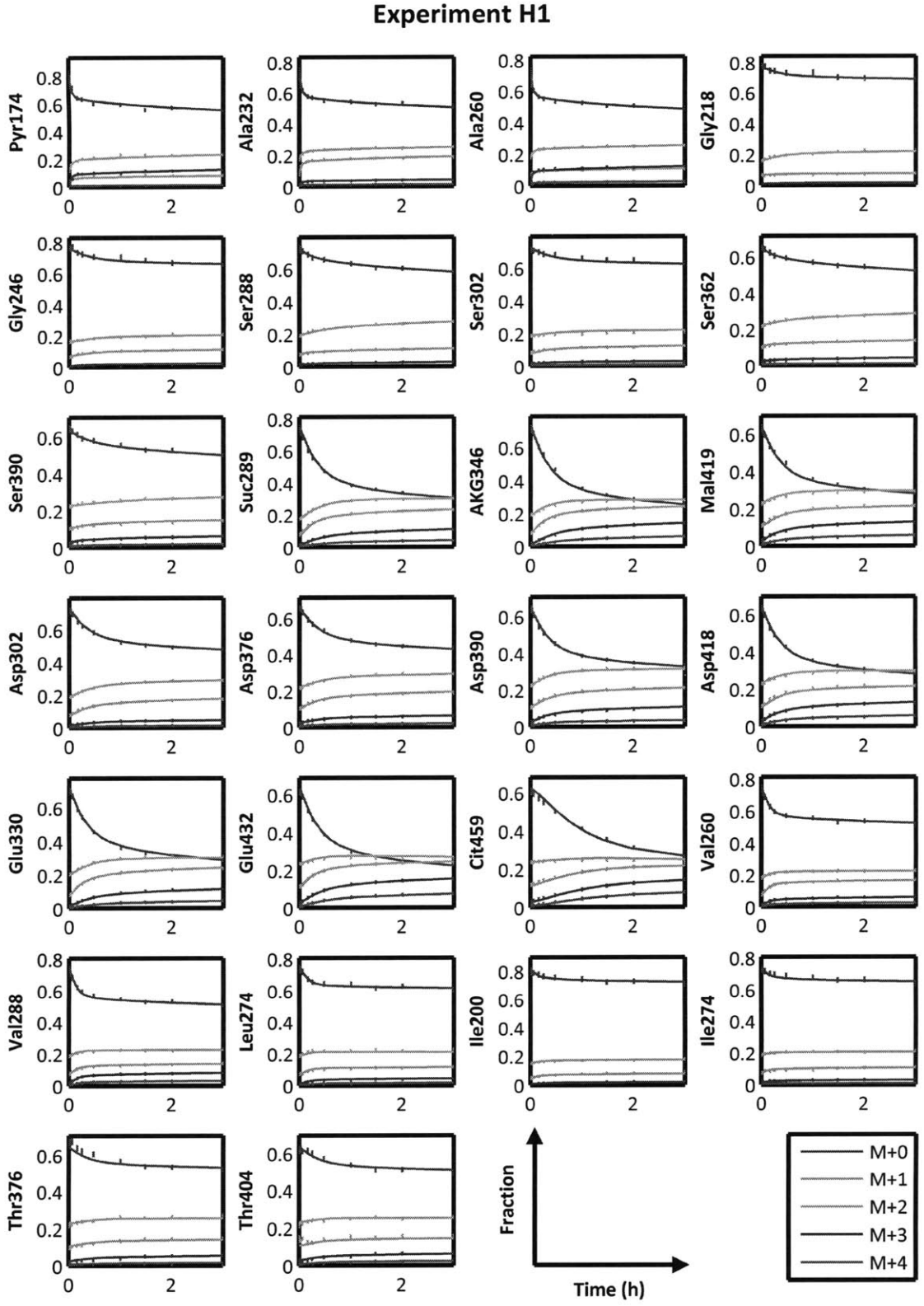

Figure E-3: Simulated isotopic labeling for experiment H1 (high aeration, linear phase) fitted to GC/MS measurements of organic and amino acid fragments.

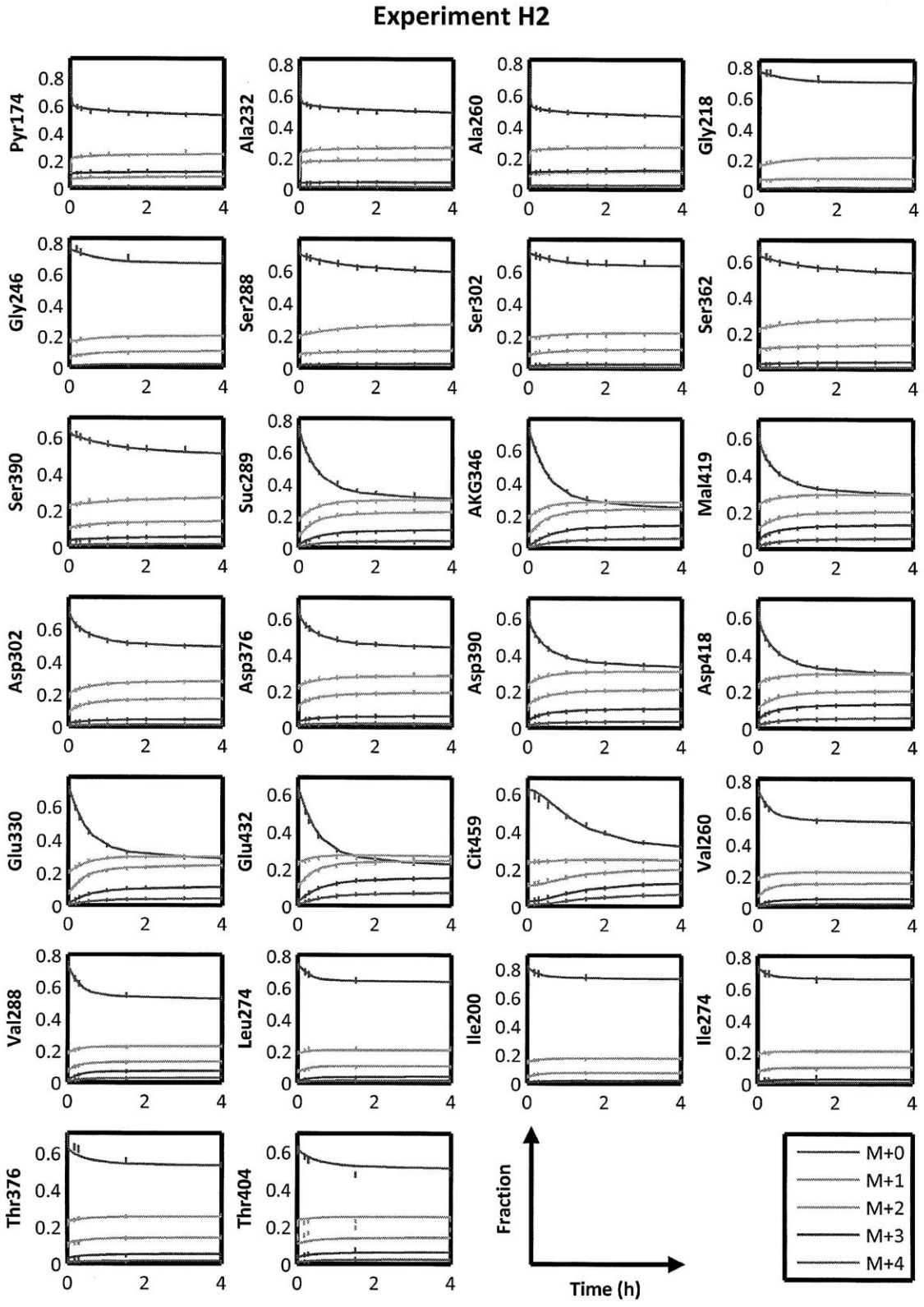

Figure E-4: Simulated isotopic labeling for experiment H2 (high aeration, stationary phase) fitted to GC/MS measurements of organic and amino acid fragments.

Flux	L1	L2	H1	H2
	97.43	98.38	97.32	97.77
$Gluc_{ext}$ → G6P	**100.00**	**100.00**	**100.00**	**100.00**
	101.80	102.61	102.94	102.07
	192.29	155.52	203.41	190.93
CO2 → $CO2_{ext}$	**200.68**	**159.84**	**208.27**	**195.70**
	205.93	160.27	210.42	200.69
	0.00	47.78	0.48	51.58
Cit → Cit_{ext}	**0.00**	**49.74**	**1.24**	**53.66**
	1.32	51.71	1.63	55.92
	44.27	7.10	21.09	0.00
Glyc → $Glyc_{ext}$	**49.45**	**11.48**	**27.75**	**4.10**
	50.92	13.18	37.47	8.72
	0.78	0.40	1.27	0.00
24 AcCoA → C16	**0.85**	**0.62**	**1.53**	**0.15**
	1.11	0.83	1.70	0.43
	1.29	0.54	2.57	0.44
27 AcCoA → C18	**1.33**	**0.77**	**2.85**	**0.61**
	1.57	0.98	3.06	0.89
	159.87	33.06	67.07	29.77
Metabolites → Biomass	**162.47**	**38.31**	**73.04**	**35.06**
	169.98	43.30	75.95	39.63
	54.51	80.86	53.27	92.36
G6P → F6P	**57.42**	**83.40**	**56.14**	**97.85**
	65.86	85.36	57.72	100.60
	596.68	0.00	230.44	0.00
G6P ↔ F6P	**9878.76**	**0.00**	**666.93**	**0.00**
	Inf	60.41	76746.00	27285.19
	77.18	91.00	80.38	94.96
F6P → DHAP + GAP	**79.64**	**93.01**	**82.60**	**97.95**
	84.48	95.14	85.62	99.99
	24.07	76.91	47.59	90.38
DHAP → GAP	**28.00**	**80.15**	**50.47**	**93.09**
	34.74	83.91	55.57	97.40
	3200.39	0.00	624.18	0.00
DHAP ↔ GAP	**40027.14**	**0.00**	**40668.99**	**0.00**
	Inf	4.68	Inf	78.84
	46.61	8.46	25.57	0.69
DHAP → Glyc	**51.64**	**12.86**	**32.13**	**4.86**
	53.48	16.89	39.40	8.49
	116.72	173.64	143.67	186.28
GAP → 3PG	**117.62**	**177.70**	**145.80**	**190.85**
	123.19	182.24	150.38	196.38
	0.00	0.00	0.00	0.00
GAP ↔ 3PG	**120.61**	**2144.36**	**259.87**	**0.53**
	Inf	Inf	Inf	Inf
	109.94	172.90	141.51	186.98
3PG → PEP	**113.22**	**176.52**	**143.83**	**190.13**
	119.26	181.68	148.46	195.68
	0.00	320.42	0.00	0.00
3PG ↔ PEP	**0.00**	**62614.56**	**112.10**	**0.00**
	Inf	Inf	Inf	Inf
	87.18	117.62	122.60	132.16
PEP → Pyr	**92.67**	**121.86**	**129.31**	**136.34**
	115.20	132.11	140.05	162.72

Table E.9: NMFA estimation results for *Y. lipolytica* at low aeration & linear growth, low aeration & stationary growth, high aeration & linear growth, and high aeration & stationary growth (L1, L2, H1, H2). Values are normalized to a glucose input flux of 100. The lower bound, the estimated value (bold), and upper bound are shown for each flux and each experiment. Additional results appear in Tables E.10, E.11, and E.12.

Flux	L1	L2	H1	H2
	84.19	117.15	125.58	127.33
Pyr → $AcCoA_{mit}$ + CO2	**87.35**	**120.68**	**130.62**	**130.55**
	93.76	124.83	134.67	135.10
	39.42	14.28	40.14	0.60
G6P → P5P	**41.61**	**16.37**	**43.43**	**1.95**
	43.45	17.66	44.05	6.59
	12.13	4.79	12.99	0.00
2 P5P → S7P + GAP	**12.79**	**5.20**	**13.99**	**0.42**
	13.40	5.65	14.64	1.99
	0.00	0.70	0.00	26.02
2 P5P ↔ S7P + GAP	**0.00**	**2.48**	**0.00**	**30.88**
	2.14	3.12	3.36	32.77
	12.13	4.59	13.56	0.00
S7P + GAP → F6P + E4P	**12.79**	**5.20**	**13.99**	**0.42**
	13.40	5.63	14.07	1.95
	30.48	0.00	1403.00	0.00
S7P + GAP ↔ F6P + E4P	**33.56**	**0.00**	**50364.68**	**0.00**
	38.14	1.09	Inf	3.10
	8.05	3.82	11.55	-0.71
P5P + E4P → F6P + GAP	**9.75**	**4.49**	**12.63**	**-0.24**
	10.37	4.91	12.85	1.31
	112.39	49.56	62.29	88.45
P5P + E4P ↔ F6P + GAP	**128.75**	**53.94**	**71.22**	**97.50**
	144.21	64.29	76.41	110.59
	93.87	119.43	128.87	129.19
$AcCoA_{mit}$ + OAA_{mit} → Cit	**95.04**	**122.22**	**134.09**	**132.67**
	101.59	124.32	138.09	137.25
	35.70	34.76	16.90	55.36
Cit → AKG	**38.70**	**37.00**	**19.18**	**59.10**
	41.05	37.84	20.63	61.44
	0.00	0.00	0.00	0.00
Cit ↔ AKG	**0.00**	**0.00**	**0.00**	**0.00**
	4.05	1.48	0.37	9.41
	30.43	32.98	16.26	53.95
AKG → Suc	**33.40**	**35.75**	**16.80**	**57.96**
	35.65	37.75	20.11	60.89
	30.43	32.98	16.26	53.95
Suc → Fum	**33.40**	**35.75**	**16.80**	**57.96**
	35.65	37.75	20.11	60.89
	0.00	34.25	1.70	23.48
Suc → Fum	**1.81**	**41.74**	**3.55**	**35.86**
	3.65	45.15	6.74	47.56
	33.55	34.15	15.57	54.33
Fum → Mal_{mit}	**36.49**	**36.48**	**18.18**	**58.62**
	38.77	39.27	21.75	61.27
	200.52	2954.18	92.53	228.16
Fum ↔ Mal_{mit}	**334.25**	**333276.21**	**133.56**	**285.12**
	489.91	Inf	225.56	344.92
	93.87	119.43	128.87	129.19
Mal_{mit} → OAA_{mit}	**95.04**	**122.22**	**134.09**	**132.67**
	101.59	124.32	138.09	137.25
	0.00	0.00	0.00	0.00
Mal_{mit} ↔ OAA_{mit}	**0.00**	**54.93**	**43.58**	**0.00**
	12.95	Inf	Inf	Inf

Table E.10: NMFA estimation results for *Y. lipolytica* at low aeration & linear growth, low aeration & stationary growth, high aeration & linear growth, and high aeration & stationary growth (L1, L2, H1, H2). Values are normalized to a glucose input flux of 100. The lower bound, the estimated value (bold), and upper bound are shown for each flux and each experiment. Additional results appear in Tables E.9, E.11, and E.12.

Flux	L1	L2	H1	H2
	0.00	0.00	0.00	0.00
Pyr → OAA	**0.00**	**0.00**	**0.00**	**24.10**
	19.66	18.51	10.48	50.33
	57.38	79.87	117.35	86.05
OAA → Mal	**58.56**	**85.82**	**119.61**	**93.49**
	74.05	92.30	121.58	98.19
	1534.31	169.66	333.41	1727.69
OAA ↔ Mal	**19054.26**	**202.93**	**5863.76**	**44923.97**
	Inf	236.87	Inf	Inf
	0.00	0.00	1.65	11.99
Mal → Pyr	**0.00**	**0.08**	**3.70**	**19.45**
	5.28	3.69	6.18	23.59
	-4.77	34.41	0.55	20.64
PEP → OAA	**14.23**	**53.17**	**11.69**	**52.42**
	20.58	55.72	14.53	55.51
	0.00	38.06	0.00	0.00
PEP ↔ OAA	**4.41**	**50.21**	**0.00**	**0.00**
	5.95	63.06	1.19	7.87
	55.11	30.37	106.09	14.76
Cit → AcCoA + OAA	**56.35**	**35.48**	**113.68**	**19.90**
	67.05	41.69	114.14	27.32
	54.29	85.62	114.83	69.19
Mal → Mal_{mit}	**58.56**	**85.75**	**115.91**	**74.04**
	70.46	92.40	121.21	82.77
	2530.03	1495.40	616.42	1927.79
Mal ↔ Mal_{mit}	**55657.14**	**143366.94**	**13039.48**	**36558.43**
	Inf	Inf	Inf	Inf
	0.97	0.20	0.41	0.18
MEETHF → METHF	**0.99**	**0.23**	**0.44**	**0.21**
	1.03	0.24	0.47	0.24
	1.84	0.39	0.79	0.36
MEETHF → FTHF	**1.91**	**0.45**	**0.86**	**0.41**
	2.00	0.50	0.92	0.47
	26.20	5.60	11.42	5.02
AKG → Glu	**26.84**	**6.46**	**12.41**	**5.90**
	28.26	7.19	13.33	6.71
	5.22	1.08	2.28	0.96
Glu → Gln	**5.33**	**1.26**	**2.39**	**1.15**
	5.55	1.32	2.55	1.29
	0.87	0.18	0.37	0.16
Glu → Pro	**0.90**	**0.21**	**0.41**	**0.19**
	0.94	0.22	0.43	0.22
	1.14	0.24	0.49	0.21
Glu → Arg	**1.17**	**0.28**	**0.53**	**0.25**
	1.23	0.31	0.56	0.28
	10.20	2.11	4.62	2.22
OAA → Asp	**10.43**	**2.46**	**5.04**	**2.59**
	10.78	2.65	5.44	2.95
	1.28	0.27	0.55	0.24
Asp → Asn	**1.33**	**0.31**	**0.60**	**0.29**
	1.39	0.33	0.64	0.33
	0.95	0.20	0.41	0.18
Pyr → Ala	**0.99**	**0.23**	**0.44**	**0.21**
	1.03	0.24	0.46	0.24

Table E.11: NMFA estimation results for *Y. lipolytica* at low aeration & linear growth, low aeration & stationary growth, high aeration & linear growth, and high aeration & stationary growth (L1, L2, H1, H2). Values are normalized to a glucose input flux of 100. The lower bound, the estimated value (bold), and upper bound are shown for each flux and each experiment. Additional results appear in Tables E.9, E.10, and E.12.

Flux	L1	L2	H1	H2
	1.29	0.43	0.49	0.08
3PG → Ser	**1.50**	**0.49**	**0.66**	**0.10**
	1.53	0.56	0.86	0.20
	-1.64	-0.28	-0.90	-0.65
Ser → Gly	**-1.60**	**-0.24**	**-0.73**	**-0.57**
	-1.58	-0.18	-0.54	-0.48
	3.34	1.37	1.82	0.00
Ser ↔ Gly	**3.49**	**1.58**	**3.11**	**0.00**
	4.08	1.87	4.75	0.63
	6.51	1.27	2.67	1.41
Gly → MEETHF	**6.57**	**1.41**	**2.96**	**1.65**
	6.76	1.47	3.19	1.85
	0.00	0.00	0.00	0.00
Gly ↔ MEETFH	**0.00**	**0.00**	**0.00**	**0.00**
	0.08	0.08	0.02	0.02
	9.26	1.64	3.73	2.04
Thr → Gly	**9.35**	**1.93**	**4.22**	**2.47**
	9.54	2.01	4.63	2.79
	1.14	0.24	0.49	0.22
Ser → Cys	**1.17**	**0.28**	**0.53**	**0.25**
	1.22	0.30	0.56	0.28
	0.64	0.13	0.27	0.12
Asp + Pyr → Lys	**0.65**	**0.15**	**0.29**	**0.14**
	0.67	0.17	0.31	0.16
	2.93	0.61	1.54	0.88
Asp → Thr	**3.05**	**0.72**	**1.72**	**1.00**
	3.19	0.75	1.94	1.13
	0.97	0.20	0.41	0.18
Asp → Met	**0.99**	**0.23**	**0.44**	**0.21**
	1.03	0.24	0.47	0.24
	0.34	0.07	0.15	0.06
2 Pyr → Val	**0.36**	**0.08**	**0.16**	**0.08**
	0.37	0.09	0.17	0.09
	1.60	0.34	0.69	0.30
$AcCoA_{mit}$ + 2 Pyr → Leu	**1.66**	**0.39**	**0.74**	**0.36**
	1.74	0.41	0.80	0.40
	1.05	0.22	0.45	0.20
Thr + Pyr → Ile	**1.09**	**0.26**	**0.49**	**0.24**
	1.14	0.27	0.52	0.27
	0.82	0.17	0.35	0.15
2 PEP + E4P → Phe	**0.84**	**0.20**	**0.38**	**0.18**
	0.88	0.20	0.40	0.20
	1.30	0.28	0.57	0.36
2 PEP + E4P → Tyr	**1.36**	**0.32**	**0.61**	**0.29**
	1.43	0.35	0.65	0.47
	0.82	0.17	0.35	0.15
Ser + P5P + PEP + E4P → Trp	**0.84**	**0.20**	**0.38**	**0.18**
	0.88	0.20	0.40	0.20
	1.84	0.39	0.79	0.36
P5P + FTHF → His	**1.91**	**0.45**	**0.86**	**0.41**
	2.00	0.50	0.92	0.47

Table E.12: NMFA estimation results for *Y. lipolytica* at low aeration & linear growth, low aeration & stationary growth, high aeration & linear growth, and high aeration & stationary growth (L1, L2, H1, H2). Values are normalized to a glucose input flux of 100. The lower bound, the estimated value (bold), and upper bound are shown for each flux and each experiment. Additional results appear in Tables E.9, E.10, and E.11.

Appendix F

Abbreviations

3PG	3-phosphoglycerate
6PG	6-phosphoglycerate
AcCoA	acetyl coenzyme A
AKG	α-ketoglutarate
Ala	alanine
Arg	arginine
Asn	asparagine
Asp	aspartate
C16	16-carbon fatty acid
C18	18-carbon fatty acid
CHO	chinese hamster ovary
CI	confidence interval
Cit	citrate
CO_2	carbon dioxide
Cys	cysteine
cyt	cytosolic
DHAP	dihydroxyacetone phosphate
dil	dilution
DMEM	Dulbecco's modified Eagle's medium
E4P	erythrose 4-phosphate

EC_2	transketolase 2-carbon intermediate
EC_3	transaldolase 3-carbon intermediate
EMU	elementary metabolite unit
exp	exponential
ext	extracellular
F6P	fructose 6-phosphate
FA	fatty acid
FBP	fructose 1,6-bisphosphate
Frag	metabolite fragment
FTHF	formyltetrahydrofolate
Fum	fumarate
G6P	glucose 6-phosphate
GAP	glyceraldehyde 3-phosphate
GC	gas chromatography
Gln	glutamine
GLP	glycerol 3-phosphate
Glu	glutamate
Gluc	glucose
Gly	glycine
Glyc	glycerol
Glyc3P	glycerol 3-phosphate
H1	linear phase high-aeration experiment
H2	stationary phase high-aeration experiment
HCMV	human cytomegalovirus
His	histidine
HPA	3-hydroxypropionaldehyde
HPLC	high-performance liquid chromatography
ICit	isocitrate
IDH	isocitrate dehydrogenase
Ile	isoleucine

KDPG	2-keto-3-deoxy-6-phosphogluconate
L1	linear phase low-aeration experiment
L2	stationary phase low-aeration experiment
Lac	lactate
lb	lower bound
LC	liquid chromatography
Leu	leucine
lin	linear
LL-DAP	LL-diaminopimelate
LNS	long nonstationary experiment
LNS w/ C	long nonstationary experiment with concentrations
Mal	malate
MEETHF	methylenetetrahydrofolate
Met	methionine
METHF	methyltetrahydrofolate
MetranCL	command line metabolic tracer analysis
MFA	metabolic flux analysis
MID	mass isotopomer distribution
mit	mitochondrial
mRNA	messenger ribonucleic acid
MS	mass spectroscopy
MSTFA	N-methyl-N-trifluoroacetamide
MTBSTFA	N-methyl-N-(tert-butyldimethyl-silyl)trifluoroacetimide
NADPH	reduced nicotinamide adenine dinucleotide phosphate
NMFA	isotopically nonstationary metabolic flux analysis
NMR	nuclear magnetic resonance
NTP	nucleotide phosphate
O	original value
OAA	oxaloacetate
OD	optical density

P5P	pentose 5-phosphate
Palm	palmitate
PCI	peripheral component interconnect
PDH	pyruvate dehydrogenase
PDO	1,3-propanediol
PEP	phosphoenolpyruvate
Phe	phenylalanine
PPP	pentose phosphate pathway
Pro	proline
Pyr	pyruvate
R5P	ribose 5-phosphate
ref	reference
RNAi	ribonucleic acid interference
Ru5P	ribulose 5-phosphate
S	stationary experiment
S7P	seduheptulose 7-phosphate
Ser	serine
SIM	selected ion monitoring
SNS	short nonstationary experiment
SNS w/ C	short nonstationary experiment with concentrations
SSR	sum of squared residuals
stdev	standard deviation
Suc	succinate
SucCoA	succinyl coenzyme A
tag	triacylglyceride
TBDMCS	tert-butyldimethylchlorosilane
TBDMS	tert-butyldimethylsilyl
TCA	tricarboxylic acid
Thr	threonine
TMCS	trimethylchlorosilane

TMS	trimethylsilyl
Trp	tryptophan
Tyr	tyrosine
U	uniform
ub	upper bound
Val	valine
X5P	xylulose 5-phosphate
xch	exchange
xpt	experiment
YNB	yeast nitrogen base

Bibliography

[1] PM Abou-Sleiman, MMK Muqit, and NW Wood. Expanding insights of mitochondrial dysfunction in Parkinson's disease. *Nat Rev Neurosci*, 7(3):207–219, 2006.

[2] B Alberts, D Bray, J Lewis, M Raff, K Roberts, and JD Watson. *Molecular biology of the cell.* New York: Garland Science, 2002.

[3] DK Allen, Y Shachar-Hill, and JB Ohlrogge. Compartment-specific labeling information in ^{13}C metabolic flux analysis of plants. *Phytochemistry*, 68(16-18):2197–2210, 2007.

[4] H Alper, C Fischer, E Nevoigt, and G Stephanopoulos. Tuning genetic control through promoter engineering. *P Natl Acad Sci USA*, 102(36):12678, 2005.

[5] H Alper, YS Jin, JF Moxley, and G Stephanopoulos. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab Eng*, 7(3):155–164, 2005.

[6] H Alper, K Miyaoku, and G Stephanopoulos. Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. *Nat Biotechnol*, 23(5):612–616, 2005.

[7] H Alper, K Miyaoku, and G Stephanopoulos. Characterization of lycopene-overproducing *E. coli* strains in high cell density fermentations. *Appl Microbiol Biot*, 72(5):968–974, 2006.

[8] H Alper, J Moxley, E Nevoigt, GR Fink, and G Stephanopoulos. Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science*, 314(5805):1565, 2006.

[9] H Alper and G Stephanopoulos. Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metab Eng*, 9(3):258–267, 2007.

[10] MR Antoniewicz, JK Kelleher, and G Stephanopoulos. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab Eng*, 8(4):324–337, 2006.

[11] MR Antoniewicz, JK Kelleher, and G Stephanopoulos. Elementary metabolite units (EMU): A novel framework for modeling isotopic distributions. *Metab Eng*, 9(1):68–86, 2007.

[12] MR Antoniewicz, DF Kraynie, LA Laffend, J González-Lergier, JK Kelleher, and G Stephanopoulos. Metabolic flux analysis in a nonstationary system: fed-batch fermentation of a high yielding strain of E. coli producing 1,3-propanediol. *Metab Eng*, 9(3):277–292, 2007.

[13] AI Archakov, YD Ivanov, AV Lisitsa, and VG Zgoda. AFM fishing nanotechnology is the way to reverse the Avogadro number in proteomics. *Proteomics*, 7(1):4–9, 2007.

[14] AA Aristidou, KY San, and GN Bennett. Metabolic engineering of *Escherichia coli* to enhance recombinant protein production through acetate reduction. *Biotechnol Progr*, 11(4):475–478, 1995.

[15] JE Bailey. Toward a science of metabolic engineering. *Science*, 252(5013):1668, 1991.

[16] G Barth and C Gaillardin. Physiology and genetics of the dimorphic fungus *Yarrowia lipolytica*. *FEMS Microbiol Rev*, 19(4):219–237, 2006.

[17] G Barth and W Künkel. Alcohol dehydrogenase (ADH) in yeasts. II. NAD+- and NADP+-dependent alcohol dehydrogenases in *Saccharomycopsis lipolytica*. *Z Allg Mikrobiol*, 19(6):381–390, 2007.

[18] CJ Baxter, JL Liu, AR Fernie, and LJ Sweetlove. Determination of metabolic fluxes in a non-steady-state system. *Phytochemistry*, 68(16–18):2313–2319, 2007.

[19] SB Biddinger and CR Kahn. From mice to men: Insights into the insulin resistance syndromes. *Annu Rev Physiol*, 68:123–158, 2006.

[20] HW Blanch, PD Adams, KM Andrews-Cramer, WB Frommer, BA Simmons, and JD Keasling. Addressing the need for alternative transportation fuels: the Joint Bioenergy Institute. *ACS Chem Biol*, 3(1):17–20, 2008.

[21] LM Blank, F Lehmbeck, and U Sauer. Metabolic-flux and network analysis in fourteen hemiascomycetous yeasts. *FEMS Yeast Res*, 5(6-7):545–558, 2005.

[22] J Börner, S Buchinger, and D Schomburg. A high-throughput method for microbial metabolome analysis using gas chromatography/mass spectrometry. *Anal Biochem*, 367(2):143–151, 2007.

[23] LG Boros, MR Lerner, DL Morgan, SL Taylor, BJ Smith, RG Postier, and DJ Brackett. [1,2-$^{13}C_2$]-D-glucose profiles of the serum, liver, pancreas, and DMBA-induced pancreatic tumors of rats. *Pancreas*, 31(4):337–343, 2005.

[24] LG Boros, MP Steinkamp, JC Fleming, WNP Lee, M Cascante, and EJ Neufeld. Defective RNA ribose synthesis in fibroblasts from patients with thiamine-responsive megaloblastic anemia (TRMA). *Blood*, 102(10):3556, 2003.

[25] GA Bray and T Bellanger. Epidemiology, trends, and morbidities of obesity and the metabolic syndrome. *Endocrine*, 29(1):109–117, 2006.

[26] PM Bruinenberg, JP Van Dijken, and WA Scheffers. A theoretical analysis of NADPH production and consumption in yeasts. *Microbiology*, 129(4):953–964, 1983.

[27] A Buchholz, R Takors, and C Wandrey. Quantification of intracellular metabolites in *Escherichia coli* K12 using liquid chromatographic-electrospray ionization tandem mass spectrometric techniques. *Anal Biochem*, 295(2):129–137, 2001.

[28] SC Burgess, TT He, Z Yan, J Lindner, AD Sherry, CR Malloy, JD Browning, and MA Magnuson. Cytosolic phosphoenolpyruvate carboxykinase does not solely control the rate of hepatic gluconeogenesis in the intact mouse liver. *Cell Metab*, 5(4):313–320, 2007.

[29] S Buziol, I Bashir, A Baumeister, W Claaßen, N Noisommit-Rizzi, W Mailinger, and M Reuss. New bioreactor-coupled rapid stopped-flow sampling technique for measurements of metabolite dynamics on a subsecond time scale. *Biotechnology Bioeng*, 80(6):632–636, 2002.

[30] B Cannon and J Nedergaard. Brown adipose tissue: function and physiological significance. *Physiol Rev*, 84(1):277–359, 2004.

[31] A Cederberg, LM Grønning, B Ahrén, K Taskén, P Carlsson, and S Enerbäck. FOXC2 is a winged helix gene that counteracts obesity, hypertriglyceridemia, and diet-induced insulin resistance. *Cell*, 106(5):563–573, 2001.

[32] YJ Chang, PF Suthers, and CD Maranas. Identification of optimal measurement sets for complete flux elucidation in metabolic flux analysis experiments. *Biotechnol Bioeng*, 100(6):1039–1049, 2008.

[33] JC Chatham, B Bouchard, and C Des Rosiers. A comparison between NMR and GCMS ^{13}C-isotopomer analysis in cardiac metabolism. *Mol Cell Biochem*, 249(1):105–112, 2003.

[34] B Christensen and J Nielsen. Isotopomer analysis using GC-MS. *Metab Eng*, 1(4):282–290, 1999.

[35] HR Christofk, MG Vander Heiden, MH Harris, A Ramanathan, RE Gerszten, R Wei, MD Fleming, SL Schreiber, and LC Cantley. The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature*, 452(7184):230–233, 2008.

[36] LB Crittenden and DW Salter. Gene insertion: Current progress and long-term goals. *Avian Dis*, 30(1):43–46, 1986.

[37] M Dauner, JE Bailey, and U Sauer. Metabolic flux analysis with a comprehensive isotopomer model in *Bacillus subtilis*. *Biotechnol Bioeng*, 76(2):144–156, 2001.

[38] M Dauner and U Sauer. GC-MS analysis of amino acids rapidly provides rich information for isotopomer balancing. *Biotechnol Progr*, 16(4):642–649, 2000.

[39] H David, AM Krogh, C Roca, M Akesson, and J Nielsen. CreA influences the metabolic fluxes of *Aspergillus nidulans* during growth on glucose and xylose. *Microbiology*, 151(7):2209–2221, 2005.

[40] W de Koning and K Dam. A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH. *Anal Biochem*, 204(1):118–123, 1992.

[41] RJ DeBerardinis, A Mancuso, E Daikhin, I Nissim, M Yudkoff, S Wehrli, and C B Thompson. Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *P Natl Acad Sci USA*, 104(49):19345–19350, 2007.

[42] RJ DeBerardinis, N Sayed, D Ditsworth, and CB Thompson. Brick by brick: metabolism and tumor cell growth. *Curr Opin Genet Dev*, 18(1):54–61, 2008.

[43] NC Duarte, MJ Herrgård, and BØ Palsson. Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. *Genome Research*, 14(7):1298–1309, 2004.

[44] AL Dulmage and NS Mendelsohn. Coverings of bipartite graphs. *Canad J Math*, 10:517–534, 1958.

[45] WB Dunn. Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes. *Phys Biol*, 5(1):11001–11024, 2008.

[46] WB Dunn and DI Ellis. Metabolomics: current analytical platforms and methodologies. *Trend Anal Chem*, 24(4):285–294, 2005.

[47] M Emmerling, M Dauner, A Ponti, J Fiaux, M Hochuli, T Szyperski, K Wuthrich, JE Bailey, and U Sauer. Metabolic flux responses to pyruvate kinase knockout in *Escherichia coli*. *J Bacteriol*, 184(1):152–164, 2002.

[48] VR Fantin, J St-Pierre, and P Leder. Attenuation of LDH-A expression uncovers a link between glycolysis, mitochondrial physiology, and tumor maintenance. *Cancer Cell*, 9(6):425–434, 2006.

[49] ME Feder and JC Walser. The biological limitations of transcriptomics in elucidating stress and stress responses. *J Evolution Biol*, 18(4):901–910, 2005.

[50] CA Fernandez, C Des Rosiers, SF Previs, F David, and H Brunengraber. Correction of ^{13}C mass isotopomer distributions for natural stable isotope abundance. *J Mass Spectrom*, 31(3):255–262, 1998.

[51] O Fiehn. Metabolomics–the link between genotypes and phenotypes. *Plant Mol Biol*, 48(1):155–171, 2002.

[52] O Fiehn, J Kopka, P Dormann, T Altmann, RN Trethewey, and L Willmitzer. Metabolite profiling for plant functional genomics. *Nat Biotechnol*, 18(11):1157–1161, 2000.

[53] O Fiehn, J Kopka, RN Trethewey, and L Willmitzer. Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. *Anal Chem*, 72(15):3573–3580, 2000.

[54] A Fire, S Xu, MK Montgomery, SA Kostas, SE Driver, and CC Mello. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, 391(6669):806–810, 1998.

[55] E Fischer and U Sauer. Metabolic flux profiling of *Escherichia coli* mutants in central carbon metabolism using GC-MS. *Eur J Biochem*, 270(5):880–891, 2003.

[56] J Folch, M Lees, and GHS Stanley. A simple method for the isolation and purification of total lipides from animal tissues. *J Biol Chem*, 226(1):497–509, 1957.

[57] NS Forbes, AL Meadows, DS Clark, and HW Blanch. Estradiol stimulates the biosynthetic pathways of breast cancer cells: Detection by metabolic flux analysis. *Metab Eng*, 8(6):639–652, 2006.

[58] ES Ford, WH Giles, and WH Dietz. Prevalence of the metabolic syndrome among US adults: findings from the third National Health and Nutrition Examination Survey. *J Am Med Assoc*, 287(3):356–359, 2002.

[59] M Fussenegger, S Schlatter, D Dätwyler, X Mazur, and JE Bailey. Controlled proliferation by multigene metabolic engineering enhances the productivity of Chinese hamster ovary cells. *Nat Biotechnol*, 16(5):468–472, 1998.

[60] PE Gill, W Murray, and MH Wright. *Practical optimization*. London: Academic Press, 1981.

[61] DE Goldberg and K Deb. A comparative analysis of selection schemes used in genetic algorithms. *Lect Notes Comput Sc*, 1:69–93, 1991.

[62] GH Golub and CF Van Loan. *Matrix Computations.* Baltimore: Johns Hopkins University Press, 1996.

[63] AK Gombert, M Moreira dos Santos, B Christensen, and J Nielsen. Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J Bacteriol*, 183(4):1441–1451, 2001.

[64] B Gonzalez, J François, and M Renaud. A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol. *Yeast*, 13(14):1347–1355, 1997.

[65] C Goudar, R Biener, C Boisart, R Heidemann, J Piret, A de Graaf, and K Konstantinov. Metabolic flux analysis of CHO cells in perfusion culture by metabolite balancing and 2D [^{13}C, ^{1}H] COSY NMR spectroscopy. *Metab Eng*, 12(2):138–149, 2009.

[66] L Guevara-Olvera, C Calvo-Mendez, and J Ruiz-Herrera. The role of polyamine metabolism in dimorphism of *Yarrowia lipolytica*. *Microbiology*, 139(3):485–493, 1993.

[67] G Hatzivassiliou, F Zhao, DE Bauer, C Andreadis, AN Shaw, D Dhanak, SR Hingorani, DA Tuveson, and CB Thompson. ATP citrate lyase inhibition can suppress tumor cell growth. *Cancer Cell*, 8(4):311–321, 2005.

[68] MK Hellerstein. New stable isotope-mass spectrometric techniques for measuring fluxes through intact metabolic pathways in mammalian systems: introduction of moving pictures into functional genomics and biochemical phenotyping. *Metab Eng*, 6(1):85–100, 2004.

[69] MK Hellerstein, M Christiansen, S Kaempfer, C Kletke, K Wu, JS Reid, K Mulligan, NS Hellerstein, and CH Shackleton. Measurement of *de novo* hepatic lipogenesis in humans using stable isotopes. *J Clin Invest*, 87(5):1841–1852, 1991.

[70] J Hill, E Nelson, D Tilman, S Polasky, and D Tiffany. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *P Natl Acad Sci USA*, 103(30):11206–11210, 2006.

[71] U Hofmann, K Maier, A Niebel, G Vacun, M Reuss, and K Mauch. Identification of metabolic fluxes in hepatic cells from transient ^{13}C-labeling experiments: Part I. Experimental observations. *Biotechnol and Bioeng*, 100(2):344–354, 2008.

[72] AL Holleran, G Fiskum, and JK Kelleher. Quantitative analysis of acetoacetate metabolism in AS-30D hepatoma cells with ^{13}C and ^{14}C isotopic techniques. *Am J Physiol*, 272(6):E945–E951, 1997.

[73] MA Hoque, H Ushiyama, M Tomita, and K Shimizu. Dynamic responses of the intracellular metabolite concentrations of the wild type and *pykA* mutant *Escherichia coli* against pulse addition of glucose or NH_3 under those limiting continuous cultures. *Biochem Eng J*, 26(1):38–49, 2005.

[74] Q Hua, C Yang, T Baba, H Mori, and K Shimizu. Responses of the central metabolism in *Escherichia coli* to phosphoglucose isomerase and glucose-6-phosphate dehydrogenase knockouts. *J Bacteriol*, 185(24):7053–7067, 2003.

[75] T Ishikura and JW Foster. Incorporation of molecular oxygen during microbial utilization of olefins. *Nature*, 192:892–893, 1961.

[76] T Kamiryo, Y Nishikawa, M Mishina, M Terao, and S Numa. Involvement of long-chain acyl coenzyme A for lipid synthesis in repression of acetyl-coenzyme A carboxylase in *Candida lipolytica*. *P Natl Acad Sci USA*, 76(9):4390–4394, 1979.

[77] J Katz, P Wals, and WN Lee. Isotopomer studies of gluconeogenesis and the Krebs cycle with ^{13}C-labeled lactate. *J Biol Chem*, 268(34):25509–2521, 1993.

[78] AT Kharroubi, TM Masterson, TA Aldaghlas, KA Kennedy, and JK Kelleher. Isotopomer spectral analysis of triglyceride fatty acid synthesis in 3T3-L1 cells. *Am J Physiol*, 263(4):E667–E675, 1992.

[79] P Kiefer, E Heinzle, O Zelder, and C Wittmann. Comparative metabolic flux analysis of lysine-producing *Corynebacterium glutamicum* cultured on glucose or fructose. *Appl Environ Microb*, 70(1):229–239, 2004.

[80] RJ Kleijn, JM Geertman, BK Nfor, C Ras, D Schipper, JT Pronk, JJ Heijnen, AJ van Maris, and WA van Winden. Metabolic flux analysis of a glycerol-overproducing *Saccharomyces cerevisiae* strain based on GC-MS, LC-MS and NMR-derived C-labelling data. *FEMS Yeast Res*, 7(2):216–231, 2007.

[81] M Klingenberg and KS Echtay. Uncoupling proteins: the issues from a biochemist point of view. *BBA Bioenergetics*, 1504(1):128–143, 2001.

[82] MJ Klug and AJ Markovetz. Degradation of hydrocarbons by members of the genus *Candida*. II. Oxidation of *n*-alkanes and 1-alkenes by *Candida lipolytica*. *J Bacteriol*, 93(6):1847–1852, 1967.

[83] G Kroemer and J Pouyssegur. Tumor cell metabolism: cancer's Achilles' heel. *Cancer Cell*, 13(6):472–482, 2008.

[84] HC Lange, M Eman, G van Zuijlen, D Visser, JC van Dam, J Frank, MJ de Mattos, and JJ Heijnen. Improved rapid sampling for in vivo kinetics of intracellular metabolites in *Saccharomyces cerevisiae*. *Biotechnol Bioeng*, 75(4):406–415, 2001.

[85] CH Lee, P Olson, and RM Evans. Minireview: lipid metabolism, metabolic diseases, and peroxisome proliferator-activated receptors. *Endocrinology*, 144(6):2201, 2003.

[86] SK Lee, H Chou, TS Ham, TS Lee, and JD Keasling. Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Curr Opin Biotech*, 19(6):556–563, 2008.

[87] M Li, PY Ho, S Yao, and K Shimizu. Effect of *lpdA* gene knockout on the metabolism in *Escherichia coli* based on enzyme activities, intracellular metabolite concentrations and metabolic flux analysis by ^{13}C-labeling experiments. *J Biotech*, 122(2):254–266, 2006.

[88] VI Litvinenko, JA Burgher, VS Vyshemirskij, and NA Sokolova. Application of genetic algorithm for optimization gasoline fractions blending compounding. *P IEEE ICAIS*, pages 391–394, 2002.

[89] BB Lowell and JS Flier. Brown adipose tissue, β3-adrenergic receptors, and obesity. *Ann Rev Med*, 48(1):307–316, 1997.

[90] K Madsen, HB Nielsen, and O Tingleff. Methods for non-linear least squares problems, 2nd edition. *Inform Math Model*, 2004.

[91] PR Maharjan and T Ferenci. Global metabolite analysis: the influence of extraction methodology on metabolome profiles of *Escherichia coli*. *Anal Biochem*, 313(1):145–154, 2003.

[92] K Maier, U Hofmann, M Reuss, and K Mauch. Identification of metabolic fluxes in hepatic cells from transient ^{13}C-labeling experiments: Part II. Flux estimation. *Biotechnol Bioeng*, 100(2):355–370, 2008.

[93] CR Malloy, AD Sherry, and FM Jeffrey. Evaluation of carbon flux and substrate selection through alternate pathways involving the citric acid cycle of the heart by ^{13}C NMR spectroscopy. *J Biol Chem*, 263(15):6964, 1988.

[94] A Marx, AA de Graaf, W Wiechert, L Eggeling, and H Sahm. Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing. *Biotechnol Bioeng*, 49(2):111–129, 1996.

[95] MR Mashego, L Wu, JC van Dam, C Ras, JL Vinke, WA van Winden, WM van Gulik, and JJ Heijnen. MIRACLE: mass isotopomer ratio analysis of U-^{13}C-labeled extracts. A new method for accurate quantification of changes in concentrations of intracellular metabolites. *Biotechnol Bioeng*, 85(6):620–628, 2004.

[96] M Mattey. The production of organic acids. *Crit Rev Biotechnol*, 12(1):87–132, 1992.

[97] JB McKinlay and C Vieille. ^{13}C-metabolic flux analysis of *Actinobacillus succinogenes* fermentative metabolism at different $NaHCO_3$ and H_2 concentrations. *Metab Eng*, 10(1):55–68, 2008.

[98] CM Metallo, JL Walther, and G Stephanopoulos. Evaluation of ^{13}C isotopic tracers for metabolic flux analysis in mammalian cells. *J Biotechnol*, 144(3):167–174, 2009.

[99] M Möllney, W Wiechert, D Kownatzki, and AA de Graaf. Bidirectional reaction steps in metabolic networks: IV. Optimal design of isotopomer labeling experiments. *Biotechnol Bioeng*, 66(2):86–103, 1999.

[100] JC Moore and FH Arnold. Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. *Nat Biotechnol*, 14(4):458–467, 1996.

[101] JF Moxley, MC Jewett, MR Antoniewicz, SG Villas-Bôas, H Alper, RT Wheeler, L Tong, AG Hinnebusch, T Ideker, J Nielsen, and G Stephanopoulos. Linking high-resolution metabolic flux phenotypes and transcriptional regulation in yeast modulated by the global regulator Gcn4p. *P Natl Acad Sci USA*, 106(16):6477–6482, 2009.

[102] J Munger, BD Bennett, A Parikh, XJ Feng, J McArdle, HA Rabitz, T Shenk, and JD Rabinowitz. Systems-level metabolic flux profiling identifies fatty acid synthesis as a target for antiviral therapy. *Nat Biotechnol*, 26(10):1179–1186, 2008.

[103] RA Neese, D Faix, C Kletke, K Wu, AC Wang, CH Shackleton, and MK Hellerstein. Measurement of endogenous synthesis of plasma cholesterol in rats and humans using MIDA. *Am J Physiol*, 264(1):E136–E147, 1993.

[104] J Nielsen. It is all about metabolic fluxes. *J Bacteriol*, 185(24):7031–7035, 2003.

[105] Y Noguchi, JD Young, JO Aleman, ME Hansen, JK Kelleher, and G Stephanopoulos. Effect of anaplerotic fluxes and amino acid availability on hepatic lipoapoptosis. *J Biol Chem*, 284(48):33425–33436, 2009.

[106] 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}C labeling experiments. *J Biotechnol*, 129(2):249–267, 2007.

[107] K Nöh, A Wahl, and W Wiechert. Computational tools for isotopically instationary ^{13}C labeling experiments under metabolic steady state conditions. *Metab Eng*, 8(6):554–577, 2006.

[108] K Nöh and W Wiechert. Experimental design principles for isotopically instationary ^{13}C labeling experiments. *Biotechnol Bioeng*, 94(2):234–251, 2006.

[109] EJ Nyns, JP Auquiere, N Chiang, and AL Wiaux. Comparative growth of *Candida lipolytica* on glucose and *n*-hexadecane. *Nature*, 215:177–178, 1967.

[110] T Okayasu, M Ikeda, K Akimoto, and K Sorimachi. The amino acid composition of mammalian and bacterial cells. *Amino Acids*, 13(3):379–391, 1997.

[111] M Oogaki, T Nakahara, H Uchiyama, and T Tabuchi. Extracellular production of D-(+)-2-hydroxyglutaric acid by *Yarrowia lipolytica* from glucose under aerobic, thiamine-deficient conditions. *Agr Biol Chem*, 47(11):2619–2624, 1983.

[112] S Papanikolaou, I Chevalot, M Komaitis, I Marc, and G Aggelis. Single cell oil production by *Yarrowia lipolytica* growing on an industrial derivative of animal fat in batch cultures. *Appl Microbiol Biot*, 58(3):308–312, 2002.

[113] D Pimentel and TW Patzek. Ethanol production using corn, switchgrass, and wood; biodiesel production using soybean and sunflower. *Nat Resource Res*, 14(1):65–76, 2005.

[114] A Pothen and CJ Fan. Computing the block triangular form of a sparse matrix. *T Math Soft*, 16(4):303–324, 1990.

[115] JD Powell, GF Franklin, and ML Workman. *Digital control of dynamic systems.* Reading, MA: Addison-Wesley, 1990.

[116] SF Previs, GW Cline, and GI Shulman. A critical evaluation of mass isotopomer distribution analysis of gluconeogenesis *in vivo*. *Am J Physiol Endoc M*, 277(1):E154–60, 1999.

[117] A Rantanen, T Mielikainen, J Rousu, H Maaheimo, and E Ukkonen. Planning optimal measurements of isotopomer distributions for estimation of metabolic fluxes. *Bioinformatics*, 22(10):1198–1206, 2006.

[118] C Ratledge. Fatty acid biosynthesis in microorganisms being used for single cell oil production. *Biochimie*, 86(11):807–815, 2004.

[119] C Ratledge and JP Wynn. The biochemistry and molecular biology of lipid accumulation in oleaginous microorganisms. *Adv Appl Microbiol*, 51:1–52, 2002.

[120] C Risso, SJ Van Dien, A Orloff, DR Lovley, and MV Coppi. Elucidation of an alternate isoleucine biosynthesis pathway in *Geobacter sulfurreducens*. *J Bacteriol*, 190(7):2266–2274, 2008.

[121] NJ Rothwell and MJ Stock. Combined effects of cafeteria and tube-feeding on energy balance in the rat. *Proc Nutr Soc*, 38(1):5A, 1979.

[122] U Sauer. Metabolic networks in motion: ^{13}C-based flux analysis. *Mol Syst Biol*, 2:62, 2006.

[123] U Sauer and JE Bailey. Estimation of P-to-O ratio in *Bacillus subtilis* and its influence on maximum riboflavin yield. *Biotechnol Bioeng*, 64(6):750–754, 1999.

[124] U Sauer, DR Lasko, J Fiaux, M Hochuli, R Glaser, T Szyperski, K Wuthrich, and JE Bailey. Metabolic flux ratio analysis of genetic and environmental modulations of *Escherichia coli* central carbon metabolism. *J Bacteriol*, 181(21):6679, 1999.

[125] U Schaefer, W Boos, R Takors, and D Weuster-Botz. Automated sampling device for monitoring intracellular metabolite dynamics. *Anal Biochem*, 270(1):88–96, 1999.

[126] J Schaub, K Mauch, and M Reuss. Metabolic flux analysis in *Escherichia coli* by integrating isotopic dynamic and isotopic stationary ^{13}C labeling data. *Biotechnol Bioeng*, 99(5):1170–1185, 2008.

[127] J Schaub, C Schiesling, M Reuss, and M Dauner. Integrated sampling procedure for metabolome analysis. *Biotechnol Progr*, 22(5):1434–1442, 2006.

[128] K Schmidt, J Nielsen, and J Villadsen. Quantitative analysis of metabolic fluxes in *Escherichia coli*, using two-dimensional NMR spectroscopy and complete isotopomer models. *J Biotechnol*, 71(1–3):175–189, 1999.

[129] P Seale, S Kajimura, and BM Spiegelman. Transcriptional control of brown adipocyte development and physiological function—of mice and men. *Gene Dev*, 23(7):788–797, 2009.

[130] BK Sekhon, RH Roubin, A Tan, WK Chan, and DM Sze. High-throughput screening platform for anticancer therapeutic drug cytotoxicity. *Assay Drug Devel Technol*, 6(5):711–722, 2008.

[131] AA Shastri and JA Morgan. A transient isotopic labeling methodology for ^{13}C metabolic flux analysis of photoautotrophic microorganisms. *Phytochemistry*, 68(16–18):2302–2312, 2007.

[132] GW Small, LM Ercoli, DHS Silverman, SC Huang, S Komo, SY Bookheimer, H Lavretsky, K Miller, P Siddarth, NL Rasgon, JC Mazziotta, S Saxena, HM Wu, MS Mega, JL Cummings, AM Saunders, MA Pericak-Vance, AD Roses, JR Barrio, and ME Phelps. Cerebral metabolic and cognitive decline in persons at genetic risk for Alzheimer's disease. *P Natl A Sci USA*, 97(11):6037–42, 2000.

[133] G Sriram, DB Fulton, and JV Shanks. Flux quantification in central carbon metabolism of *Catharanthus roseus* hairy roots by ^{13}C labeling and comprehensive bondomer balancing. *Phytochemistry*, 68(16-18):2243–2257, 2007.

[134] G Sriram, L Rahib, JS He, AE Campos, LS Parr, JC Liao, and KM Dipple. Global metabolic effects of glycerol kinase overexpression in rat hepatoma cells. *Mol Genet Metab*, 93(2):145–159, 2008.

[135] G. Stephanopoulos. Metabolic engineering. *Curr Opin Biotech*, 5(2):196–200, 1994.

[136] G Stephanopoulos. Metabolic fluxes and metabolic engineering. *Metab Eng*, 1(1):1–11, 1999.

[137] G Stephanopoulos. Challenges in engineering microbes for biofuels production. *Science*, 315(5813):801, 2007.

[138] G Stephanopoulos, AA Aristidou, and J Nielsen. *Metabolic engineering: principles and methodologies.* Academic Press, 1998.

[139] PF Suthers, AP Burgard, MS Dasika, F Nowroozi, S Van Dien, JD Keasling, and CD Maranas. Metabolic flux elucidation for large-scale models using ^{13}C labeled isotopes. *Metab Eng*, 9(5-6):387–405, 2007.

[140] Y Tang, F Pingitore, A Mukhopadhyay, R Phan, TC Hazen, and JD Keasling. Pathway confirmation and flux analysis of central metabolic pathways in *Desulfovibrio vulgaris* Hildenborough using gas chromatography-mass spectrometry and Fourier transform-ion cyclotron resonance mass spectrometry. *J Bacteriol*, 189(3):940–949, 2007.

[141] YJ Tang, JS Hwang, DE Wemmer, and JD Keasling. *Shewanella oneidensis* MR-1 fluxome under various oxygen conditions. *Appl Environ Microb*, 73(3):718–729, 2007.

[142] U Theobald, W Mailinger, M Baltes, M Rizzi, and M Reuss. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. *Biotechnol Bioeng*, 55(2):305–316, 1997.

[143] Y Toya, N Ishii, T Hirasawa, M Naba, K Hirai, K Sugawara, S Igarashi, K Shimizu, M Tomita, and T Soga. Direct measurement of isotopomer of intracellular metabolites using capillary electrophoresis time-of-flight mass spectrometry for efficient metabolic flux analysis. *J Chromatogr A*, 1159(1–2):134–141, 2007.

[144] R Tsugawa, T Nakase, T Koyabashi, K Yamashita, and S Okumura. Fermentation of *n*-paraffins by yeast. Part III. α-Ketoglutarate productivity of various yeasts. *Agr Biol Chem Tokyo*, 33:929–938, 1969.

[145] KE Tyo, HS Alper, and GN Stephanopoulos. Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends Biotechnol*, 25(3):132–137, 2007.

[146] J.J. Vallino and G. Stephanopoulos. Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine overproduction. *Biotechnology and Bioengineering*, 41:633–646, 1993.

[147] J Van Gerpen. Biodiesel processing and production. *Fuel Process Technol*, 86(10):1097–1107, 2005.

[148] WA van Winden, JC van Dam, C Ras, RJ Kleijn, JL Vinke, WM van Gulik, and JJ Heijnen. Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN. PK113-7D based on mass isotopomer measurements of ^{13}C-labeled primary metabolites. *FEMS Yeast Res*, 5(6–7):559–68, 2005.

[149] A Varma and BO Palsson. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl Environ Microb*, 60(10):3724–3731, 1994.

[150] R Vega and A Domínguez. Cell wall composition of the yeast and mycelial forms of *Yarrowia lipolytica*. *Arch Microbiol*, 144(2):124–130, 1986.

[151] SG Villas-Bôas and P Bruheim. Cold glycerol–saline: the promising quenching solution for accurate intracellular metabolite analysis of microbial cells. *Anal Biochem*, 370(1):87–97, 2007.

[152] SG Villas-Bôas, J Hojer-Pedersen, M Akesson, J Smedsgaard, and J Nielsen. Global metabolite analysis of yeast: evaluation of sample preparation methods. *Yeast*, 22(14):1155–1170, 2005.

[153] CA Voigt, C Martinez, ZG Wang, SL Mayo, and FH Arnold. Protein building blocks preserved by recombination. *Nat Struct Biol*, 9(7):553, 2002.

[154] SA Wahl, K Nöh, and W Wiechert. ^{13}C labeling experiments at metabolic non-stationary conditions: An exploratory study. *BMC Bioinformatics*, 9(1):152–169, 2008.

[155] D Whitley. A genetic algorithm tutorial. *Stat Comput*, 4(2):65–85, 1994.

[156] W Wiechert and AA de Graaf. Bidirectional reaction steps in metabolic networks: I. Modeling and simulation of carbon isotope labeling experiments. *Biotechnol Bioeng*, 55(1):101–117, 1997.

[157] W Wiechert, M Möllney, N Isermann, M Wurzel, and AA de Graaf. Bidirectional reaction steps in metabolic networks: III. Explicit solution and analysis of isotopomer labeling systems. *Biotechnol Bioeng*, 66(2):69–85, 1999.

[158] W Wiechert and K Nöh. From stationary to instationary metabolic flux analysis. *Adv Biochem Eng/Biotechnol*, 92:145–172, 2005.

[159] C Wittmann, JO Krömer, P Kiefer, T Binz, and E Heinzle. Impact of the cold shock phenomenon on quantification of intracellular metabolites in bacteria. *Anal Biochem*, 327(1):135–139, 2004.

[160] L Yang, T Kasumov, RS Kombu, SH Zhu, AV Cendrowski, F David, VE Anderson, JK Kelleher, and H Brunengraber. Metabolomic and mass isotopomer analysis of liver gluconeogenesis and citric acid cycle: II. Heterogeneity of metabolite labeling pattern. *J Biol Chem*, 283(32):21988–21996, 2008.

[161] H Yoo, MR Antoniewicz, G Stephanopoulos, and JK Kelleher. Quantifying reductive carboxylation flux of glutamine to lipid in a brown adipocyte cell line. *J Biol Chem*, 283(30):20621–20627, 2008.

[162] H Yoo, G Stephanopoulos, and JK Kelleher. Quantifying carbon sources for *de novo* lipogenesis in wild-type and IRS-1 knockout brown adipocytes. *J Lipid Res*, 45(7):1324–1332, 2004.

[163] Y Yoshikuni, TE Ferrin, and JD Keasling. Designed divergent evolution of enzyme function. *Nature*, 440(7087):1078–1082, 2006.

[164] JD Young, JL Walther, MR Antoniewicz, H Yoo, and G Stephanopoulos. An elementary metabolite unit (EMU) based method of isotopically nonstationary flux analysis. *Biotechnol Bioeng*, 99(3):686–699, 2008.

[165] Z Zhao, K Kuijvenhoven, C Ras, WM van Gulik, JJ Heijnen, PJT Verheijen, and WA van Winden. Isotopic non-stationary ^{13}C gluconate tracer method for accurate determination of the pentose phosphate pathway split-ratio in *Penicillium chrysogenum*. *Metab Eng*, 10(3–4):178–186, 2008.

[166] C Zupke and G Stephanopoulos. Intracellular flux analysis in hybridomas using mass balances and *in vitro* ^{13}C NMR. *Biotechnol Bioeng*, 45(4):292–303, 1995.