^{13}C metabolic flux ratio analysis by direct measurement of free metabolic intermediates in *L. mexicana* using gas chromatography-mass spectrometry

Milica Ng[1,2], Eleanor C. Saunders[1,2], Moshe Olshansky[3], Malcolm J. McConville[1,2], and Vladimir A. Likić[2]

[1] Department of Biochemistry and Molecular Biology, University of Melbourne, Parkville, Victoria 3010, Australia

[2] Bio21 Institute of Molecular Science and Biotechnology, University of Melbourne, Parkville, Victoria 3010, Australia

[3] Walter and Eliza Hall Institute of Medical Research, 1 G Royal Parade, Parkville, Victoria 3052, Australia

Abstract

Until recently, ^{13}C-based flux analyses have almost exclusively relied on analysis of labelled amino acids in proteins. This approach is not directly applicable to *Leishmania*, as these parasites scavenge most of their amino acids from the media. *Leishmania* are also unusual in that they i) share little genomic similarity with other organisms ii) constitutively express their metabolic genes and iii) display minimal changes in the enzyme levels throughout their life cycle stages. The three factors have contributed to an early development of comprehensive and reproducible ^{13}C-based metabolomics approaches in these parasites. The work presented here contributes to the creation of new ^{13}C-based metabolic flux approaches based on the isotopologue analysis of free metabolite pools in *Leishmania mexicana*. Namely, a new approach is presented for simultaneous calculation of *in vivo* fractional fluxes (or flux ratios) into two or more metabolite nodes with carbon dioxide condensation, based on isotopologue analysis of free metabolite pools. This method is used to perform the first quantitative *in vivo* fractional flux calculation of central carbon metabolism in any human parasite.

Introduction

Current metabolic research focuses on how to harness the newly acquired omics data in order to be able to i) restore metabolic function when things go wrong (e.g. in numerous metabolic diseases), ii) disable metabolism of pathogens, and iii) optimise metabolic production rates of biotechnologically relevant organisms. Most importantly we want to be able to alter cellular metabolism in a predictable, rather than previously ad hoc, manner. The cellular metabolism comprises a large number of biochemical reactions, and the reconstruction of the metabolic network topology is an important first step in understanding metabolic function in a particular cell. This static picture gives information on the cell's metabolic capacity. What is of more interest however, is the amount of activity throughout different parts of the network. Metabolic flux, defined as the rate of flow of metabolite atoms through the metabolic network, is the direct measure of a metabolic pathway's (or individual reaction's) activity level.

In general, metabolic flux inside the cell is regulated by altering the number and/or the capacity of enzymes that catalyse a particular reaction. Therefore flux increase or decrease is often hypothesized based on the measurements of the relevant enzyme or metabolite levels. This can often lead to incorrect conclusions, as an increase in enzyme levels does not necessarily indicate an increase in the metabolic flux, as enzymes can be present but inactive [1,2]. Alternatively, an enzyme can be active without having any significant control of the overall pathway flux (this scenario is the subject of metabolic control analysis [3]). Similarly, metabolite levels alone are not accurate indicators of enzymatic activity. The tendency of an organism, or a cell, to maintain internal equilibrium by adjusting its physiological processes (homeostasis) often results in metabolite pool levels remaining constant even though a coordinated increase, or decrease, of enzymatic activity throughout a pathway is taking place [4]. Moreover, metabolite levels can change in the opposite direction to the related enzyme's activity, if enzymes upstream or downstream of them have been activated or deactivated. Therefore, data on enzyme and metabolite levels are only able to generate predictions about the rate of flow of metabolite atoms between different nodes. Metabolic flux on the other hand, is the true measure of this rate.

Only a small number of fluxes, such as nutrient uptake, product secretion and biomass production can be measured directly. The intracellular fluxes, which carry the most important information, are not directly measurable. Different mathematical modelling strategies have been developed to solve the problem. According to whether they measure fluxes as a function of time or constant fluxes at metabolic steady state, they can be divided into dynamic and static (or steady-state) modelling approaches respectively. This work is focused on the latter.

All living cells maintain steady state by regulatory mechanisms in order to compensate for changes in their external environments, e.g. temperature, nutrient content or pH level. It has been observed that even after relatively large environmental perturbations, cells reach a new steady state within minutes. Measuring fluxes in metabolic equilibrium, reached after perturbations of interest, provides valuable information on the adjustments that cells make to regulate metabolic function. At metabolic steady state, metabolite levels in the network are constant (or assumed to be near enough constant compared to the relevant flux values). Thus, based on the law of mass conservation, the sum of all the fluxes that add to a particular intracellular metabolite pool must equal the sum of all the fluxes that deplete the same pool. Therefore, utilising known reaction stoichiometry, a set of linear equations can be written for each intracellular metabolite pool inside the metabolic network. The number of equations in the model will be equal to the number of metabolites in the network, and the number of unknowns will be equal to the number of reactions (or flux values to be determined). In a typical metabolic network, metabolites are shared between reactions and hence the number of metabolites will be smaller than the number of reactions. This means that the overall linear equation system will be underdetermined. In other words, there will be more unknowns than the equations to solve for them. Two different strategies have been devised to deal with this: i) flux balance analysis [5,6,7] and ii) ^{13}C-based flux analysis. Our focus is on the latter.

The two most prominent ^{13}C-based approaches are: ^{13}C metabolic flux analysis (^{13}C-MFA) [8,9,10,11], and ^{13}C metabolic flux ratio analysis (^{13}C-MFRA) [12,13,14]. Typically, ^{13}C-labelled substrates are fed to the cells under study, allowing different metabolic pathways to cleave and scramble substrates' carbon backbone. Under the conditions of metabolic and isotopic steady-state, isotopic composition of metabolic intermediates depends on the labelling state of the input substrates, network topology, and fluxes operating in the network. The isotopic composition of intermediates (and products) can be measured by NMR [15,16,17] or mass spectrometry [18,19,20], and given that the input substrates and the network topology are known, fluxes operating in the metabolic network can be inferred by mathematical modelling. However, underlying this is a significant mathematical modelling effort, and a number of challenges remain. Firstly, ^{13}C-MFA requires that the topology of the metabolic network is known, i.e., as a part of mathematical modelling the network topology is assumed. This in turn implies that it is possible to assume incorrect network topology, and obtain a seemingly reasonable solution (fluxes that fit the experimental data), without any warning signs. Secondly, because of the integral nature of the model, any errors in the input data will propagate to affect the entire solution. Once the solution fluxes are obtained by ^{13}C-MFA a detailed statistical evaluation of the results is recommended [21].

^{13}C metabolic flux ratio analysis (^{13}C-MFRA) is a ‘divide and conquer’ variation on ^{13}C-MFA. Namely, the method relies on a set of measurements from ^{13}C-labelling experiments, at isotopic steady state, to resolve relative flux contributions of metabolic pathways and/or reactions to a single metabolic network node of interest. Unlike ^{13}C-MFA, it does not require comprehensive network topology. This comes at the price of being able to compute only relative and not net flux values. However, in organisms where topology is yet unknown this strategy can prove invaluable in deciphering topology before embarking on a more elaborate ^{13}C-MFA work. The widely available detailed drawings of metabolic networks can often leave a novice with a belief that very little, if anything, is unknown due to the evolutionary preservation of metabolic pathways. However, a recent review estimates that 30-40% of metabolic activities are missing enzyme and/or gene allocation in an average organism [22]. The view of enzymes as a single substrate converting catalyst is also challenged by the existence of so called promiscuous enzymes which have the ability to catalyse more than one reaction [23,24]. Therefore, although often unjustifiably neglected, ^{13}C-MFRA offers an important insight into *in vivo* flux activity around individual network nodes. This approach is also essential to reduce the overall computational effort when the biological question of interest can be answered by considering only a few metabolic nodes.

Leishmania cause the human disease leishmaniasis, the clinical symptoms of which range from cutaneous and mucocutaneous forms through to visceral leishmaniasis that is fatal unless treated. It is estimated that 12 million people are currently infected worldwide, and that more than 350 million people are at risk [25,26]. The leishmaniasis is thought to result in 60,000 deaths each year [27], making it the second deadliest protozoan infection after malaria. There is no vaccine against leishmaniasis [28] and current chemotherapies are often costly, cause severe side effects, and demonstrate increasing rates of resistance [29,30,31]. There is an urgent need to identify new drug targets, with metabolic pathways that are necessary for growth, differentiation and virulence of these important human pathogens being likely candidates. These pathways must, at the same time, not present or at least be minimally utilised by their host. As well as targeting these novel parasite-specific pathways, it may be possible to create metabolic ‘bottlenecks’ leading to the accumulation of toxic intermediates. Alternatively, inhibition of an enzyme with slow turnover in a parasite and fast turnover in the host is an option (e.g. [32]). Studying mechanisms by which these pathogens acquire drug resistance is another important area of research.

In microorganisms, ^{13}C-based flux analyses have almost exclusively relied on analysis of labelled amino acids isolated from hydrolysed proteins. In prokaryotes, this approach is favoured as these organisms synthesise many of their amino acids *de novo*. Even though incorporation of ^{13}C atoms into protein takes much longer then into the cellular precursors (i.e. proteinogenic amino acids), the abundance of protein within the cell makes it an easy substrate to isolate and analyse. Furthermore, in ^{13}C-based flux analyses, out of 20 amino acids, 8 metabolite patterns are usually discernible and this is often enough to constrain the stoichiometric model of central carbon metabolism [33]. Furthermore, metabolite extraction from bacteria is still an unsolved problem [33]. These analyses, however, are hampered in organisms that scavenge extracellular amino acids, such as *Leishmania spp*. For example *L. mexicana* have been shown to scavenge most of their amino acids from the media and, with the exception of alanine, glutamate and aspartate, the labelling of other amino acids is negligible [34].

In common with analyses in other eukaryotes, accurately determining flux in *Leishmania spp.* is further complicated by unknown gene functionality and network topology as well as the intracellular compartmentalisation of metabolites and pathways. For example, the GeneDB database indicates that 60% of *L. major* genes currently have no known functionality [35,36]. The knowledge of metabolic gene function is necessary, but not sufficient. The connectivity and arrangement of metabolic pathways need to be experimentally elucidated for each organism of interest. Nonetheless some progress has been made in identifying important carbon sources and delineating the network topology of central carbon metabolism [34]. This was made possible by using a range of ^{13}C-substrates and detecting the ^{13}C-enrichment of metabolic intermediates. Despite these advances, ^{13}C-based flux analysis has not been applied to *Leishmania spp*.

The study of flux throughout central carbon metabolism in *Leishmania spp.* is of particular interest given the pathways' central roles in parasite growth and virulence [37]. This is despite the observation that central carbon metabolism is largely conserved between species as it is anticipated that, given the significant differences in compartmentalisation and regulatory mechanisms between host and parasite, central carbon metabolism remains a good drug target in *Leishmania spp*. ^{13}C-MFRA is attractive means to delineate central carbon metabolism flux in *Leishmania spp*. Firstly, ^{13}C-MFRA requires only the knowledge of reactions which top up a metabolite pool of interest and, as a result, the calculation is often largely independent of the overall network topology. This is particularly useful given methods to experimentally ascertain *Leishmania*'s metabolic network topology are still evolving. Secondly, ^{13}C-MFRA is amenable to high-throughput analysis, and coupled with the advances in GC-MS techniques, can lead to the development of rapid methods capable of scanning a multitude of different experimental set ups within reasonable time frames.

We present here a novel approach for simultaneous calculation of *in vivo* fractional fluxes (or flux ratios) into two or more metabolite nodes with carbon dioxide condensation, based on isotopologue analysis of free metabolite pools. This method highlights the yet unexplored potential of combining ^{13}C-MFRA and direct metabolomics measurements for the quantitative study of metabolism. In particular the approach will benefit the study of i) the organisms whose metabolic network topology is poorly understood, and
ii) the organisms in which a sufficient number of metabolite mass isotopic distributions cannot be inferred from the measurements of proteinogenic amino acids. The utility of the approach is demonstrated by its application on the data collected for the malate and oxaloacetate nodes in central carbon metabolism of rapidly dividing *L. mexicana* parasites, providing the first in vivo quantitative fractional flux values for *Leishmania spp.*, or any other protozoan. These are also, to the best of our knowledge, the first fractional flux results based on the mass isotopic distribution measurements of free intracellular metabolic intermediates.

Methods

Notation. In this section we introduce the notation used throughout. The property that will be of central interest in considerations below is mass isotopic distribution vector **I**. Let **s** denote the vector of a single chemical species that differ in their isotopic composition s = [s_0, s_1, …, s_n] (here chemical species can be either a single atom or a molecule). Let **m** denote the vector of corresponding masses **m** = [m_0, m_1, …, m_n], such that m_0 is the mass of the isotopic species s_0, m_1 is the mass of the isotopic species s_1, and so on. Let **I** denote the isotopic distribution vector **I** = [a_0, a_1, …, a_n] where the numbers a_0, a_1 … a_n represent abundances of individual isotopic species with masses m_0, m_1, …, m_n, such that $a_0 + a_2 + \ldots + a_n = 1$. In the standard convention the vectors **s**, **m** and **I** are ordered according to the increasing mass ($m_0 < m_1 < \ldots < m_n$) such that a_0 is the abundance of the isotopic species s_0 which has the smallest mass m_0. Mass spectrometry (MS) measures *m/z* ratios to produce intensity patterns called mass spectra. The mass spectrum observed experimentally by MS, is directly related to the mass isotopic distribution of the chemical species under observation. In case of singly charged species (z=1) MS measures directly atomic/molecular masses of the species involved. Most mass spectrometers can distinguish masses of 0.05 u apart, however detected masses are commonly binned in integer increments. In metabolomics, all masses detected in the range from −0.3 to +0.7 relative to an integer mass are usually summed and recorded as the integer mass intensity.

There are two distinct ways in which two isotopic populations of atoms can combine into a complex mixture: simple mixing and chemical bonding. In simple mixing of two chemical species with well-defined isotopic populations $\mathbf{I}_1$ and $\mathbf{I}_2$, the mass isotopic distribution of the resulting population is given by:

$$\boldsymbol{I} = f_1\boldsymbol{I}_1 + f_2\boldsymbol{I}_2$$

where f_1 and f_2 are mixing fractions of the two populations.

In mixing due to chemical bonding (i.e. mixing of isotopic species occurs due to chemical reactions) atoms combine in fixed stoichiometric ratios. If the two populations of chemical species are characterized by mass isotopic distributions $\mathbf{I}_1$ and $\mathbf{I}_2$, the resulting chemical species will have the mass isotopic distribution:

$$\boldsymbol{I} = \boldsymbol{I}_1 \otimes \boldsymbol{I}_2$$

where '⊗' denotes discrete convolution (also known as Cauchy product [38]). This simple relationship is a direct result of integer binning of detected masses.

A metabolite node without condensation or cleavage. Consider two metabolic pathways converging to a single metabolite within a larger metabolic network. Figure 1A shows two metabolites, D and E, both converted to G. Let us assume that isotopic populations of D and E are well defined, and given by $\mathbf{I}_D$ and $\mathbf{I}_E$ respectively. If no condensation or cleavage occurs in either of the two reactions D → G and E → G, then G inherits the carbon backbone directly from D or E. If the metabolic network is in metabolic and isotopic steady state, the resulting mass isotopic distribution of G will be a result of simple mixing, given as a combination of the two mass isotopic distributions in proportions determined by relative contributions of fluxes v_D and v_E:

$$\boldsymbol{I}_G = \frac{v_D}{v_D + v_E}\boldsymbol{I}_D + \frac{v_E}{v_D + v_E}\boldsymbol{I}_E$$

or

$$\boldsymbol{I}_G = f_D\boldsymbol{I}_D + f_E\boldsymbol{I}_E$$

where f_D and f_E are fractional contributions of D → G and E → G reactions, respectively. The above equation can be rewritten in a matrix form:

$$\boldsymbol{I}_G^T = \boldsymbol{Z}f^T$$

where $\mathbf{f} = [\, f_D, f_E \,]$ contains mixing fractions, and $f_D + f_E = 1$. The matrix **Z** contains mass isotopic distributions and its columns are mass isotopic distributions for individual chemical species arranged in the same order as the fractional contributions in the vector f (i.e. $\mathbf{Z} = [\mathbf{I}_D^T, \mathbf{I}_E^T]$). If one is able to measure $\mathbf{I}_D$, I_E, and $\mathbf{I}_G$, the fractional contributions of the two reactions f_D and f_E can be obtained by solving the above matrix equation for **f**. In general, **Z** is not a square matrix. However, in the majority of central carbon metabolism reactions which are of interest in the labeling experiments, the number of distinct isotopic populations being mixed is at most 3 (number of **Z** columns), and the smallest number of carbon atoms in their skeletons is at least 3, hence the number of **Z** rows is at least 4. Therefore, the number of rows will be greater than the number of columns. Hence, provided that isotopic distribution vectors of the populations entering the node are linearly independent, **Z** will have a left inverse:

$$\boldsymbol{f}^{\mathrm{T}} = \left(\boldsymbol{Z}^T\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}^T\boldsymbol{I}_G^T \qquad [1]$$

We note that the Equation 1 can be extended to a node where k metabolites converge without condensation or cleavage, with mass isotopic distributions $\mathbf{I}_D$, I_E, …, $\mathbf{I}_M$, and respective fractional contributions vector is $\mathbf{f} = [\, f_D, f_E, ..., f_M \,]$, provided that left inverse of **Z** exists.

A metabolite node with CO_2 condensation. The situation where two or more metabolites converge to a single node without condensation or cleavage is rare. Consider a prototypical model of the TCA cycle with several anaplerotic reactions, as shown in Figure 2. In this schematic, five metabolite nodes can be identified as candidates for flux ratio determination: PEP, Pyr, OAA, Mal, and Suc (on the basis that these metabolites' carbon backbones are produced in more than one reaction, and hence potentially display different isotopic distributions). At each of these metabolite nodes one contributing reaction involves CO_2, either as a cleavage or condensation. More specifically, three metabolite nodes involve CO_2 cleavage (PEP, Pyr, and Suc) and two nodes involve CO_2 condensation (Mal and OAA). The latter reaction, where a N-1 carbon backbone is converted into a N carbon backbone involving CO_2 condensation, will be analysed in more detail below.

Consider a model reaction involving five metabolites D, E, F, G and CO_2, assumed to be a part of a larger metabolic network, and to combine as shown in Figure 1B. Metabolites D and E are converted to G without a loss or gain or carbon atoms, while the metabolite F combines with CO_2 to produce G (therefore, metabolites D, E, and G have N carbon atoms, while metabolite F has N-1 carbon atoms). The fractional contributions to the total influx into the node G are:

$$f_D = \frac{v_D}{v_D + v_E + v_F},\ f_E = \frac{v_E}{v_D + v_E + v_F},\ f_F = \frac{v_F}{v_D + v_E + v_F}$$

the mass isotopic distribution vector for the metabolite G can be written as:

$$\boldsymbol{I}_G = f_D \boldsymbol{I}_D + f_E \boldsymbol{I}_E + \left(1 - f_D - f_E\right)[\boldsymbol{I}_F \otimes \boldsymbol{I}_{CO2}] \qquad [2]$$

where $f_D + f_E \leq 1$, and $\mathbf{I}_F \otimes \mathbf{I}_{CO2}$ represents the discrete convolution of mass isotopic distribution vectors for metabolites F and CO_2.

Since the elements of mass isotopic distribution vectors are normalized to one, we can replace $\mathbf{I}_{CO2}$ = [x, 1-x]. If $\mathbf{I}_D$, $\mathbf{I}_E$, and $\mathbf{I}_F$ are measured, the Equation 2 has three unknowns on the right-hand side: f_E, f_D, and x. The quantity on the left hand side is the isotopic distribution vector $\mathbf{I}_G$ of the metabolite G. Furthermore, we denote the measured value of $\mathbf{I}_G$ as $\mathbf{I}_{Gm}$, and define function Φ_1 such that:

$$\Phi_1 = || \boldsymbol{I}_G - \boldsymbol{I}_{Gm} ||^2 \qquad [3]$$

where double bars denote a vector 2-norm. If $\mathbf{I}_D$, $\mathbf{I}_E$, $\mathbf{I}_F$, and $\mathbf{I}_G$ are measured, Φ_1 is a function of f_E, f_D, and x. Hence, the values for f_E, f_D, and x can be obtained by constrained optimization of the function Φ_1 (f_E, f_D, x):

$$min\ \Phi_1\ \left(f_D, f_E, x\right)\ such\ that\ 0\ \leq f_D, f_E, x\ \leq\ 1\ and\ f_D + f_E \leq\ 1 \quad [4]$$

where Φ_1 is defined by the Equation 3. Typically an optimisation algorithm is employed to solve Equation 4 (e.g. Matlab *FMINCON* function [39]) , and the main challenge then is to ensure that the algorithm has reached a global, rather than a local minimum.

Two metabolite nodes with CO_2 condensation. The extended version of the previous scenario is depicted in Figure 1C. The relationship between metabolites D, E, F, and G is as described previously. We also have a mirror metabolite node S, which is coupled to the node G via CO_2. The metabolites P and Q convert to S without a loss or gain or carbon atoms, while the metabolite R combines with CO_2 to produce G (therefore, metabolites P, Q, and S have N carbon atoms, while metabolite R has N-1 carbon atoms). The mass isotopic distribution vector for the metabolite G is as per Equation 2, and the mass isotopic distribution vector for the metabolite S can be written as:

$$\boldsymbol{I}_S = f_P\boldsymbol{I}_P + f_Q\boldsymbol{I}_Q + \left(1 - f_P - f_Q\right)[\boldsymbol{I}_R \otimes \boldsymbol{I}_{CO2}] \qquad [5]$$

We denote the measured value of $\mathbf{I}_S$ as $\mathbf{I}_{Sm}$, and define function Φ_2 such that:

$$\Phi_2 = || \boldsymbol{I}_S - \boldsymbol{I}_{Sm} ||^2 \qquad [6]$$

If $\mathbf{I}_P$, $\mathbf{I}_Q$, $\mathbf{I}_R$, and $\mathbf{I}_S$ are measured, Φ is a function of f_p, f_Q, and x. And the values for f_P, f_Q, and x can be obtained by constrained optimization of the function Φ_2 (f_A, f_B, x):

$$min\ \Phi_2\ \left(f_P, f_Q, x\right) such\ that\ 0 \leq f_P, f_Q, x \leq 1\ and\ f_P + f_Q \leq 1 \qquad [7]$$

Recognizing that the mass isotopic distribution of CO_2 is a property of the system that arises from mass balances in the metabolic and isotopic steady-state (Figure 2), one can combine Equations [4] and [7] into a single optimization problem:

$$\begin{aligned} &min\ \Phi\ \left(f_D, f_E, f_P, f_Q, x\right) = || \boldsymbol{I} - \boldsymbol{I}_m ||^2 \\ &s.t.\ 0 \leq f_D, f_E, f_P, f_Q, x \leq 1, f_D + f_E \leq 1, and\ f_P + f_Q \leq 1 \end{aligned} \qquad [8]$$

where $\mathbf{I}$ = [$\mathbf{I}_G$, $\mathbf{I}_S$], $\mathbf{I}_m$ = [$\mathbf{I}_{Gm}$, $\mathbf{I}_{Sm}$] (see Equations 2 and 5).

Parasite strains and labelling conditions. *L. mexicana* wild type (WT) (MNYC/BZ/62/M379) promastigotes were cultured in RPMI medium (Sigma) supplemented with 10% heat inactivated foetal bovine serum (iFBS, GibcoBRL) at 27 °C. Parasites were passaged twice weekly (1/100 and 1/1000) into fresh media to maintain log phase growth. Mid-log phase promastigotes (1 x 10^7 cell/ml) were harvested two days after inoculation of the media. Cells were pelleted and resuspended in a pre-labelling media to equilibrate (at 2 x 10^7 cells/mL, 1 hour, 27 °C). Briefly, this completely defined media (modified from [40]) included glucose (6 mM, ^{12}C-U-glc), and most other amino acids except for the non-essential amino acid, alanine. Fatty acids were included as lipid containing bovine serum albumin (BSA, 0.5 % final). After equilibration, cells were pelleted (805 x g, 10 minutes) and resuspended in the labelling media containing 3 mM [U-$^{13}C_6$] glucose and 3 mM [1-$^{13}C_1$] glucose for 15 hr in order to reach isotopic equilibrium (2 x 10^7 cells/mL).

Sampling and metabolite extraction. Parasite metabolism was quenched as described in [41]. Briefly, aliquots of promastigote batch cultures (8 x 10^8

parasites total, approximately 40 mL) were metabolically quenched by immersing the 75 cm^2 flask in a dry ice-ethanol slurry to rapidly cool the cell suspension to 0 ^oC. Immediately, the chilled parasites were centrifuged (805 x g, 10 minutes, 0 ^oC) and the resulting cell pellet was suspended in ice-cold PBS and 16 replicates (4 x 10^7 cell total) were transferred to a microfuge tube and washed twice with cold PBS (10,000 x g, 1 min, at 0 ^oC). The cell pellet was extracted in chloroform:methanol:water ($CHCl_3$:CH_3OH:H_2O, 1:3:1 v/v, 250 µL, vortex mixed), containing 1 nmol scyllo-inositol as internal standard (60 ^oC, 15 min). Insoluble material was removed by centrifugation (16,100 x g, 0 ^oC, 5 minutes) and the supernatant adjusted to $CHCl_3$:CH_3OH:H_2O 1:3:3 v/v, with addition of H_2O, vortex mixed and centrifuged (16,000 x g, 5 min) to induce phase separation. The upper phase of the replicates were transferred to a fresh microfuge tube for analysis of polar metabolites.

Analysis of polar and apolar metabolites by GC-MS. Polar phases were dried into 250 µL glass vial inserts (*in vacuo*, 55 ^oC) and free aldehyde groups protected by derivitization in methoxyamine chloride (Sigma, 20 µl, 20 mg/ml in pyridine) with continuous mixing (14 hrs, 25 ^oC). Replicates were derivitized (silylation) with either BSTFA (N,O-bis(trimethylsilyl) trifluoroacetamide containing 1% trimethylchlorosilane (TMCS), Pierce, 20 µl, 1 hr, 25 ^oC) or MTBSTFA (N-Methyl-N-[tert-butyldimethyl-silyl]trifluoroacetimide containing 1% tert-butyldimethylchlorosilane (TBDMCS), Pierce, 20 µl, 1 hr, 25 ^oC). All samples were analysed by GC-MS with a DB5 capillary column (J&W Scientific, 30 m, 250 µm i.d., 0.25 µm film thickness), with a 10 m inert duraguard and fitted with a gerstel autosampler. The injector insert and GC-MS transfer line temperatures were 270 ^oC and 250 ^oC, respectively. The oven temperature gradient was programmed as follows: 70 ^oC (1 min); 70 ^oC to 295 ^oC (12.5 ^oC/min); 295 ^oC to 320 ^oC (25 ^oC/min); 320 ^oC (2 min). Data was collected in both SCAN and selected ion-monitoring (SIM) modes for metabolite identification and quantification (dwell time 20 ms), respectively. For TMS derivitised samples, the metabolites and ions monitored were: PEP (369.1, 370.1, 371.1, 372.1, and 373.1 *m/z*), citrate (465.15, 466.15, 467.15, 468.15, 469.15, 470.15, 471.15, and 472.15 *m/z*), succinate (247.1, 248.1, 249.1, 250.1, 251.1 and 252.1 *m/z*), and malate (335.1, 336.1, 337.1, 338.1, 339.1, and 340.1 *m/z*). For TBDMS derivitised samples, the metabolites and ions monitored were: alanine (260, 261, 262, and 263 *m/z*), aspartate (418, 419, 420, 421, and 422 *m/z*) and glutamate (432, 433, 434, 435, and 436 *m/z*). For the TBDMS samples, SIM data was collected in split/splitless mode Replicates were randomised and similarly prepared standards were also analysed.

Results

Validation using a model network. A simple model network (shown in Figure 2), comprising of lower glycolysis (v_1 and v_3), TCA cycle (v_4-v_8, v_{12}), glycosomal succinate fermentation (v_9, v_{13}, and v_{15}) pathway [42] and several anaplerotic reactions (v_{10}, v_{11}, v_{14}), was constructed based on previously published data for *L. mexicana* [34,37]. Glucose and aspartate are assumed to be the only carbon sources (alanine was not present in the media). And carbon leaves the network via carbon dioxide, alanine, acetate, glutamate, and succinate. There are a total of 20 fluxes in this reaction network, and assuming a metabolic steady-state it results in nine equations as shown in Table 2. Assuming flux values for the reactions participating in the network and isotopic composition of input substrates it is possible to calculate exact mass isotopic distributions of individual metabolites at metabolic and isotopic steady-state. To validate the method for calculating flux ratios described above, we assumed the reaction fluxes shown in Figure 3, the composition of input substrates, and calculated the steady-state mass isotopic distributions for the carbon backbone of phosphoenolpyruvte (PEP), pyruvate (Pyr), succinate (Suc), malate (Mal), oxaloacetate (OAA), and CO_2 in metabolic and isotopic steady-state (see Table 3). Then we applied the Equation 4 to back-calculate the fractional fluxes around Mal and OAA nodes.

It is straightforward to show that for the assumed fluxes in Figure 3, the steady-state equations hold (Table 2). The substrate labelling was assumed to be 50% [U-$^{13}C_6$] + 50% [1-$^{13}C_1$] glucose and unlabelled aspartate (Glc^{in} and Asp^{in} in Figure 3). Mass spectrometry measurements yield mass isotopic distributions of molecular fragments that can be corrected to obtain mass isotopic distribution vectors of the carbon backbone alone. Therefore for the purpose of this work, the values in Table 3 represent idealized measurements (no experimental error), obtained on the metabolic network shown in Figure 2 and a specific set of fluxes (Figure 3).

For Mal node, Equations 3 and 4 can be written explicitly as follows:

$$\boldsymbol{I}_{Mal} = f_8\boldsymbol{I}_{Suc} + f_{13}\boldsymbol{I}_{OAA} + \left(1 - f_8 - f_{13}\right)[\boldsymbol{I}_{Pyr} \otimes \boldsymbol{I}_{CO2}] \qquad [9]$$

$$min\ \Phi_{Mal}\ \left(f_8, f_{13}, x\right)\ such\ that\ 0\ \le f_8, f_{13}, x\ \le\ 1\ and\ f_8 + f_{13} \le\ 1 \qquad [10]$$

where f_8, f_{11}, and f_{13} are fractional fluxes corresponding to net fluxes v_8, v_{11}, and v_{13} respectively (see Figure 2), and $f_{11} = 1 - f_8 - f_{13}$. Similarly, for OAA node Equations 3 and 4 can be written as:

$$\boldsymbol{I}_{OAA} = f_2\boldsymbol{I}_{Asp^{in}} + \left(1 - f_2 - f_{15}\right)\boldsymbol{I}_{Mal} + f_{15}\,[\boldsymbol{I}_{PEP} \otimes \boldsymbol{I}_{CO2}] \qquad [11]$$

$$min\ \Phi_{OAA}\ \left(f_2, f_{15}, x\right)\ such\ that\ 0\ \le f_2, f_{15}, x\ \le\ 1\ and\ f_2 + f_{15} \le\ 1 \qquad [12]$$

where f_2, f_{12}, and f_{15} are fractional fluxes corresponding to v_2, v_{12}, and v_{15}, respectively (Figure 2), and $f_{12} = 1 - f_2 - f_{15}$. In the above equations, the variable x refers to m_0 in the mass isotopic distribution vector of CO_2 (i.e. $\mathbf{I}_{CO2}$ = [x, 1-x], see above). As before, recognizing that the mass isotopic distribution of CO_2 is a property of the system that arises from mass balances in the metabolic and isotopic steady-state, one can combine Equations [10] and [12] into a single optimization problem:

$$\begin{aligned} &min\ \Phi\ \left(f_2, f_8, f_{13}, f_{15}, x\right) = \|\boldsymbol{I}\ - \boldsymbol{I}_m\|^2 \\ &s.t.\ 0\ \le f_2, f_8, f_{13}, f_{15}, x\ \le\ 1, f_8 + f_{13}\ \le\ 1, and\ f_2 + f_{15} \le\ 1 \end{aligned} \qquad [13]$$

where I = [I_{Mal} , I_{OAA}] is the concatenation of calculated mass isotopic distribution vectors for Mal and OAA respectively, Equations 9 and 11 or explicitly:

$$
\begin{aligned}
I_{Mal,0} &= f_{13}I_{OAA,0} + f_8 I_{Suc,0} - I_{Pyr,0}x(f_{13} + f_8 - 1)\\
I_{Mal,1} &= f_{13}I_{OAA,1} + f_8 I_{Suc,1} + (I_{Pyr,0}(x-1) - xI_{Pyr,1})(f_{13} + f_8 - 1)\\
I_{Mal,2} &= f_{13}I_{OAA,2} + f_8 I_{Suc,2} + (I_{Pyr,1}(x-1) - xI_{Pyr,2})(f_{13} + f_8 - 1)\\
I_{Mal,3} &= f_{13}I_{OAA,3} + f_8 I_{Suc,3} + (I_{Pyr,2}(x-1) - xI_{Pyr,3})(f_{13} + f_8 - 1)\\
I_{Mal,4} &= f_{13}I_{OAA,4} + f_8 I_{Suc,4} + I_{Pyr,3}(x-1)(f_{13} + f_8 - 1)\\
I_{OAA,0} &= f_{12}I_{Mal,0} + f_2 I_{Asp,0} - I_{PEP,0}x(f_{12} + f_2 - 1)\\
I_{OAA,1} &= f_{12}I_{Mal,1} + f_2 I_{Asp,1} + (I_{PEP,0}(x-1) - xI_{PEP,1})(f_{12} + f_2 - 1)\\
I_{OAA,2} &= f_{12}I_{Mal,2} + f_2 I_{Asp,2} + (I_{PEP,1}(x-1) - xI_{PEP,2})(f_{12} + f_2 - 1)\\
I_{OAA,3} &= f_{12}I_{Mal,3} + f_2 I_{Asp,3} + (I_{PEP,2}(x-1) - xI_{PEP,3})(f_{12} + f_2 - 1)\\
I_{OAA,4} &= f_{12}I_{Mal,4} + f_2 I_{Asp,4} + I_{PEP,3}(x-1)(f_{12} + f_2 - 1)
\end{aligned}
$$

and I_m is the concatenation of the same measured vectors.

This formulation of the problem allows a single optimization of the objective function of five variables, subject to upper and lower boundary constraints as well as two independent inequality constraints, to obtain fractional distribution of incoming fluxes around Mal and OAA nodes. Since the proposed approach delineates flux distributions around two nodes simultaneously, it falls between the traditional ^{13}C MFA (resolves all fluxes in the underlying network topology) and metabolic flux ratio analysis (resolves distribution of fluxes around a single node).

Stability of the solution. We investigated stability of the optimization described by Equation 13, by repeated optimizations on idealized data. The fluxes were assumed as shown in Figure 3, and an equimolar mix of [U-$^{13}C_6$] and [1-$^{13}C_1$] glucose and unlabelled aspartate were assumed as substrates. Isotopic steady-state labelling of PEP, Pyr, Suc, Mal, and OAA were calculated by EMU methodology [43]. The calculated values of mass isotopic vectors were confirmed by forward labelling calculations carried in OpenFlux [44] and Metran [45]. We employed 100 randomly initialised runs of Matlab *FMINCON* optimiser to compute the fractional fluxes for the malate and oxaloacetate nodes. The results (Figure 4) showed that the sum of squared residuals (SSR) value returned was different each time. This means that either objective function surface is very flat causing the optimisation algorithm to terminate in a vicinity of the global minimum (but rarely, if ever, at the global minimum) or the surface of the objective function had numerous closely spaced shallow local minima in the vicinity of the global minimum. Moreover, the existence of the global minimum is not guaranteed. Further tests with 1000, 10000, 100000 and 1000000 randomly initialised runs of *FMINCON* optimiser were conducted, and although each time the results got sufficiently close to the expected (synthetic) values, SSR value distribution was similar to that of 100 runs (Figure 4). In the case of experimental data (where the results are not known) it would have been very difficult to ascertain how many randomly initialised optimisation runs were sufficient. Decreasing the value of the objective function stopping criteria and increasing the number of iterations per optimisation run was also tested with unsatisfactory results. Noting that if x was known both fractional fluxes could be solved for analytically via Equation 1, an alternative approach is to vary x from 0 to 1, compute fractional fluxes and the objective function Φ, and then plot the value of Φ against the value of x. This will reveal the existence and the location of the global minimum, and the corresponding values. Firstly, we utilised quadratic programming to calculate the objective function value for a range of, for example, 100,000 equidistant x values. The results are shown in Figure 5A, and they indicate that the objective function has a global minimum. The minimum value of this new range is shown in Figure 5B, and has been observed to produce more precise results than 1, 000, 000 runs of *FMINCON* optimiser in the previous section (Table 4). Moreover, on a computer where 1, 000, 000 runs of *FMINCON* optimiser takes 10 hours, this approach runs in less than 10 minutes and is therefore suitable for high-throughput applications.

Validation in the presence of noise. Experimental data will always be affected by noise, various artifacts due to machine imperfections, and possibly human error. Therefore an important question is, are the fractional fluxes computable in the presence of realistic levels of experimental noise? In instruments that deploy detectors with electron multipliers, a proportional relationship between the measured mean intensity and variance is often observed (e.g. in mass spectrometry [46] and microarrays [47]. In our data (five metabolites, a total of 23 *m/z* channels measured), a strong relationship between the mean intensity and variance was not apparent (Figure 6). The observed correlation coefficient between the mean intensity and variance was 0.448, and the p-value for testing the hypothesis of no correlation was 0.03. Furthermore, the data for each *m/z* channel appeared normally distributed in the first approximation: for only one out of 23 *m/z* channels the hypothesis that the observed values are drawn from a normal distribution could be rejected at the 95% confidence level by Lilliefors test (specifically for Mal m_0+2 data set, which showed two groups of tightly grouped values). However, the distribution of variances calculated for 23 *m/z* channels (Figure 6, y-axis) failed Kolmogorov-Smirnov and Lilliefors tests for normality. Therefore, the observed distribution of variances across different *m/z* channels is unlikely to be normal. To approximate this distribution, we fitted the parameters of the Weibull distribution to the calculated variances across 23 *m/z* channels. The Weibull distribution is widely used in reliability engineering and life data analysis because it can model a wide range of behaviours [48]. Best fit to the experimental data resulted in α=0.0031 (scale) and β=1.5315 (shape) of the Weibull distribution. Therefore to simulate noise similar to that observed experimentally, we assumed the following: (a) no correlation between the noise variance and m/z channel intensity; (b) noise for each *m/z* channel was normally distributed; and (c) the variances of individual *m/z* channels follow the Weibull distribution with α=0.0031 and β=1.5315.

In summary, the data with noise was simulated in the following way. Isotopic steady-state labelling of PEP, Pyr, Suc, Mal, and OAA were calculated by EMU methodology [43], assuming the fluxes as given in Figure 3, and an equimolar mix of [U-$^{13}C_6$] and [1-$^{13}C_1$] glucose and unlabelled aspartate as substrates. This produced backbone labelling patterns of PEP, Pyr, Suc, Mal, and OAA, as idealized intensities for 23 *m/z* channels (m_0, m_0+1, m_0+2, m_0+3 for three carbon backbones of Pyr and PEP, and m_0, m_0+1, m_0+2, m_0+3, m_0+4 for four carbon backbones of Suc, Mal, and OAA). To simulate noise, a set of variances was drawn from the Weibull distribution with α=0.0031 and β=1.5315, and variances were randomly assigned to 23 *m/z* channels. In the next step, for each *m/z* channel a noise value was drawn from a normal distribution characterized with the assigned variance, and the noise was added to the channel intensity. The plot of mean intensity vs standard deviation for the experimental data, as well as for an example simulated data with noise, is shown in Figure 6. The fractional flux estimation method (Equation 13) was run with the noisy synthetic data, and it was able to retrieve the original fractional fluxes to within the confidence intervals (Table 5). This validates the approach in the presence of the level of noise found in our experimental data.

Application to experimental data. We applied the optimizaition via Equation 13 to determine fractional fluxes around Mal and OAA nodes in *L. mexicana*. Rapidly dividing *L. mexicana* parasites were cultured in completely defined media containing equimolar [U-$^{13}C_6$] and [1-$^{13}C_1$] glucose for 15 hrs. Metabolites were extracted, derivitised and analysed by GC-MS. The normalized data obtained for PEP, Pyr, Suc, Mal, and OAA, corrected for the presence of natural isotopes and heteroatoms [49] is shown in Table 6. For these five metabolites a total of 23 *m/z* channels were measured (m_0, m_0+1, m_0+2, m_0+3 for three carbon backbones of Pyr and PEP; and m_0, m_0+1, m_0+2, m_0+3, m_0+4 for four carbon backbones of Suc, Mal, and OAA). Pyr labelling was measured via free intracellular alanine, and OAA labelling via intracellular aspartate. A total of 8 experimental replicates were measured. The experimental data acquired on *L. mexicana* is shown in Table 6. We assumed the network topology shown in Figure 2, and proceeded to determine fractional fluxes around Mal and OAA nodes according to Equation 13. Table 7 lists best solutions for f_2, f_8, f_{13}, and f_{15}, for each experimental replicate achieved by minimization of the objective function. A simple measure of the uncertainty in the solution is given as standard deviation in Table 7, calculated from eight experimental replicates. The estimated fractional fluxes for Mal node (f_8, f_{11}, and f_{13}) and OAA node (f_2, f_{12}, and f_{15}) are shown graphically in Figure 7, and a pictorial representation of the fractional fluxes in *L. mexicana* determined from the experimental data is shown in Figure 8. We observed that OAA is chiefly derived from malate (f_{12}, 86%) rather then from aspartate (f_2, 4%) or PEP (f_{15}, 10%). Malate was primarily derived from succinate (f8, 61%) and OAA (f_{13}, 38%) while pyruvate contribution was found to be negligible (f_{11}, 1%).

Discussion

A total of six flux ratios into the oxaloacetate and malate pools, as well as the mass isotopic distribution of carbon dioxide, have been computed in the rapidly dividing *L. mexicana* promastigotes at metabolic and isotopic steady state (Table 7). These results, shown in Figures 7 and 8, represent the first *in vivo* quantitative fractional flux values for *Leishmania spp.* or any other protozoan. They are also, to the best of our knowledge, the first fractional flux results based on the mass isotopic distribution measurements of free intracellular metabolic intermediates. Together these two firsts highlight the yet unexplored potential of combining ^{13}C-MFRA and direct metabolomics measurements for the quantitative study of metabolism. The newly developed approach for simultaneous calculation of *in vivo* fractional fluxes into two or more nodes with carbon dioxide condensation falls in between traditional ^{13}C-MFA (computes a number of net fluxes in the network) and ^{13}C-MFRA (computes fractional fluxes around a single node). The metabolic reactions where carbon dioxide from the environment is incorporated into a metabolite's carbon skeleton are commonly encountered in biology. Specifically, the new approach will be essential in experimental set-ups where the carbon dioxide mass isotopic distribution cannot be accurately and precisely quantified (e.g. bioreactors operated in batch mode).

The fractional flux results show that the oxaloacetate pool is predominantly topped up from the malate pool (~86% of the oxaloacetate carbon backbone comes from the malate carbon skeleton), a small amount (~10%) of the oxaloacetate is generated from PEP (produced by glycolysis) through the glycosomal succinate fermentation (GSF) pathway, and a very small amount of the oxaloacetate (~4%) comes from the uptake of exogenous aspartate (Figure 8).

Fractional flux results for the malate node indicate that the malic enzyme operation in the direction from pyruvate to malate is negligible, and almost twice as much malate is created through the TCA cycle from succinate than via the GSF pathway from oxaloacetate. This result confirms the irreversibility of the malic enzyme reaction reported previously [34]. Fractional influxes into the malate pool display considerably larger confidence intervals than fractional influxes into the oxaloacetate pool (Figure 7). Analysis of additional metabolites, such as the mass isotopic distribution of fumarate which is closer to malate than the measured succinate, may improve the fit for the malate node. Alternatively, the larger confidence may actually arise from metabolite compartmentation in which the mass isotopic distributions of malate may differ according to its subcellular localisation.

In *Leishmania spp.* malate serves as both an intermediate of the TCA cycle, within the mitochondria, and as an intermediate in PEP fermentation to succinate within the glycosome. Indeed several intermediates (fumarate, OAA, and succinate) are located in both the glycosome and mitochondria. Indeed OAA and malate may also be found in the cytosol. This compartmentalisation adds another level of complexity to the analysis of eukaryotic cells that is not found in prokaryotes. Unfortunately, currently available techniques are unable to measure compartment-specific metabolite labelling patterns. Commonly, the existence of compartments is ignored in ^{13}C-based metabolic flux modelling by assuming uniform labelling patterns of individual metabolites inside multiple compartments. No studies currently exist to prove or disprove this assumption [50]. Indeed, the analyses presented in this work have only been made feasible with the recent development of the novel experimental methods for the detection of labelling patterns inside free intracellular metabolites. It is expected that within the next decade, further advances in the experimental and measurement techniques will allow compartment specific metabolite measurements, and this level new resolution will make ^{13}C-based metabolic approaches even more powerful.

Due to potential reversibility of all reactions involved in topping up the OAA and malate nodes, no definitive conclusions can be drawn about the net flux values through the TCA cycle and GSF pathway. However in the future, if enough metabolic network nodes are analysed and the extracellular fluxes are measured, the resulting information can be assembled to compute net fluxes. Moreover, future studies into enzyme reversibility using approaches such as those recently developed by Kleijn and co-workers [51], could be incorporated into this model. Nevertheless, the results for both nodes point to relatively high activity in the TCA cycle, which is also supported by the high production of carbon dioxide reported earlier [34]. Therefore, while some of the relative flux differences can be explained by reaction reversibility, the flux through the TCA cycle might still be higher than that through the GSF pathway. Hence, while the flux through the TCA cycle is not expected to be eight times the value of the GSF flux, as this would interfere with the energetics inside the glycosome, it could still be relatively higher. Its currently believed that approximately half of the PEP produced by glycolysis is imported back into the glycosome for fermentation to succinate in order to regenerate the investment of ATP and NADPH made in the preparatory stages of glycolysis. This would severely restrict the availability of PEP for import into the TCA cycle as OAA (f_{15}) [34].

In summary, methodologies developed in this work contribute to the rapid estimation of *in vivo* fractional fluxes (or flux ratios) in *L. mexicana* parasites, based on GC-MS ^{13}C-based metabolomics measurements of free metabolic intermediates. Ultimately, ^{13}C-based flux modelling provides a high resolution on key areas of metabolism, and the approach presented here and its further development will result in a comprehensive tool for the elucidation of molecular pathogenicity, and precise measurements of the parasite's metabolic response to different perturbations of interest, such as genetic modifications, nutritional and other environmental changes, including, the effect of the existing and new candidate drugs, and in different life cycle stages. It is hoped that due to the generality of the problem formulation, the approach will also find a wider application to ^{13}C-MFRA studies in various other organisms.

Acknowledgments

The authors would like to acknowledge Metabolomics Australia and David de Souza for the use of their GC-MS in the collection of the *Leishmania* data sets.

References

1. Bouwman J, Kiewiet J, Lindenbergh A, van Eunen K, Siderius M, et al. (2010) Metabolic regulation rather than de novo enzyme synthesis dominates the osmo-adaptation of yeast. Yeast 28: 43-53.
2. Suarez RK, Moyes CD (2012) Metabolism in the age of 'omes'. J Exp Biol 215: 2351-2357.
3. Fell DA (1992) Metabolic control analysis: a survey of its theoretical and experimental development. Biochem J 286 (Pt 2): 313-330.
4. Fernie AR, Geigenberger P, Stitt M (2005) Flux an important, but neglected, component of functional genomics. Curr Opin Plant Biol 8: 174-182.
5. Kauffman KJ, Prakash P, Edwards JS (2003) Advances in flux balance analysis. Curr Opin Biotechnol 14: 491-496.
6. Palsson B (2004) Two-dimensional annotation of genomes. Nat Biotechnol 22: 1218-1219.
7. Price ND, Reed JL, Palsson BO (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. Nat Rev Microbiol 2: 886-897.
8. Zamboni N, Sauer U (2009) Novel biological insights through metabolomics and 13C-flux analysis. Curr Opin Microbiol 12: 553-558.
9. Wiechert W (2001) 13C metabolic flux analysis. Metab Eng 3: 195-206.
10. Stephanopoulos G (1999) Metabolic fluxes and metabolic engineering. Metab Eng 1: 1-11.
11. Tang YJ, Martin HG, Myers S, Rodriguez S, Baidoo EE, et al. (2009) Advances in analysis of microbial metabolic fluxes via (13)C isotopic labeling. Mass Spectrom Rev 28: 362-375.
12. Sauer U (2004) High-throughput phenomics: experimental methods for mapping fluxomes. Curr Opin Biotechnol 15: 58-63.
13. Nanchen A, Fuhrer T, Sauer U (2007) Determination of metabolic flux ratios from 13C-experiments and gas chromatography-mass spectrometry data: protocol and principles. Methods Mol Biol 358: 177-197.
14. Zamboni N, Fendt SM, Ruhl M, Sauer U (2009) (13)C-based metabolic flux analysis. Nat Protoc 4: 878-892.
15. Szyperski T (1998) 13C-NMR, MS and metabolic flux balancing in biotechnology research. Q Rev Biophys 31: 41-106.
16. Sherry AD, Jeffrey FM, Malloy CR (2004) Analytical solutions for (13)C isotopomer analysis of complex metabolic conditions: substrate oxidation, multiple pyruvate cycles, and gluconeogenesis. Metab Eng 6: 12-24.
17. Fan TW, Lane AN (2011) NMR-based stable isotope resolved metabolomics in systems biochemistry. J Biomol NMR 49: 267-280.

18. Christensen B, Nielsen J (1999) Isotopomer analysis using GC-MS. Metab Eng 1: 282-290.
19. Hellerstein MK, Neese RA (1999) Mass isotopomer distribution analysis at eight years: theoretical, analytic, and experimental considerations. Am J Physiol 276: E1146-1170.
20. Wittmann C (2002) Metabolic flux analysis using mass spectrometry. Adv Biochem Eng Biotechnol 74: 39-64.
21. Antoniewicz MR, Kelleher JK, Stephanopoulos G (2006) Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. Metab Eng 8: 324-337.
22. Chen L, Vitkup D (2007) Distribution of orphan metabolic activities. Trends Biotechnol 25: 343-348.
23. Hult K, Berglund P (2007) Enzyme promiscuity: mechanism and applications. Trends Biotechnol 25: 231-238.
24. Nam H, E LN, A LJ, H LD, L CR, et al. (2012) Network Context and Selection in the Evolution to Enzyme Specificity. Science 337: 1101-1104.
25. Cox FEG (1996) The Wellcome Trust illustrated history of tropical diseases. London: The Wellcome Trust. 452 p.
26. Hotez PJ, Fenwick A, Savioli L, Molyneux DH (2009) Rescuing the bottom billion through control of neglected tropical diseases. Lancet 373: 1570-1575.
27. Neuber H (2008) Leishmaniasis. J Dtsch Dermatol Ges 6: 754-765.
28. Nagill R, Kaur S (2011) Vaccine candidates for leishmaniasis: A review. Int Immunopharmacol.
29. Ait-Oudhia K, Gazanion E, Vergnes B, Oury B, Sereno D (2011) Leishmania antimony resistance: what we know what we can learn from the field. Parasitol Res 109: 1225-1232.
30. Murray HW, Berman JD, Davies CR, Saravia NG (2005) Advances in leishmaniasis. Lancet 366: 1561-1577.
31. Ponte-Sucre A, Faber JH, Gulder T, Kajahn I, Pedersen SE, et al. (2007) Activities of naphthylisoquinoline alkaloids and synthetic analogs against Leishmania major. Antimicrob Agents Chemother 51: 188-194.
32. Breton JC, Bouteille B, Sonan T (1988) [DFMO: an alternative therapeutic agent for human African trypanosomiasis]. Bull Soc Pathol Exot Filiales 81: 571-577.
33. Bolten CJ, Kiefer P, Letisse F, Portais JC, Wittmann C (2007) Sampling for metabolome analysis of microorganisms. Anal Chem 79: 3843-3849.
34. Saunders EC, Ng WW, Chambers JM, Ng M, Naderer T, et al. (2011) Isotopomer profiling of Leishmania mexicana promastigotes reveals important roles for succinate fermentation and aspartate uptake in tricarboxylic acid cycle (TCA) anaplerosis, glutamate synthesis, and growth. J Biol Chem 286: 27706-27717.
35. Choi J, El-Sayed NM (2012) Functional genomics of trypanosomatids. Parasite Immunol 34: 72-79.
36. Scheltema RA, Decuypere S, T'Kindt R, Dujardin JC, Coombs GH, et al. (2010) The potential of metabolomics for Leishmania research in the post-genomics era. Parasitology 137: 1291-1302.

37. Saunders EC, DP DES, Naderer T, Sernee MF, Ralton JE, et al. (2010) Central carbon metabolism of Leishmania parasites. Parasitology 137: 1303-1313.
38. Kreyszig E (1988) Advanced Engineering Mathematics. New York: John Wiley & Sons, Inc. 819 p.
39. Fischer E, Sauer U (2003) Metabolic flux profiling of Escherichia coli mutants in central carbon metabolism using GC-MS. Eur J Biochem 270: 880-891.
40. Merlen T, Sereno D, Brajon N, Rostand F, Lemesre JL (1999) Leishmania spp: completely defined medium without serum and macromolecules (CDM/LP) for the continuous in vitro cultivation of infective promastigote forms. Am J Trop Med Hyg 60: 41-50.
41. Saunders EC, Macrae JI, Naderer T, Ng M, McConville MJ, et al. (2012) LeishCyc: A Guide to Building a Metabolic Pathway Database and Visualization of Metabolomic Data. Methods Mol Biol 881: 505-529.
42. Michels PA, Bringaud F, Herman M, Hannaert V (2006) Metabolic functions of glycosomes in trypanosomatids. Biochim Biophys Acta 1763: 1463-1477.
43. Antoniewicz MR, Kelleher JK, Stephanopoulos G (2007) Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. Metab Eng 9: 68-86.
44. Quek LE, Wittmann C, Nielsen LK, Kromer JO (2009) OpenFLUX: efficient modelling software for 13C-based metabolic flux analysis. Microb Cell Fact 8: 25.
45. Yoo H, Antoniewicz MR, Stephanopoulos G, Kelleher JK (2008) Quantifying reductive carboxylation flux of glutamine to lipid in a brown adipocyte cell line. J Biol Chem 283: 20621-20627.
46. House LL, Clyde MA, Wolpert RL (2011) Bayesian Nonparametric Models for Peak Identificationin (MALDI-TOF) Mass Spectroscopy. Annals of Applied Statistics 5: 1488-1511.
47. Novak JP, Sladek R, Hudson TJ (2002) Characterization of variability in large-scale gene expression data: implications for study design. Genomics 79: 104-113.
48. Khan MSA, Khalique A, Abouammoh AM (1989) On estimating parameters in a discrete Weibull distribution. Reliability, IEEE Transactions on 38: 348-350.
49. Lee WN, Byerley LO, Bergner EA, Edmond J (1991) Mass isotopomer analysis: theoretical and practical considerations. Biol Mass Spectrom 20: 451-458.
50. Wahrheit J, Nicolae A, Heinzle E (2011) Eukaryotic metabolism: measuring compartment fluxes. Biotechnol J 6: 1071-1085.
51. Kleijn RJ, Buescher JM, Le Chat L, Jules M, Aymerich S, et al. (2010) Metabolic fluxes during strong carbon catabolite repression by malate in Bacillus subtilis. J Biol Chem 285: 1587-1596.

Figures

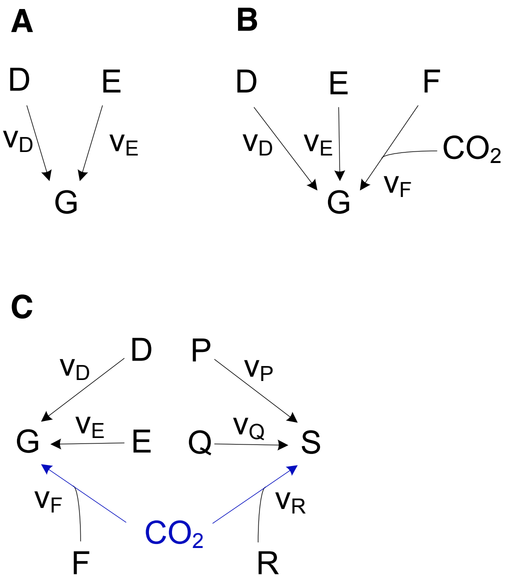

Figure 1. Three scenarios for converging metabolites.

Panel (A): metabolites D and E are converted to metabolite G without condensation or cleavage. In this case G inherits the carbon backbone directly from both D and E. Panel (B): three metabolites D, E, and F contribute carbon backbones to metabolite G. In this case metabolites D and E carbon backbones are form metabolite's G backbone without condensation or cleavage, while the metabolite F's backbone combines with the carbon in CO_2 to form the carbon skeleton of G. Panel (C): similar to (B) except that there are now two nodes coupled via CO_2.

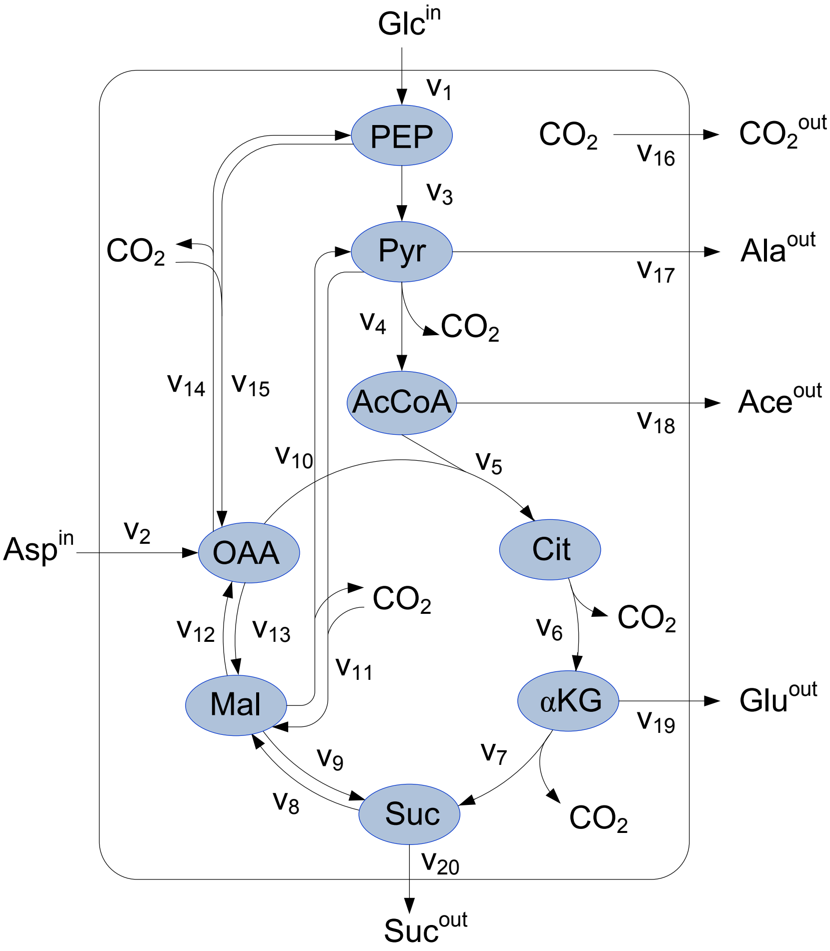
Glc^in
v1
PEP
CO2
v16
CO2^out
v3
CO2
Pyr
Ala^out
v17
v4
CO2
v14
v15
AcCoA
Ace^out
v18
v10
v5
Asp^in
v2
OAA
Cit
CO2
CO2
v12
v13
v6
v11
Mal
αKG
Glu^out
v19
v9
v7
v8
Suc
CO2
v20
Suc^out

Figure 2. The model reaction network.

The network includes the prototypical TCA cycle, lower glycolysis, and several anaplerotic reactions. It is based on central carbon metabolism network topology published by Saunders et al., in 2011 [34]. The solid square represents the system boundary; Glc^{in} and Asp^{in} represent glucose and aspartate, assumed to be only substrates; Suc^{out}, Glu^{out}, Ace^{out}, Ala^{out}, and CO_2^{out} represent secreted succinate, glutamate, acetate, alanine, and carbon dioxide, respectively; PEP, Pyr, AcCoA, Cit, αKG, Suc, Mal and OAA represent intracellular phosphoenolpyruvate, pyruvate, acetyl coenzyme A, citrate, α-ketoglutarate, succinate, malate and oxaloacetate, respectively. The reactions are as given in Table 1.

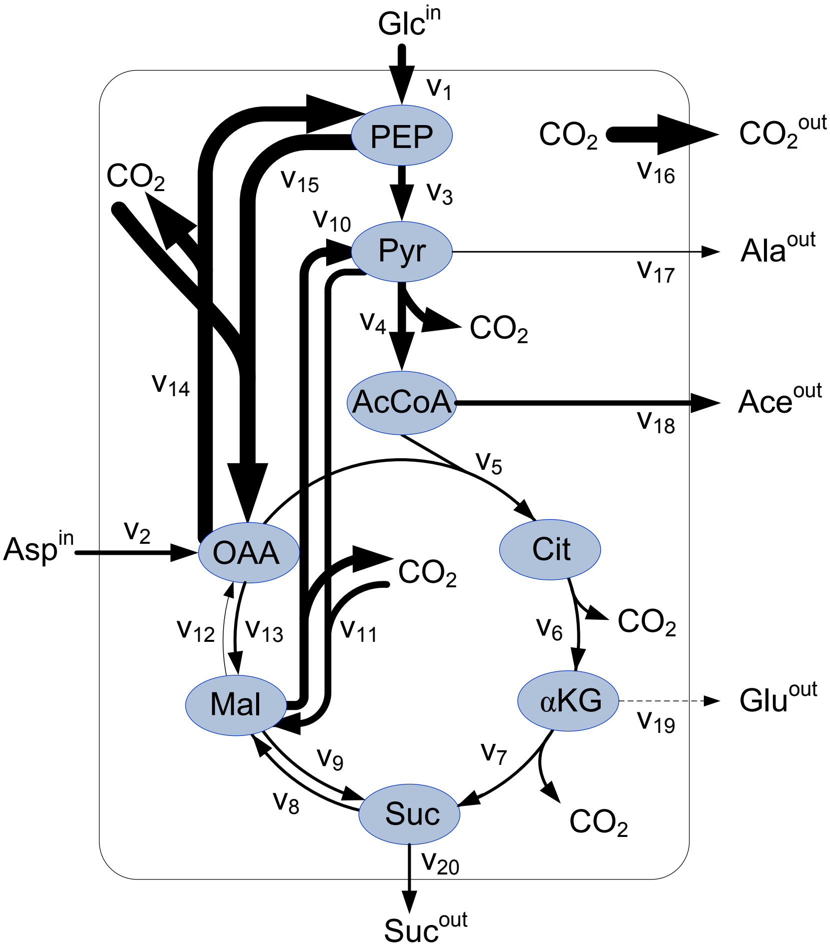
Glc^in
v1
PEP
CO2
CO2^out
v16
CO2
v15
v3
v10
Pyr
Ala^out
v17
v4
CO2
v14
AcCoA
Ace^out
v18
v5
Cit
Asp^in
v2
OAA
CO2
v12
v13
v11
v6
CO2
Mal
αKG
Glu^out
v19
v9
v7
v8
Suc
CO2
v20
Suc^out

Figure 3. Graphical representation of synthetic fluxes assumed to operate in the reaction network in Figure 2.

The assumed fluxes are as follows: v_1 = 1.1, v_2 = 0.6, v_3 = 1.0, v_4 = 1.1, v_5 = 0.4, v_6 = 0.4, v_7 = 0.4, v_8 = 0.4, v_9 = 0.4, v_{10} = 1.2, v_{11} = 0.9, v_{12} = 0.1, v_{13} = 0.4, v_{14} = 2.0, v_{15} = 2.1, v_{16} = 2.1, v_{17} = 0.2, v_{18} = 0.7, v_{19} = 0, v_{20} = 0.4. The zero flux (i.e. v_{19}) is shown as a dashed line; the thickness of the lines representing the rest of the fluxes is directly proportional to their value.

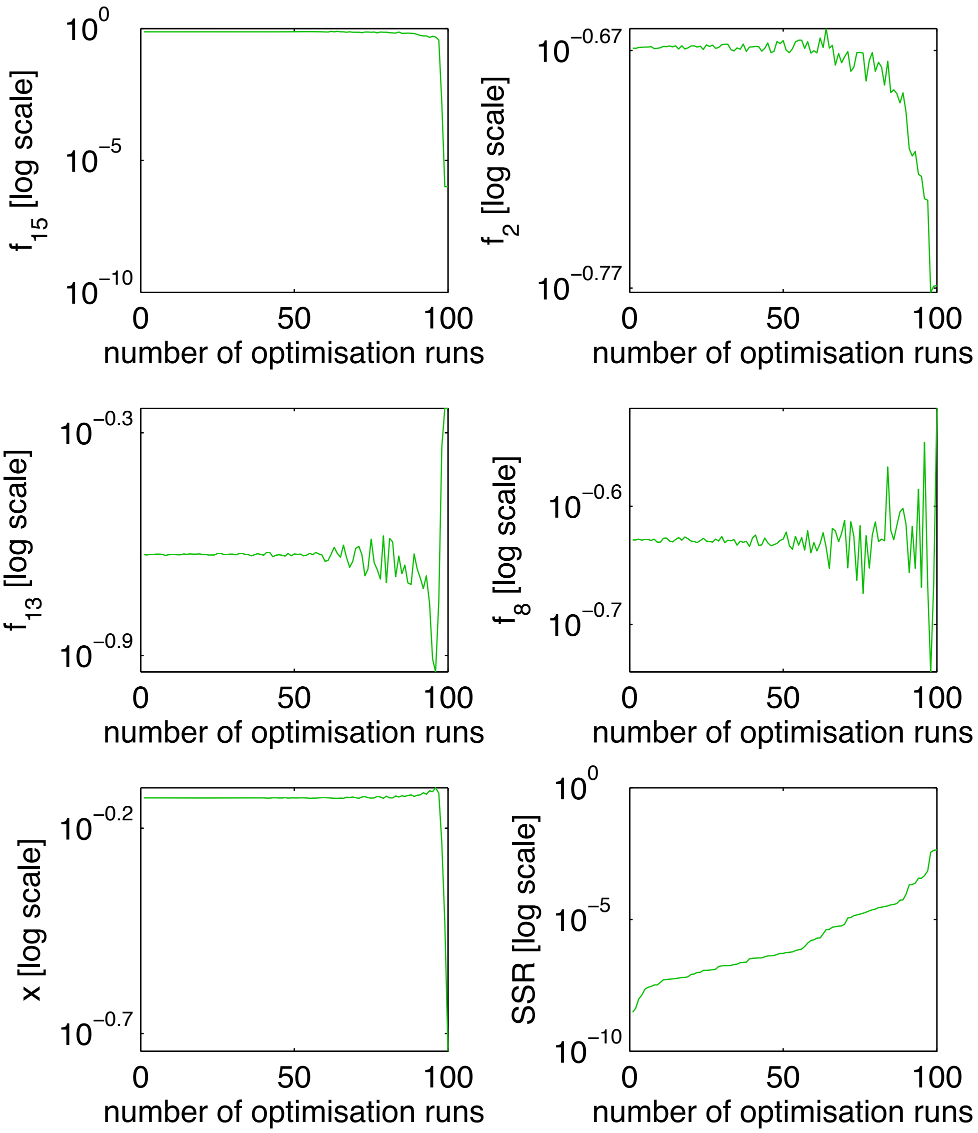

Figure 4. Results based on ideal synthetic data and 100 runs of Matlab's *FMINCON* optimisation function.

Optimisation run numbers are based on the results sorted by the value of the objective function (i.e. 1st optimisation run has the smallest objective function value, and 100th optimisation run has the largest objective function value. Similar results were obtained for 1000, 10000, 100000 and 1000000 optimisation runs, in that no single value for the objective function is reached in more than one run.

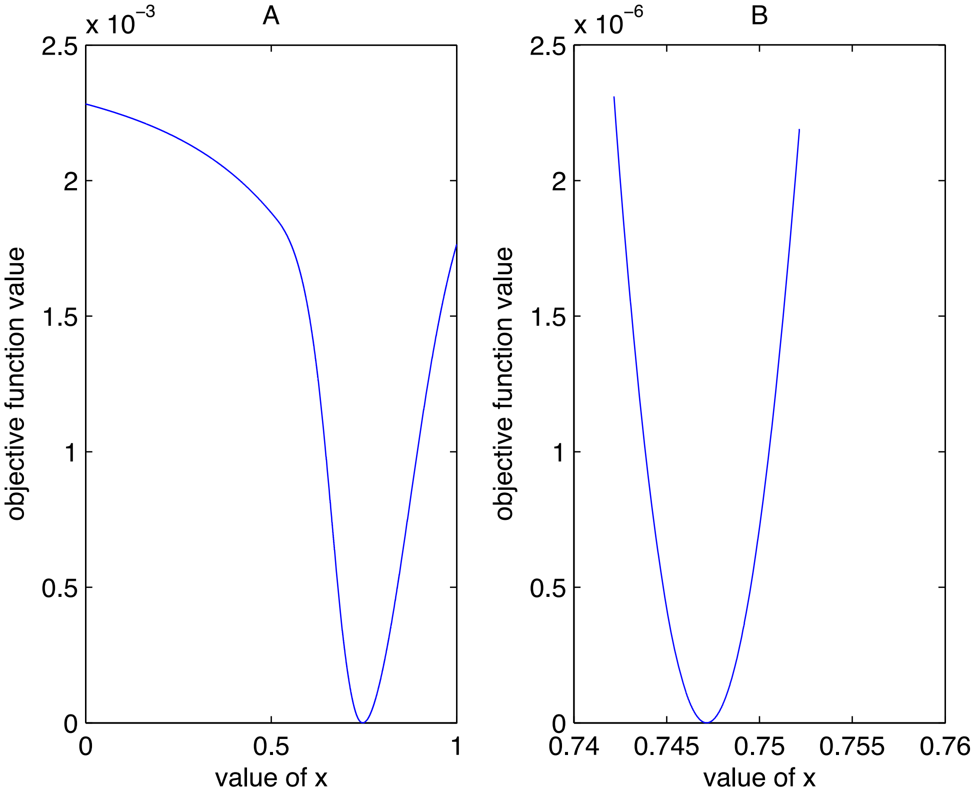

Figure 5. Results based on ideal synthetic data without an optimisation function.

Panel (A) shows the value of the objective function for 1000 equidistant values of x. Panel (B) zooms in on 10 of the previous points belonging to the values of x closest to the x value corresponding to the objective function minimum.

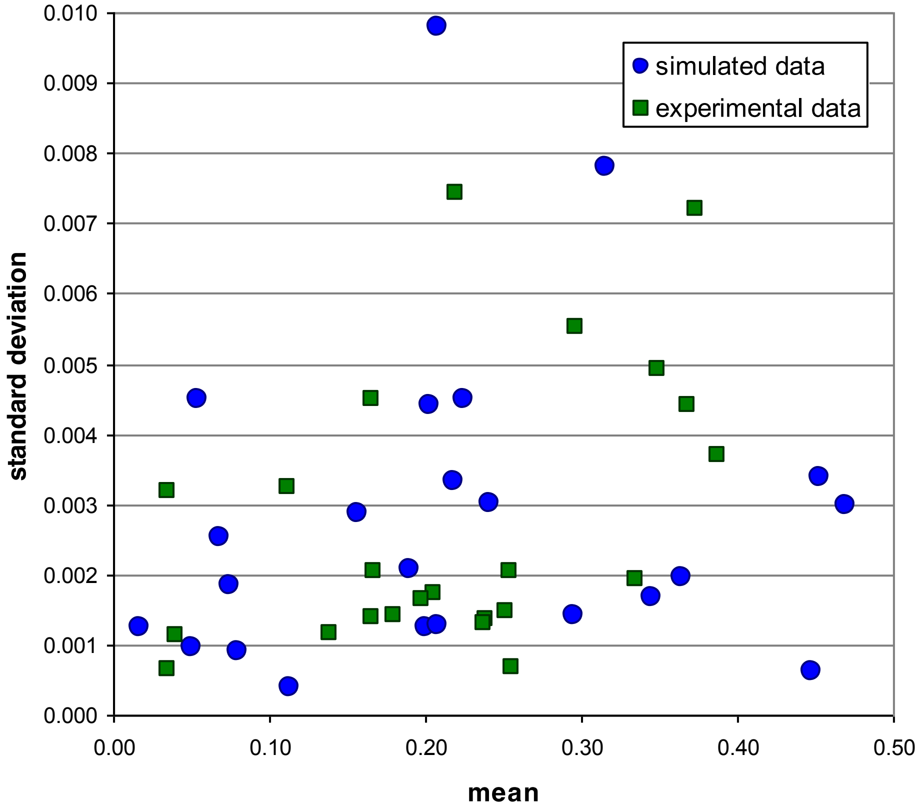

Figure 6. Intensity mean vs. standard deviation for 23 m/z channels across five metabolites.

Intensity mean vs. standard deviation for 23 m/z channels across five metabolites (PEP, Pyr, Mal, Suc, and OAA). For experimental data plotted are the mean and standard deviation calculated from eight replicate experiments. For the simulated data set, *m/z* channel intensities were taken from mass isotopic distributions of five metabolites calculated based on fluxes shown in Figure 3, and standard deviations were derived from Weibull distribution whose parameters were derived from the best fit to experimental data. Only a weak correlation between the mean intensity and variance in the experimental data was observed (a correlation coefficient of 0.448). The plot suggests that the strategy adopted for the simulation of noise in data approximates well noise in experimental data.

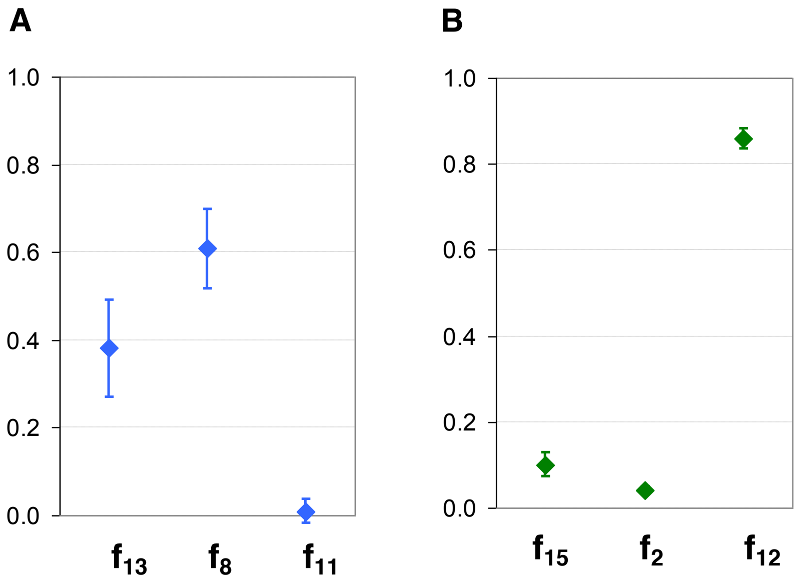

Figure 7. Fractional fluxes f_{13}, f_8, f_{11} (Mal node) and f_{15}, f_2, f_{12} (OAA node).

Fractional fluxes f_{13}, f_8, f_{11} (Mal node) and f_{15}, f_2, f_{12} (OAA node) calculated by optimization via Equation 11 for rapidly dividing *L. mexicana* promastigotes. Mean and standard deviation from eight replicate experiments are shown.

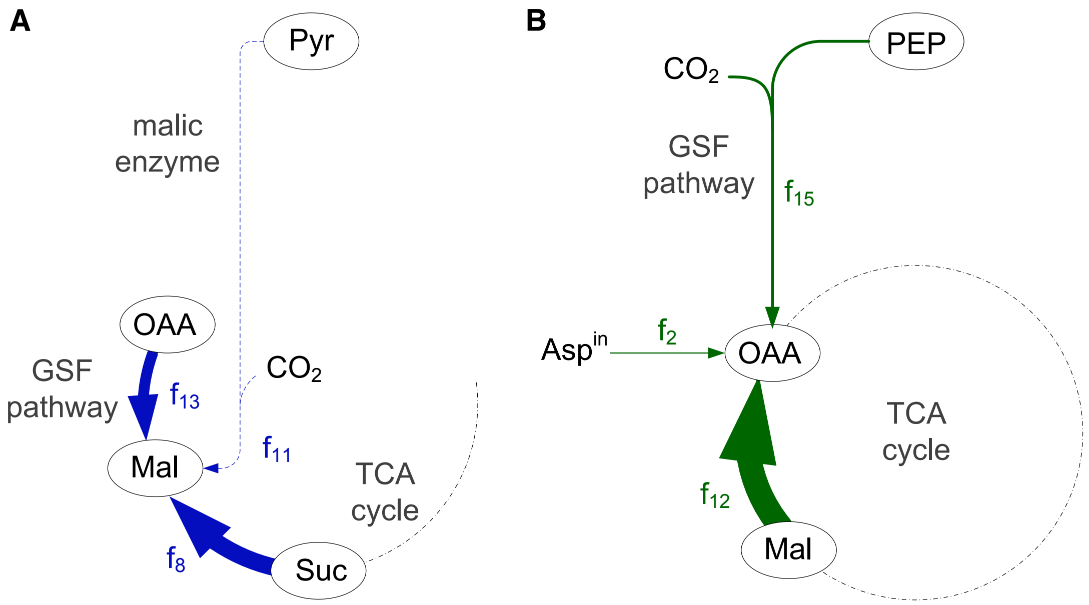
A
Pyr
malic
enzyme
OAA
GSF
pathway
f13
CO2
f11
Mal
TCA
cycle
f8
Suc
B
CO2
PEP
GSF
pathway
f15
Asp^in
f2
OAA
TCA
cycle
f12
Mal

Figure 8. Incoming fractional fluxes around OAA and Mal nodes obtained from experimental data.

Shown is the mean based on eight independent replicate experiments. The thickness of the reaction arrows is proportional to the obtained mean values of fractional fluxes in Table 6. Graphical representation is derived from Figure 2 by focusing only on the relevant metabolic subnetwork, all other metabolites and reactions are omitted.

Tables

Table 1. The reactions for the model shown in Figures 2 and 3, with their stoichiometries.

Flux	**Description**	**Stoichiometry**
v_1	PEP produced from glucose via glycolysis	Glc^{in} = PEP
v_2	Aspartate uptake	Asp^{in} = OAA
v_3	Pyruvate kinase	PEP = Pyr + ATP
v_4	Pyruvate dehydrogenase complex	Pyr = AcCoA + CO_2 + NADH
v_5	Citrate synthase	AcCoA + OAA = Cit
v_6	Aconitase and Isocitrate dehydrogenase	Cit = αKG + CO_2 + NADPH
v_7	α-ketoglutarate dehydrogenase and succinyl-CoA synthase	αKG = Suc + CO_2 + NADH + ATP
v_8	Succinate dehydrogenase and fumarase	Suc = Mal + $CoQ_{10}H_2$
v_9	Fumarate hydratase and fumarate reductase	Mal + NADH = Suc
v_{10}	Malic enzyme	Mal = Pyr + CO_2 + NADPH
v_{11}	Malic enzyme	Pyr + CO_2 + NADPH = Mal
v_{12}	Malate dehydrogenase	Mal = OAA + NADH
v_{13}	Malate dehydrogenase	OAA + NADH = Mal
v_{14}	PEP carboxykinase	OAA + ATP = PEP + CO_2
v_{15}	PEP from glycolysis	PEP + CO_2 = OAA + ATP
v_{16}	Carbon dioxide secretion	$CO_2 = CO_2^{out}$
v_{17}	Alanine secretion	Pyr = Ala^{out}
v_{18}	Acetate secretion	AcCoA = Ace^{out}
v_{19}	Glutamate secretion	αKG = Glu^{out}
v_{20}	Succinate secretion	Suc = Suc^{out}

Table 2. Metabolic steady-state equations for the network shown in Figures 2 and 3.

Node	**Metabolic steady-state equation**
PEP	$v_1 + v_{14} = v_3 + v_{15}$
Pyr	$v_3 + v_{10} = v_4 + v_{11} + v_{17}$
AcCoA	$V_4 = v_5 + v_{18}$
Cit	$V_5 = v_6$
αKG	$V_6 = v_7 + v_{19}$

Suc	$V_7 + v_9 = v_8 + v_{20}$
Mal	$V_8 + v_{11} + v_{13} = v_9 + v_{10} + v_{12}$
OAA	$v_2 + v_{12} + v_{15} = v_5 + v_{13} + v_{14}$
CO_2	$v_4 + v_6 + v_7 + v_{10} + v_{14} = v_{11} + v_{15} + v_{16}$

Table 3. Calculated steady-state mass isotopic distributions for the carbon backbone of the key metabolites for the assumed fluxes in Figure 3.

	m_0	m_1	m_2	m_3	m_4
CO_2	0.747152	0.252848			
$PEP_{(3)}$	0.447893	0.189819	0.017475	0.344812	
$Pyr_{(3)}$	0.453369	0.200297	0.052123	0.294211	
$Asp_{(4)}^{in}$	0.957882	0.041441	0.000672	0.000005	0.00000
$Suc_{(4)}$	0.314743	0.220152	0.226117	0.159334	0.079653
$Mal_{(4)}$	0.363796	0.240843	0.112386	0.208842	0.074134
$OAA_{(4)}$	0.469236	0.208786	0.049947	0.203994	0.068036

Assumed input substrates consisted of 50% [U-$^{13}C_6$] + 50% [1-$^{13}C_1$] glucose and unlabelled aspartate (Glc^{in} and Asp^{in} in Figure 3). Given the assumed fluxes, the forward labelling of key metabolites was calculated using the EMU framework [43]. The numbers in the subscript refer to the number of carbon atoms in the fragment.

Table 4. Comparison of solutions with and without the use of the optimisation function.

Variable	**Actual values**	***FMINCON***	***QUADPROG***
SSR	0.0000	$4.89 \cdot 10^{-11}$	$2.05 \cdot 10^{-12}$
f_8	0.2353	0.235320	0.235298
f_{13}	0.2353	0.235298	0.235281
f_2	0.2143	0.214273	0.214281
f_{15}	0.7500	0.749971	0.749979

Comparing our approach (without the use of an optimisation function) with Matlab's *FMINCON* optimisation function. The second column shows the best of 1,000,000 randomly initialised *FMINCON* optimisation runs. The third column shows the values obtained via the quadratic programming approach. Note that the values in both columns were produced using ideal synthetic data.

Table 5. Synthetic fractional flux estimation in the presence of noise.

	Oxaloacetate node			**Malate node**			
	$\mathbf{f_{12}}$	$\mathbf{f_2}$	$\mathbf{f_{15}}$	$\mathbf{f_{13}}$	$\mathbf{f_8}$	$\mathbf{f_{11}}$	**SSR**
actual	**0.0357**	**0.2143**	**0.7500**	**0.2353**	**0.2353**	**0.5294**	
#1	0.0495	0.2112	0.7393	0.2185	0.2352	0.5462	$3.21 \cdot 10^{-05}$

#2	0.1031	0.2130	0.6840	0.2226	0.1819	0.5955	$1.14\cdot10^{-04}$
#3	0.0087	0.2140	0.7773	0.1957	0.1820	0.6223	$2.20\cdot10^{-05}$
#4	0.0000	0.2204	0.7796	0.2490	0.2568	0.4943	$9.55\cdot10^{-05}$
#5	0.0026	0.2263	0.7712	0.2873	0.2707	0.4420	$2.36\cdot10^{-05}$
#6	0.0746	0.2135	0.7119	0.2202	0.2473	0.5326	$1.06\cdot10^{-05}$
#7	0.0516	0.2206	0.7277	0.2465	0.2386	0.5149	$2.15\cdot10^{-05}$
#8	0.0449	0.2070	0.7481	0.2545	0.2559	0.4896	$8.84\cdot10^{-06}$
mean	**0.0419**	**0.2157**	**0.7424**	**0.2368**	**0.2336**	**0.5297**	
stdev	**0.0366**	**0.0062**	**0.0339**	**0.0283**	**0.0337**	**0.0585**	

The results shown are based on simulated noisy data (8 synthetic replicates).

Table 6. Mass isotopic distributions for PEP, Pyr, Suc, Mal, and OAA in rapidly dividing *Leishmania mexicana* promastigotes at isotopic and metabolic steady state measured by GC-MS.

	m_0	**$m_0 + 1$**	**$m_0 + 2$**	**$m_0 + 3$**	**$m_0 + 4$**
PEP					
#1	0.38455	0.20987	0.11943	0.28615	
#2	0.37930	0.21352	0.11075	0.29644	
#3	0.36596	0.23151	0.11043	0.29209	
#4	0.37409	0.21702	0.10981	0.29909	
#5	0.37149	0.21954	0.11077	0.29819	
#6	0.37592	0.21428	0.11387	0.29593	
#7	0.36610	0.21884	0.10979	0.30527	
#8	0.36393	0.22847	0.11218	0.29542	
Pyr					
#1	0.38218	0.23950	0.14000	0.23832	
#2	0.38705	0.23748	0.13820	0.23727	
#3	0.39041	0.23660	0.13741	0.23558	
#4	0.38387	0.23952	0.13962	0.23698	
#5	0.39245	0.23608	0.13719	0.23428	
#6	0.38715	0.23756	0.13855	0.23674	
#7	0.38345	0.23938	0.13952	0.23765	
#8	0.39008	0.23710	0.13718	0.23564	
Suc					
#1	0.33850	0.25392	0.20144	0.16545	0.04070
#2	0.33424	0.25585	0.20493	0.16515	0.03982
#3	0.33238	0.25613	0.20762	0.16223	0.04164
#4	0.33631	0.25507	0.20521	0.16500	0.03841
#5	0.33504	0.25559	0.20507	0.16586	0.03844
#6	0.33312	0.25563	0.20422	0.16668	0.04035
#7	0.33600	0.25541	0.20533	0.16436	0.03890
#8	0.33544	0.25512	0.20575	0.16356	0.04013
Mal					
#1	0.35553	0.25190	0.19805	0.15909	0.03543
#2	0.34944	0.25284	0.19578	0.16669	0.03525

#3	0.34035	0.25339	0.19901	0.16815	0.03909
#4	0.35047	0.25311	0.19523	0.17021	0.03098
#5	0.35427	0.25523	0.19517	0.16446	0.03088
#6	0.34552	0.25818	0.19886	0.16142	0.03601
#7	0.34642	0.25226	0.19846	0.17282	0.03004
#8	0.34664	0.25460	0.19817	0.16466	0.03594
OAA					
#1	0.36490	0.25232	0.17912	0.16952	0.03414
#2	0.36846	0.25170	0.17974	0.16568	0.03442
#3	0.37368	0.24910	0.17746	0.16435	0.03542
#4	0.36197	0.25387	0.18154	0.16793	0.03469
#5	0.37238	0.25012	0.17976	0.16376	0.03398
#6	0.36467	0.25087	0.18096	0.16801	0.03549
#7	0.36321	0.25248	0.18173	0.16808	0.03450
#8	0.37055	0.25107	0.17918	0.16553	0.03366

Heavy carbon was fed through 50%[U-$^{13}C_6$] + 50%[1-$^{13}C_1$] glucose for 15 hrs (see Methods). PEP, Mal and Suc were TMS derivatised, while alanine (used for Pyr) and aspartate (used for OAA) were TBDMS derivatised before GC-MS analysis. Measured intensities were obtained by manual integration using ChemStation (Agilent Technologies). For each metabolite raw intensities were normalized, and corrected for natural isotopic contributions originating from all atoms except for the backbone carbon atoms.

Table 7. Flux ratio estimation results in rapidly dividing *L. mexicana* promastigotes.

	Oxaloacetate node			**Malate node**			
	f_{12}	**f_2**	**f_{15}**	**f_{13}**	**f_8**	**f_{11}**	**SSR**
#1	0.8172	0.0246	0.1582	0.4271	0.5729	0.0000	$1.13\cdot10^{-04}$
#2	0.8827	0.0384	0.0789	0.4257	0.5743	0.0000	$4.05\cdot10^{-05}$
#3	0.8611	0.0560	0.0829	0.1694	0.7523	0.0783	$4.84\cdot10^{-05}$
#4	0.8846	0.0339	0.0816	0.5190	0.4810	0.0000	$5.26\cdot10^{-05}$
#5	0.8630	0.0440	0.0930	0.4821	0.5179	0.0000	$2.92\cdot10^{-05}$
#6	0.8353	0.0377	0.1270	0.3268	0.6732	0.0000	$3.98\cdot10^{-05}$
#7	0.8701	0.0445	0.0854	0.3925	0.6075	0.0000	$9.43\cdot10^{-05}$
#8	0.8568	0.0450	0.0983	0.3129	0.6871	0.0000	$3.33\cdot10^{-05}$
mean	**0.8589**	**0.0405**	**0.1006**	**0.3819**	**0.6083**	**0.0098**	
stdev	**0.0229**	**0.0093**	**0.0279**	**0.1108**	**0.0910**	**0.0277**	

The results shown are based on experimental data (Table 6) and Equation 13. Mass isotopic distribution vector of CO_2, $[x, 1 - x]$, was also computed ($x = 0.7537 \pm 0.1058$).