Accepted for publication in Molecular Biosystems, doi: 10.1039/C3MB70348G

Flux and reflux: Metabolite reflux in plant suspension cells and its implications on isotope-assisted metabolic flux analysis

Shilpa Nargund[1,¶,&], Ashish Misra[1,#,&], Xiaofeng Zhang[1,&], Gary D. Coleman[2] and Ganesh Sriram[1,*]

[1]Department of Chemical and Biomolecular Engineering and [2]Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD 20742, USA

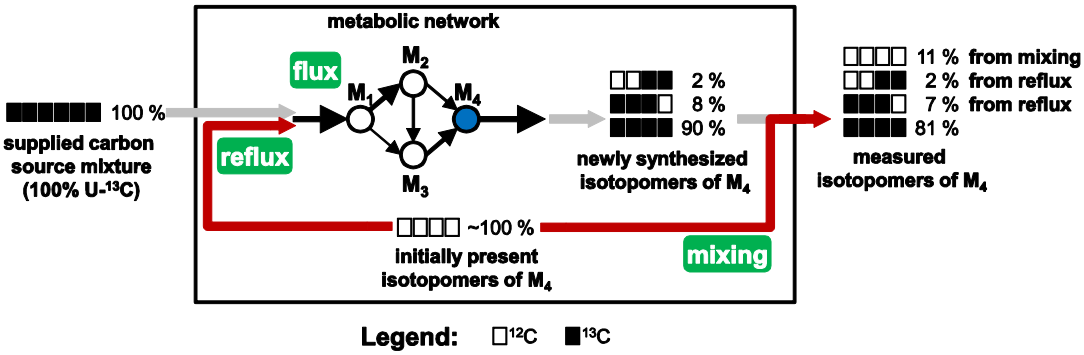
Current affiliations: [¶]Amgen Inc., Thousand Oaks, CA, USA; [#]DBT-ICT Centre for Energy Biosciences, Institute of Chemical Technology, Mumbai, India

[&]These three authors contributed equally to this work.

[*]**Corresponding author:** Ganesh Sriram. Address: 1208D, Chemical and Nuclear Engineering Building 090, College Park, MD 20742. Phone: +1 301 405-1261; Fax: +1 301 405-0523; Email: gsriram@umd.edu.

Running head: Metabolite reflux in isotope-assisted flux analysis

Graphical Abstract



Graphical abstract description. Metabolite reflux is a common phenomenon in isotope labeling experiments employed for metabolic flux analysis, but its effects on the results of flux analysis have not been investigated. Through targeted isotope labeling experiments, we investigate this phenomenon in poplar cell suspensions. Furthermore, we show that accurate modeling of reflux improves the accuracy of flux estimation.

Accepted for publication in Molecular Biosystems, doi: 10.1039/C3MB70348G

Abstract

Isotope-assisted metabolic flux analysis (MFA) is a powerful methodology to quantify intracellular fluxes through isotope labeling experiments (ILEs). In batch cultures, which are often convenient, inexpensive or inevitable especially for eukaryotic systems, MFA is complicated by the presence of the initially present biomass. This unlabeled biomass may either mix with the newly synthesized labeled biomass or reflux into the metabolic network, thus masking the true labeling patterns in the newly synthesized biomass. Here, we report a detailed investigation of such metabolite reflux in cell suspensions of the tree poplar. In ILEs supplying 28% or 98% U-^{13}C glucose as the sole organic carbon source, biomass components exhibited lower ^{13}C enrichments than the supplied glucose as well as anomalous isotopomers not explainable by simple mixing of the initial and newly synthesized biomass. These anomalous labeling patterns were most prominent in a 98% U-^{13}C glucose ILE. By comparing the performance of light- and dark- grown cells as well as by analyzing the isotope labeling patterns in aspartic and glutamic acids, we eliminated photosynthetic or anaplerotic fixation of extracellular $^{12}CO_2$ as explanations for the anomalous labeling patterns. We further investigated four different metabolic models for interpreting the labeling patterns and evaluating fluxes: (i) a carbon source (glucose) dilution model, (ii) an isotopomer correction model with uniform dilution for all amino acids, (iii) an isotopomer correction model with variable dilution for different amino acids, and (iv) a comprehensive metabolite reflux model. Of these, the metabolite reflux model provided a substantially better fit for the observed labeling patterns (sum of squared residues: 538) than the other three models whose sum of squared residues were (i) 4,626, (ii) 4,983, and (iii) 1,748, respectively. We compared fluxes determined by the metabolite reflux model to those determined by an independent methodology involving an excessively long ILE to wash out initial biomass and a minimal reflux model. This comparison showed identical or similar distributions for a majority of fluxes, thus validating our comprehensive reflux model. In summary, we have demonstrated the need for quantifying interactions between initially present biomass and newly synthesized biomass in batch ILEs, especially through the use of $\approx$100% U-^{13}C carbon sources. Our ILEs reveal a high amount of metabolite reflux in poplar cell suspensions, which is well explained by a comprehensive metabolite reflux model.

Keywords

metabolic flux analysis, isotope labeling experiment, reflux, proteinogenic amino acid, poplar, cell suspension

Accepted for publication in Molecular Biosystems, doi: 10.1039/C3MB70348G

Introduction

Isotope-assisted metabolic flux analysis (isotope MFA) is a powerful tool for investigating metabolic networks because it can estimate intracellular chemical reaction rates (fluxes)[1–4] and elucidate new metabolic pathways[5,6]. The basis of this methodology is the isotope labeling experiment (ILE), which involves supplying one or more isotopically labeled substrates to a cell or tissue culture, followed by measurement of isotope labeling patterns (*isotope* i*somers*, or isotopomers) in synthesized products such as biomass components or intracellular metabolites[7]. Isotope MFA involves interpretation of these labeling patterns through a mathematical model of metabolism to ultimately result in metabolic flux maps for the cell or tissue of interest. This modeling component serves as a bridge that connects the measurable quantities in an ILE, including isotopic labeling patterns and extracellular or biomass fluxes, to the underlying intracellular fluxes.

Two major approaches in isotope MFA are (i) steady-state MFA[e.g. 7], in which the labeling states of cellular products are measured after isotopic steady state is reached, and (ii) instationary MFA[e.g. 8,9], in which the labeling states are intentionally measured during the approach to isotopic steady state. Of these, steady-state MFA is currently more widespread due to two reasons: first, its experimental implementation is easier and second, it is computationally less challenging as it does not involve solutions of large systems of ordinary differential equations within a global optimization algorithm[8,10]. For both types of MFA, it is ideal to measure isotope labeling patterns in intracellular metabolites. However, in steady-state isotope MFA, biomass components such as proteinogenic amino acids and sugars are used as surrogates for intracellular metabolites[e.g. 11–14]. This is because these compounds are synthesized from intracellular metabolite precursors distributed throughout primary metabolism. Furthermore, cellular protein and sugars are much more abundant than intracellular metabolites and can be hydrolyzed to amino acids[15] or sugar hydrolysates[16] directly without employing any extraction steps. In this case, information on fluxes is contained in the isotopically labeled biomass synthesized *de novo* during the ILE from the supplied carbon sources. Biomass present before the commencement of the ILE (hereafter, “initial biomass”) is labeled at the natural ^{13}C abundance of 1.1%, and is thus effectively unlabeled. Often, this initial biomass can constitute up to ~10% of the biomass harvested at the end of the ILE and can interfere with the interpretation of labeling patterns. Continuous culture, such as in a chemostat, affords the washing out of the initial biomass and has been employed for many microbial systems. However, continuous culture is infeasible, unsuitable or expensive for many plant, mammalian and even microbial cell types, leaving batch culture as the only available option. In fact, steady-state isotope MFA in batch culture has been reported in studies investigating the metabolic landscapes of the bacterium *E. coli*[17,18], cell cultures of the plant *Arabidopsis thaliana*[19,20] and various mammalian cells[21–23].

Unless batch culture is performed for long periods of time extending into multiple subculture cycles, the naturally abundant initial biomass can never be washed out. Thus, the biomass harvested at the conclusion of the ILE will consist of a mixture of initial (naturally abundant) and newly synthesized (^{13}C-labeled) portions (**Fig. 1**). This raises an important question: does the initial biomass interact with the newly synthesized biomass? If there is no interaction, but only simple mixing between the initial and newly synthesized biomass, the ratio of the initial biomass to the newly synthesized biomass can be used to factor out the isotopomer abundances of the initial biomass from those measured in the harvested biomass. However, interaction between the initial biomass and the newly synthesized biomass (or “reflux” of the initial biomass into the metabolic network) can result in additional labeling patterns that can only be factored out by a metabolic model that mimics this interaction. Previous isotope MFA studies have used data correction strategies toward factoring out the effect of the initial biomass. A straightforward

method is to assume that the initial biomass does not reflux into the metabolic network and that the measured isotopomers are a weighted average of the known, naturally abundant isotopomers of the initial biomass and the unknown, isotopomers of the newly synthesized biomass. In this case, a simple algebraic operation decouples the isotopomer abundances of the initial biomass from those of the newly synthesized biomass. This "dilution rate" method has been reported in multiple studies[12,13,6]. Another approach is to assume a small influx of the naturally abundant version of the supplied carbon source (e.g. glucose), the value of this influx being a parameter estimated by MFA[24]. Recently, researchers have begun hypothesizing interactions between the initial biomass and the newly synthesized biomass, and modeling these interactions as influxes of unlabeled amino acids. Such models either assume identical influxes for each amino acid on the basis of the culture growth rate[25], or different influxes for each amino acid, which are estimated as part of the flux evaluation procedure[26,19]. Although these solutions have partially solved the intrinsic shortcoming of steady state MFA in batch-like cell cultures resulted from initial biomass, their validity, performance and ability to obtain accurate flux estimates have not yet been comprehensively investigated.

We report an experimental and computational investigation of the interactions between initial and newly synthesized biomass, or metabolite reflux, in batch-grown cell suspensions of the forest tree poplar, a potential biofuel feedstock[27]. We performed ILEs by supplying three labeled varieties of ^{13}C glucose. Of these, a 98% U-^{13}C glucose ILE revealed the presence of numerous, anomalous mass isotopomers that appeared to have been formed due to metabolite reflux. By comparing labeling patterns in light- and dark- grown cells as well as by interpreting the labeling patterns of tricarboxylic acid (TCA) cycle metabolites, we determined that these mass isotopomers were neither contributed by photosynthesis nor by anaplerotic CO_2 assimilation. On elimination of these processes, metabolite reflux remained the sole cause for the anomalous mass isotopomers. We then compared the ability of three previously reported flux-isotopomer models to account for the observed mass isotopomer data. These models accounted for the anomalies by (i) using a uniform or non-uniform (for different amino acids) dilution rate to correct the isotopomer abundances, (ii) assuming a small influx of naturally abundant glucose, or (iii) accounting for the reflux by full-fledged reflux model in which reflux is mimicked by influxes of several amino acids and metabolites. Our results show significant metabolite reflux in poplar suspension cells, with the 98% U-^{13}C glucose ILE being most sensitive to the effects of reflux. Among the four models we examined, the comprehensive reflux model accounts best for the observed anomalies in the labeling patterns. While flux results from this model compare qualitatively with those obtained from the labeling patterns of intracellular metabolites, there are some numerical differences between fluxes. This points to the need for future extensions of the metabolite reflux model.

Materials and methods

Poplar cell suspension culture and growth rate measurement

Poplar (*Populus. tremula* × *Populus. alba* clone 717-1B4) cell suspensions were grown at 20 °C in 125 mL Erlenmeyer flasks shaken at 125 rpm on an orbital shaker. The flasks were placed in continuous light from cool-white fluorescent light (Ecolux® Technology Plant and Aquarium F40T12 bulbs (photosynthetically active radiation of 200-300 μmol m^{-2} s^{-1}, incident radiation on flasks measured to be 28.3 ± 2.6 μmol $m^{-2}s^{-1}$) or continuous dark; the dark condition was realized by wrapping the flasks in aluminum foil. The suspensions were subcultured every 7 d by transferring 600 mg cells into 30 mL Murashige and Skoog medium (Phytotechnology Laboratories, Shawnee Mission, KS) containing 20 g L^{-1} glucose as the sole organic carbon source.

To quantify growth, several flasks of cells were subcultured in parallel. Three biological replicates were harvested at 7 d by vacuum filtration through glass microfiber filter paper (Whatman, Piscataway, NJ). Fresh and dry weights of the harvested cell pellets were obtained by weighing them before and after lyophilization, respectively.

Parallel ILEs

We performed three parallel ILEs by growing cells in continuous light, replacing the usual 20 g L^{-1} naturally abundant glucose in the medium with 20 g L^{-1} of: (i) 98% U-^{13}C glucose, (ii) 100% 1-^{13}C glucose or (iii) a mixture of 28% U-^{13}C and 72% naturally abundant glucose. The 98% U-^{13}C glucose ILE was also performed in continuous dark. Isotopically labeled glucose was obtained from Cambridge Isotope Laboratories (Andover, MA). Each ILE was represented by three biological duplicates, which were harvested on day 7 of culture. Cells were harvested from the ILEs by vacuum filtration through glass microfiber filter paper, and were immediately frozen in liquid nitrogen to arrest metabolism. The frozen cell pellets were then lyophilized and stored at –80°C until further analysis. We further performed three 21-d ILEs by supplying the same three types of ^{13}C labeled glucose. In these ILEs, we transferred cells to fresh media every 7 d, harvested cells at 21 d and analyzed the mass isotopomers of soluble metabolites (free amino acids) from the harvested cells.

Measurement of isotopomer abundances of proteinogenic amino acids and intracellular metabolites

To analyze proteinogenic amino acids, 20 mg of ground, lyophilized cells were vacuum-hydrolyzed with 6 N HCl for 5 h at 160 °C in hydrolysis tubes (Thermo Fisher Scientific, Rockford, IL). The hydrolysate was dried overnight in a RapidVap evaporator (Labconco, Kansas City, MO), reconstituted in 1 mL water, filtered, lyophilized and stored in -80 °C until further analysis. To improve their volatility in the GC, the amino acids were derivatized with 100 µL of N-(tert-butyldimethylsilyl)-N-methyltrifluoroacetamide (MTBSTFA; Sigma-Aldrich, St. Louis, MO) in 100 µL dimethylformamide (Thermo Fisher Scientific) by heating at 70 °C for 1 h.

Intracellular metabolites were extracted and quantified as described by Fiehn and co-workers[28]. Briefly, 20 mg of ground, lyophilized cells were contacted with 1 mL methanol and 50 µL water, heated at 75 °C for 15 min and then centrifuged, after which the supernatant containing the soluble metabolites was collected. These steps were repeated once, following which the supernatant was dried, re-dissolved in 100 µL of a 20 mg mL^{-1} solution of methoxyamine hydrochloride (MP Biomedicals, Solon, OH) in pyridine (EMD Chemicals, Gibbstown, NJ) and heated at 30 °C for 90 min. The resulting soluble metabolites were derivatized by adding 100 µl MTBSTFA and heating at 70 °C for 1 hr. Soluble metabolites were quantified by gas-chromatography as described below, using norleucine (Sigma-Aldrich, St. Louis, MO) as an internal standard.

A Varian 450 gas chromatograph (GC) connected to a Varian 300 mass spectrometer (MS) (Bruker Corporation, Fremont, CA) was used to measure the abundances and mass isotopomer distributions of amino acids. The GC was equipped with a VF-5ms column (30 m × 0.25 mm × 0.25 µm; Bruker). One microliter of each sample was injected in to the GC at a split ratio of 50:1, with helium as the carrier gas at a flow rate at 1.0 mL min^{-1}. To analyze proteinogenic amino acids, the oven temperature was programmed as follows: an initial temperature of 150 °C for 2 min, increased to 230 °C at 3 °C min^{-1}, then to 240 °C at 2 °C min^{-1} and finally to 275 °C at 10 °C min^{-1} and a constant temperature of 275 °C for 6 min, amounting to a run time of 43.2 min. To analyze intracellular metabolites, the oven temperature was programmed as follows: an initial temperature of 150 °C for 2 min, increased to 250 °C at 3 °C min^{-1}, then finally to 275 °C at 10 °C min^{-1} and a constant temperature of 275 °C for 5 min. The MS was operated in

electron ionization mode, with the ion source at 280 °C. Isotopomer abundances were recorded under selected ion monitoring mode.

The Varian MS Workstation software (version 6.9.3) was used together with the NIST mass spectral library (National Institute of Standards and Technology, Gaithersburg, MD) to identify and quantify the peaks in the MS spectra. An in-house MATLAB (The Mathworks, Natick, MA) script[14] was used to correct the mass isotopomer distributions for the natural abundances of isotopes of hydrogen, nitrogen, oxygen, sulfur, silicon and non-metabolic carbon. By acquiring and processing spectra of amino acid standards of defined compositions, we verified that both the MS measurements and the natural abundance correction program were accurate (data not shown). The thus corrected mass isotopomer distributions from the three parallel ILEs are listed in Supplementary **Table S1**. This correction different from the data correction strategy described in Introduction because the data correction strategy aims at elucidating the effects of initial ^{12}C naturally abundant biomass in a ILE, while this MATLAB program corrects for all non-metabolic carbon atoms in a measurable molecule resulted from derivatization.

Flux evaluation from isotopomer data

To estimate intracellular metabolic fluxes, the corrected mass isotopomer distributions were processed by our flux evaluation program NMR2Flux+. Details about NMR2Flux+ and its capabilities are described elsewhere[13,14,20]. Briefly, this program accepts metabolic network stoichiometries and carbon atom rearrangements supplied by the user, employs cumomer balancing to simulate isotopomer distributions and isotopomer abundances corresponding to any feasible set of fluxes in the network. Given a set of isotopomer abundances, the program uses a global optimization routine based on simulated annealing to iteratively evaluate flux values that best account for the supplied isotopomer abundances. NMR2Flux+ assesses goodness-of-fit using the sum of squared residuals (SSR) metric, calculated as:

$$\mathrm{SSR} = \sum_j \frac{I_j - I_j^x{}^2}{\sigma_j{}^2} \qquad [1]$$

where the index j cycles through all the (mass) isotopomer abundances, I^x represents measured isotopomer abundances, I represents simulated isotopomer abundances corresponding to the evaluated set of fluxes and σ represents the measured standard deviations of the isotopomer abundances. For all metabolic models, the SSR was evaluated for 405 isotopomer abundances (see Supplementary **Table S1**) measured from three ILEs on 98% 1-^{13}C glucose, 28% U-^{13}C glucose and 98% U-^{13}C glucose. The dataset used for flux evaluation includes all the proteinogenic amino acid fragments detected by GC-MS (a total of 30 fragments across 12 amino acids: Ala, Asp, Gly, Glu, His, Ile, Leu, Phe, Pro, Ser, Thr, Val; Supplementary **Table S1**).

To obtain standard deviations for the evaluated fluxes, we used a bootstrap Monte-Carlo algorithm that operated as follows: (i) the original isotopomer abundance dataset was perturbed 280 times, resulting in 280 synthetic datasets normally distributed around the original dataset; (ii) NMR2Flux+ independently evaluated fluxes from each synthetic dataset to obtain 280 synthetic flux distributions; (iii) these synthetic distributions were used to generate standard deviations for the evaluated fluxes. Fluxes and standard deviations evaluated by NMR2Flux+ for the four metabolic models in this work – uniform isotopomer correction (Iso_corr[U]), variable isotopomer correction model (Iso_corr[V]), glucose dilution (Gluc_dilu) and amino acid reflux (AA_in) (see Results for details) – are listed in Supplementary **Table S3**. The metabolic networks corresponding to these four models are listed in Supplementary **Tables S4, S5, S6** and **S7**, respectively. All NMR2Flux+ computations were performed on three identical workstations equipped with an Intel Xeon E5620 CPU and 32 GB of RAM.

Accepted for publication in Molecular Biosystems, doi: 10.1039/C3MB70348G

Results

Proteinogenic amino acids from a 7-d batch ILE exhibit anomalous ^{13}C enrichments and mass isotopomer distributions

The net ^{13}C enrichments (evaluated as explained in **Fig. 2**) of proteinogenic amino acid fragments from poplar suspension cultures grown on 28% or 98% U-^{13}C glucose for 7 d in batch cultures were considerably lower (22%-28%) (**Fig. 2**) than that of the supplied glucose. For growth on 28% U-^{13}C glucose, the enrichments ranged from 22%-23% for isoleucine, lysine and proline fragments to 27%-28% for glutamate and alanine fragments. Proportionally, for growth on 98% U-^{13}C glucose, the enrichments ranged from 76% for isoleucine, lysine and proline fragments to 92%-98% for glutamate and alanine fragments. These enrichment data were reproduced in several labeling experiments performed weeks apart from each other. The observed ^{13}C enrichments are anomalous because they contradict the expectation that the non-photosynthetic poplar cell suspensions should synthesize biomass components entirely from the supplied glucose. A recent study[29] shows that enzymes can discriminate against ^{13}C atoms principally because ^{13}C atoms form stronger bonds than ^{12}C atoms, and that this isotopic effect can contribute up to 0.8% in GC-MS measurements when particular ^{13}C tracers are used. Nevertheless, The observed discrepant ^{13}C enrichments (up to 20%) are too large to be explained by the difference in enzyme affinities for ^{12}C and ^{13}C. Additionally, a previous study[30] demonstrated that feeding ^{13}C-labeled carbon sources to Arabidopsis cell suspensions does not perturb metabolism.

One explanation for the discrepant ^{13}C enrichments is that they were due to mixing of initial biomass with the newly synthesized biomass without reflux of the initial biomass. However, an examination of the mass isotopomer abundances from a 98% U-^{13}C ILE (**Fig. 3**, Supplementary **Fig. S1** and Supplementary **Table S1**) revealed that this explanation is invalid. Mixing between the initial biomass and the newly synthesized biomass at a ratio of 1:9 (as obtained from dry weight measurements before and after the ILE) would result in the least massive (entirely ^{12}C) and most massive (entirely ^{13}C) mass isotopomers for each fragment (left panels of **Fig. 3** and Supplementary **Fig. S1**). Contrastingly, reflux of the initial biomass would result in breakdown of the initial biomass and condensation of the ensuing fragments with ^{13}C labeled fragments from the supplied glucose, thus producing a range of mass isotopomers for each fragment (right panels of **Fig. 3** and Supplementary **Fig. S1**). The measured mass isotopomer abundances (center panels of **Fig. 3** and Supplementary **Fig. S1**) clearly agree with the latter possibility. This suggests reflux of the initial biomass in addition to simple mixing. It may appear that two competing hypotheses – photosynthetic or anaplerotic fixation of $^{12}CO_2$ – could potentially explain the anomalous isotopomers. However, as explained below, further experiments and a detailed examination of isotopomer data did not support these hypotheses.

Photosynthetic $^{12}CO_2$ assimilation does not explain anomalous isotopomers

To assess if photosynthetic $^{12}CO_2$ fixation explained the anomalous isotopomers, we performed 98% U-^{13}C glucose ILEs for 7 d in continuous light or continuous dark. Photosynthetic $^{12}CO_2$ fixation would occur in the light but not in the dark. Therefore, if it were responsible for the anomalous ^{13}C enrichments of **Fig. 2**, only light-grown cells would exhibit substantially lower enrichments than the supplied glucose. Furthermore, since photosynthesis will result in $^{12}CO_2$ incorporation at C-1 of the photosynthetic product 3-phosphoglycerate, the light-grown cells should exhibit lower ^{13}C enrichment on C-1 of Ser (an amino acid derived from 3-phosphoglycerate) than dark-grown cells. Contrary to this expectation, intracellular metabolites (amino acids) from light- and dark-grown cell suspensions exhibited nearly identical ^{13}C enrichments (**Fig. 4**), including Ser. This clearly demonstrates that the discrepant ^{13}C enrichments cannot be explained by photosynthetic fixation of extracellular $^{12}CO_2$. Although the

light- and dark-grown cell suspensions exhibit some differences between their glutamate and tyrosine enrichments, these differences cannot be explained by photosynthetic CO_2 fixation as these metabolites are not generated in the Calvin cycle.

Anaplerotic $^{12}CO_2$ assimilation does not fully explain anomalous isotopomers

Anaplerotic reactions interconvert three-carbon glycolytic metabolites such as phosphoenolpyruvate or pyruvate and four-carbon tricarboxylic acid (TCA) cycle metabolites such as oxaloacetate or malate, by fixing or releasing CO_2. Significant flux through these reactions can be expected during growth as they are necessary for biomass synthesis and for shuttling reductants between organelles[6]. To determine whether net anaplerotic incorporation of extracellular $^{12}CO_2$ was solely responsible for the anomalous isotopomers, we scrutinized the mass isotopomers of aspartate and glutamate (**Fig. 5**). The significant anomalous isotopomers of these amino acids include the *M*+1 isotopomers of aspartate[12] and aspartate[234], the *M*+2 and *M*+3 isotopomers of glutamate[2345] as well as the *M*+3 and *M*+4 isotopomers of glutamate[12345] (**Fig. 5c**).

In a 98% U-^{13}C glucose ILE, anaplerotic fixation of $^{12}CO_2$ would cause ^{13}C depletion of oxaloacetate (OAA; the keto acid precursor of aspartate) at C-4 and at C-1 due to exchange between OAA and the symmetrical molecules fumarate and succinate. It would also cause ^{13}C depletion of α-ketoglutarate (AKG; the keto acid precursor of glutamate) at C-1 due to the forward TCA cycle (**Fig. 5a**). Anaplerotic $^{12}CO_2$ fixation, even if followed by infinite turns of the TCA cycle, cannot result in ^{13}C depletion at C-2 or C-3 of OAA or aspartate nor at C-2, C-3, C-4 or C-5 of AKG or glutamate. Hence, anaplerotic $^{12}CO_2$ fixation cannot produce the *M*+1 isotopomer of aspartate[12] or aspartate[234], nor the *M*+1, *M*+2 or *M*+3 isotopomers of glutamate[2345] or glutamate[12345] (**Fig. 5b**). However, the significant observed abundances of these isotopomers imply that even if significant anaplerotic $^{12}CO_2$ occurred during the ILEs, it would not explain all anomalous isotopomers of aspartate and glutamate.

On exploring various metabolic models that could explain the anomalous isotopomers of aspartate and glutamate, we arrived at a model that included reflux of acetyl-CoA (ACA) in addition to anaplerotic $^{12}CO_2$ fixation (**Fig. 5d**). This is consistent with amino acid reflux, as the degradation of many amino acids including glutamate, serine, alanine, threonine, valine and tyrosine produces pyruvate, ACA or fatty acids, and the degradation of fatty acids produces ACA (degradation pathways obtained from the MetaCyc metabolic pathway database[31]). By tracing carbon atom rearrangements, it is clear that this model results in ^{13}C depletions at appropriate carbon atoms in the TCA cycle metabolites to satisfactorily explain all the anomalous isotopomers of aspartate and glutamate (**Fig. 5e**). Simulations using a metabolic model (development explained below) also showed that the anomalous isotopomers were best explained by the influx of naturally abundant ACA, but neither 1-^{13}C nor 2-^{13}C ACA. In summary, these results eliminated photosynthesis and anaplerotic $^{12}CO_2$ fixation as explanations for the anomalous ^{13}C enrichments and isotopomers, leaving only metabolite reflux as an explanation. In particular, the reflux of ACA, that can be derived from degradation of amino acids and lipids, was crucial to explain several anomalous isotopomers.

Metabolic model incorporating reflux quantitatively explains isotopomer data better than previously reported models

We developed a comprehensive flux-isotopomer model to simulate isotopomer abundances and infer a carbon traffic in the batch-grown poplar cell suspensions (**Fig. 6**). Our model consists of plant primary metabolic pathways including glycolysis, the pentose phosphate pathway (PPP), the TCA cycle, multiple anaplerotic reactions as well as reactions leading from intracellular metabolites to biomass components such as proteinogenic amino acids and lipogenic fatty acids. Furthermore our model is compartmentalized, with the compartmental organization of

pathways based on previous plant metabolic network models published by one of the us (GS)[13,6,32,20] and one of the models evaluated by Ratcliffe and co-workers[33]. Accordingly, our model features duplicated (cytosolic and plastidic) glycolysis and pentose phosphate pathways connected by multiple intercompartmental transport fluxes as well as a mitochondrial TCA cycle. U-^{13}C glucose and $^{12}CO_2$ (though anaplerotic fixation) are the principal carbon sources in the model. The model incorporates metabolite reflux in the form of influxes of naturally abundant amino acids and metabolites resulting from the degradation of biomass components. As suggested by the anomalous isotopomers (**Fig. 3**, Supplementary **Fig. S1**, Supplementary **Table S1**), the amino acids entering the network are histidine, serine, glycine, leucine, alanine, aspartate/asparagine, threonine, methionine, isoleucine, leucine and glutamate/glutamine. The biomass component degradation products entering the model are fructose-6-phosphate (degradation product of sugars), cytosolic and plastidic pyruvate (degradation of serine, alanine, leucine and glutamate via citramalate pathway[34]), 3-phosphoglycerate (degradation of serine and glycine), succinate (degradation of valine) and ACA (degradation of fatty acids and acetate/propanoate resulting from the degradation of glutamate, threonine, valine and tyrosine). These influxes are marked in red in **Fig. 6**. The complete metabolic model incorporating metabolite reflux, hereafter designated “AA_in”, is listed in Supplementary **Table S7**.

We constructed this metabolic network model iteratively. Beginning with a simpler network, we attempted various refinements to the network and accepted those refinements that better accounted for the isotopomer abundance measurements (mathematically, those refinements that lowered the SSR of the model; Eq. [1]). In particular, a single-compartment model did not yield a satisfactory SSR, and introducing cytosol-versus-plastid compartmentalization lowered the SSR. Similarly, we introduced the amino acid influxes iteratively into successive versions of the model. We allowed the flux evaluation program NMR2Flux+ to find, through a simulated annealing global optimization algorithm, values of the metabolic pathway fluxes as well as the influxes, so as to result in the lowest possible SSR. We constrained the amino acid influxes to a lower bound of zero and an upper bound of the order of 12% of their rate of incorporation into biomass. We calculated the upper bound from measurements of (i) the growth rate (data not shown), (ii) the ratio of the initial biomass to newly synthesized biomass (1:9) as evaluated from growth rate measurements and (iii) the biomass composition (data not shown). However, the upper bounds for some influxes had to be increased, especially in case of the ACA influx, to achieve a low SSR. The evaluated fluxes and influxes, as well as their standard deviations (estimated as described in Materials and Methods) are listed in Supplementary Tables **S3 and S7**.

We compared the ability of the AA_in model to account for the measured isotopomer abundances with that of previously reported models[e.g. 25] that employed other mechanisms to model the effect of the initial biomass. One previously employed approach is adjustment of the purity (^{13}C enrichment) of the labeled carbon source. Our model “Gluc_dilu” (listed in Supplementary **Table S5**) uses this approach to model the effect of the initial biomass instead of the metabolite influxes described above. Another previous approach is the correction of the measured MIDs by a factor based on the ratio of the initial to final biomass. Our uniform isotopomer correction model “Iso_corr(U)” (listed in Supplementary **Table S4**) uses this approach. Instead of introducing additional metabolite influxes, it assumes that all proteinogenic amino acid isotopomer measurements are uniformly diluted by natural abundance in the ratio 1:9, which is the measured ratio of initial biomass to newly synthesized biomass. An improved version of this model, the variable isotopomer correction model “Iso_corr(V)”, allows different dilution ratios or mixing coefficients for different amino acids. These dilution ratios are parameters in the flux model and are estimated by our flux evaluation program NMR2Flux+ (the evaluated ratios are in Supplementary **Table S6**). This new Iso_corr(V) model is clearly an improvement over the original isotopomer correction model, as it reflects the expectation that

different amino acids may be diluted by naturally abundant intracellular pools to different extents upon introduction of the labeled carbon source.

To evaluate the performance of the models, we calculated an SSR based on a set of 405 isotopomer abundances of all the 30 amino acid fragments detected on GC-MS measured from three ILEs using 98% 1-^{13}C glucose, 28% U-^{13}C glucose and 98% U-^{13}C glucose. As evident from the parity plot **Fig. 7**, the AA_in model fits the isotopomer measurements significantly better than the other three models. Quantitatively, the AA_in model yielded an SSR of 538, whereas the Gluc_dilu model gave an SSR of 4,626, the Iso_corr(U) model gave an SSR of 4,983, and the Iso_corr(V) gave an SSR of 1,748 (**Fig. 7**, inset). Remarkably, SSR corresponding to the 98% U-^{13}C glucose ILE was the lowest in case of the AA_in model, whereas the two alternative models provided substantially poorer fits and thereby larger SSRs. The statistically acceptable SSR value for 405 redundant measurements is 462. The AA_in model comes reasonably close to this, given that it modeled a batch culture at a pseudo-steady state.

Highlights of the flux map evaluated by using the AA_in model (fluxes and statistics shown as blue boxes in **Fig. 8**, Supplementary **Table S3**) include substantial net flux through the plastidic PPP (g6pdhp: 0.43 ± 0.23) and relatively lower flux through the cytosolic PPP (0.16 ± 0.07), as well as relatively low anaplerotic cycling (ana_net: 0.04 ± 0.01). To assess the efficacy of the metabolite reflux model, we also performed a second set of three parallel ILEs on 98% 1-^{13}C, 28% U-^{13}C and 98% U-^{13}C glucose on poplar cell suspensions over a 21-d period, by seeding the cells in fresh, appropriately ^{13}C-labeled medium every 7 d. These ILEs, although excessively long, can be expected to be relatively reflux-free because of (i) their length and (ii) the fact that intracellular metabolites will be less affected by reflux and unaffected by mixing (**Fig. 1**) as compared to biomass components such as proteinogenic amino acids. We isolated and measuring isotopomer abundances of intracellular metabolites from this experiment, and evaluated fluxes using a metabolic model that lacked metabolite reflux (except the ACA influx), glucose dilution or isotopomer correction. The resulting flux evaluation yielded a statistically acceptable SSR of 382 for 357 isotopomer abundances, and the corresponding fluxes and their statistics are depicted in **Fig. 8** (red boxes). Clearly, the results of both methodologies are qualitatively the same, although there are numerical differences between them for some flux values, especially the CO_2 influx (CO_2in, reflux model: 0.18 ± 0.16, model without reflux: 0.03 ± 0.01) and the net flux through the plastidic PPP (g6pdhp, reflux model: 0.43 ± 0.23, model without reflux: 1.50 ± 0.11). Despite this, the estimates of many fluxes (CO_2out, pgif, pgibp, pdhp, tktBb, sdhr, mdhr, ana_net) are similar.

Discussion

Isotope MFA is a powerful systems biology tool for studying cellular metabolism. It relies on the measurement of labeling patterns of metabolites (isotopomers) generated by the metabolic processing of specific ^{13}C-labeled carbon sources, usually at isotopic steady state. In conjunction with information on carbon atom rearrangements in metabolic reactions, this methodology quantitatively tracks the trajectories of the ^{12}C and ^{13}C atoms through the metabolic network to simulate isotopomer abundances. By fitting these simulated isotopomer abundances to the experimentally measured ones, it evaluates flux values and distributions that best explain the labeling data. Consequently, the flux distributions estimated by isotope MFA are very sensitive to the isotopomer abundances of metabolites. In batch ILEs, an important biological phenomenon that could influence the isotopomer abundance measurements is the presence of the initial biomass and interactions between this and the newly synthesized biomass (Fig. 1). Because batch ILEs are convenient or inevitable for many biological systems,

it is necessary to quantify the effects of the initial biomass on the isotopomer abundance measurements.

In this work, we show that the use of ≈100% U-^{13}C carbon sources can provide much information on the interactions between the initial and newly synthesized biomass. An ILE using a single U-^{13}C carbon source will typically yield no flux information, as all isotopomers will be ^{13}C-labeled. However, reflux of the naturally abundant initial biomass into metabolism is amplified in such an ILE. The anomalous ^{13}C enrichments (Fig. 2) and mass isotopomer abundances (**Fig. 3**, Supplementary **Fig. S1**) from our 98% U-^{13}C glucose ILEs suggested that the initial biomass did not remain isolated from the newly synthesized biomass and that it refluxed into the metabolic network. If the initial biomass had remained isolated and had only mixed with the newly synthesized biomass, we would only have observed the least massive (*M*+0) and most massive (*M*+*n*) isotopomers of each fragment. However, isotopomer measurements from the 98% U-^{13}C ILE clearly disagreed with this expectation. We eliminated two competing causes for these isotope labeling trends – photosynthetic and anaplerotic $^{12}CO_2$ incorporation – by performing follow-up ILEs (**Fig. 4**) or examining mass isotopomers of aspartate and glutamate (**Fig. 5**). Although these results were not able to exclude the "mixing" phenomenon, they did point at reflux of the initial biomass as the major cause of the anomalous isotope labeling patterns.

Previous reports[12,13,6,24] of isotope MFA in batch cultures have used two approaches to account for the initial biomass: (i) adjustment of the ^{13}C enrichment of the labeled carbon source, or (ii) correcting the isotopomer abundance measurements assuming that they are produced by mixing of the initial and newly synthesized biomass. Only one report[26] used amino acid influxes to model the effect of initial biomass. However, there is no previous report quantifying the extents to which these models account for isotopomer measurements. This is a pressing need, given our 98% U-^{13}C glucose ILEs that are sensitive to metabolite reflux. Our comparison of all these models clearly showed that the model incorporating amino acid and metabolite reflux (AA_in; **Fig. 6**) is more accurate than the other three models. It especially outperforms them in explaining the data from the 98% U-^{13}C glucose ILE (**Fig. 7**). Two reasons may explain the superior performance of the metabolite reflux model. Foremost, the glucose dilution and isotopomer correction models only use a single dilution factor to account for anomalous isotopomers. In contrast, the reflux model uses several amino acid and metabolite influxes as flux parameters. In other words, we are introducing additional degrees of freedom at reasonable places of our metabolic network. These additional degrees of freedom can explain the anomalous isotopomers observed by us. For example, the reflux model estimated the acetyl-CoA reflux (~0.02 units) to be much higher than histidine reflux (~0.0001 units), indicating that the reflux of acetyl-CoA is much more responsible for the anomalous mass isotopomers than the reflux of His. Nevertheless, increased degrees of freedom are not the sole explanation for the anomalous isotopomer measurements. Specifically, the improved isotopomer correction model Iso_corr(V) does have a large number of degrees of freedom, just as the reflux model. Despite this, the Iso_corr(V) model failed to acceptably fit the labeling data, indicating that the dilution or mixing effect is insufficient to account for the anomalous isotopomers. The fact that the AA_in model is able to acceptably fit the labeling data, including most of the anomalous isotopomers, emphasizes the evidence for reflux. Moreover, our observation of different ^{13}C enrichments for amino acid fragments sharing a common metabolic precursor (e.g. 27.1%, 27.4%, 27.8% for glutamate[2345] fragments and 22.7%, 23.0% for proline[2345] fragments despite their both being derived from the same metabolic precursor AKG) can only be explained by a comprehensive reflux model.

Being tree cells, poplar cell suspensions are especially appropriate for this study because trees are known to be capable of conserving and reusing nutrients over small and large timescales.

Previous studies show that poplar efficiently manages its nitrogen reserves by recycling nitrogen from leaves during senescence in fall, storing this recycled nitrogen in the form of a protein (bark storage protein [BSP]) during winter and remobilizing this stored nitrogen for *de novo* protein synthesis in spring[35,36]. Exploring how poplar cells' metabolic landscapes are altered by this nitrogen cycling and storage process is a fascinating research question and metabolic flux analysis holds promise in shedding light into this question. Therefore, our study of biomass reflux model in poplar suspension cells serves as a forerunner of studying the larger problem of nitrogen recycling mechanism in poplar.

The flux distribution obtained by using isotopomer measurements of proteinogenic amino acids from a 7-d ILE and the reflux model agreed qualitatively with the flux distribution obtained using intracellular metabolites by an independent methodology – 21-d ILEs and a model with minimal reflux (**Fig. 8**). Although there were numerical differences between flux values estimated through the two methodologies, many flux values (e.g. ACA reflux, net anaplerotic flux) and all flux trends (e.g. large cyclic PPP in the plastid) were identical for both methodologies. This consistency justifies the application of reflux methodology in analyzing anomalous labeling patterns of proteinogenic amino acids.

Conclusion and future outlook

The development and isotope MFA in plants and mammalian cell or tissue cultures, combined with the inevitability of batch culture, necessitates accurate modeling to account for the initial biomass. In this regard, our work addresses two important gaps in knowledge: the use of an ILE supplying a ≈100% U-^{13}C carbon source to accurately quantify the extent of reflux, and the demonstration that a comprehensive reflux model accounts for the isotopomer abundances from such an ILE much more satisfactorily than previously employed models. Furthermore, flux results from our methodology using the reflux model were similar to those from another methodology involving minimal reflux. We anticipate our methodology to be especially useful for isotope MFA of batch cultures of bacterial, plant or mammalian cell or tissue where it is not practical to remove the initial biomass.

Future extensions to our work include targeted ILEs to improve estimates of reflux, especially of the biomass components that reflux to the greatest extent. Additionally, investigation and explicit incorporation of more nontrivial amino acid degradation pathways similar to the valine ≫ succinate and glutamate ≫ ACA pathways in our model is also necessary.

Finally, the reflux model discussed in this study is an isotopic steady-state model. However, a steady-state model of metabolite reflux is not exact as reflux of the initial biomass in a batch culture is strictly an instationary process. For example, the assumption that refluxed biomass components are naturally abundant throughout the ILE is only a first approximation. In reality, the aggregate (sum of initial and newly synthesized) biomass gets more and more ^{13}C-enriched during the course of the ILE; correspondingly, the refluxed biomass components get gradually enriched during the ILE. Here we used an isotopic steady state model as an approximation of dynamic MFA to simulate the reflux process and demonstrated its ability to explain the anomalous isotopomer measurements. Nevertheless, it is still desirable to incorporate the metabolite reflux model into instationary isotope MFA[8,10,24]. Instationary isotope MFA is a technique that involves measuring the instationary transients of the labeling states of the metabolites and solving ordinary differential equations instead of algebraic equations to simulate the corresponding isotopomers. Although this is experimentally and computationally more challenging than steady-state modeling, it holds much promise in producing reliable models of metabolite reflux. One challenge is that in instationary MFA metabolite pool sizes need to be considered. They can either be measured or be treated as parameters to estimate in the optimization procedure. However, usually pool size measurements are not very accurate,

especially for intermediate metabolites with low concentrations. In this study, we proved the amino acid reflux phenomenon through mass isotopomer measurements. By integrating reflux concept in instationary MFA, we should expect more accurate pool size estimations as well as intracellular flux estimations. We anticipate the future studies of metabolite reflux in isotope MFA will incorporate these extensions.

Author contributions

SN, AM, XZ, GDC and GS conceived this study; SN, AM, XZ and GS designed it. GDC provided poplar cell suspension cultures; SN, AM and XZ performed experiments; SN, AM, XZ and GS analyzed data and drafted the manuscript. XZ and GS revised the manuscript after peer review. SN, AM, XZ contributed equally; their names appear in the author list in the chronological order of their joining this project.

Acknowledgments

The authors thank Margaret N. Simons and Whitney Hollinshead for assistance with experiments. This work was funded by the U. S. National Science Foundation (award IOS-0922650) award to GDC and GS.

References

1. W. S. Ahn and M. R. Antoniewicz, *Metabolic Engineering*, 2011, **13**, 598–609.
2. N. J. Kruger and R. G. Ratcliffe, *Biochimie*, 2009, **91**, 697–702.
3. R. W. Leighty and M. R. Antoniewicz, *Metabolic Engineering*, 2012, **14**, 533–541.
4. J. Schwender, *Current Opinion in Biotechnology*, 2008, **19**, 131–137.
5. J. Schwender, F. Goffman, J. B. Ohlrogge, and Y. Shachar-Hill, *Nature*, 2004, **432**, 779–782.
6. G. Sriram, D. B. Fulton, and J. V. Shanks, *Phytochemistry*, 2007, **68**, 2243–2257.
7. W. Wiechert, *Metabolic Engineering*, 2001, **3**, 195–206.
8. W. Wiechert and K. Noh, *Advances in Biochemical Engineering/Biotechnology*, 2005, **92**, 145–172.
9. W. Wiechert and K. Nöh, *Current Opinion in Biotechnology*, in press, **0**, 0.
10. K. Nöh and W. Wiechert, *Biotechnology and Bioengineering*, 2006, **94**, 234–251.
11. T. Szyperski, *European Journal of Biochemistry*, 1995, **232**, 433–448.
12. M. Dauner, J. E. Bailey, and U. Sauer, *Biotechnology and Bioengineering*, 2001, **76**, 144–156.
13. G. Sriram, D. B. Fulton, V. V. Iyer, J. M. Peterson, R. Zhou, M. E. Westgate, M. H. Spalding, and J. V. Shanks, *Plant Physiology*, 2004, **136**, 3043–3057.
14. G. Sriram, L. Rahib, J.-S. He, A. E. Campos, L. S. Parr, J. C. Liao, and K. M. Dipple, *Molecular Genetics and Metabolism*, 2008, **93**, 145–159.
15. M. I. Klapa, J. C. Aon, and G. Stephanopoulos, *Biotechniques*, 2003, **34**, 832–6, 838, 840 passim.
16. G. Sriram, V. V. Iyer, D. B. Fulton, and J. V. Shanks, *Metabolic Engineering*, 2007, **9**, 442–451.
17. T. Shen, B. Rui, H. Zhou, X. Zhang, Y. Yi, H. Wen, H. Zheng, J. Wu, and Y. Shi, *Molecular BioSystems*, 2013, **9**, 121.
18. X. Chen, A. P. Alonso, D. K. Allen, J. L. Reed, and Y. Shachar-Hill, *Metabolic Engineering*, 2011, **13**, 38–48.
19. S. K. Masakapalli, N. J. Kruger, and R. G. Ratcliffe, *The Plant Journal*, 2013, **74**, 569–582.
20. S. Nargund and G. Sriram, *Molecular BioSystems*, 2013, **9**, 99.
21. C. M. Metallo, J. L. Walther, and G. Stephanopoulos, *Journal of Biotechnology*, 2009, **144**, 167–174.
22. T. Omasa, K. Furuichi, T. Iemura, Y. Katakura, M. Kishimoto, and K. Suga, *Bioprocess and Biosystems Engineering*, 2009, **33**, 117–125.

23. L.-E. Quek, S. Dietmair, J. O. Krömer, and L. K. Nielsen, *Metabolic Engineering*, 2010, **12**, 161–171.
24. M. Antoniewicz, D. Kraynie, L. Laffend, J. Gonzalezlergier, J. Kelleher, and G. Stephanopoulos, *Metabolic Engineering*, 2007, **9**, 277–292.
25. J. Lonien and J. Schwender, *Plant Physiology*, 2009, **151**, 1617–1634.
26. N. J. Kruger, S. K. Masakapalli, and R. G. Ratcliffe, *Journal of Experimental Botany*, 2011, **63**, 2309–2323.
27. P. Sannigrahi, A. J. Ragauskas, and G. A. Tuskan, *Biofuels, Bioproducts and Biorefining*, 2010, **4**, 209–226.
28. O. Fiehn, J. Kopka, R. N. Trethewey, and L. Willmitzer, *Analytical Chemistry*, 2000, **72**, 3573–3580.
29. T. M. Wasylenko and G. Stephanopoulos, *Biotechnology Journal*, 2013, **8**, 1080–1089.
30. N. J. Kruger, J. E. Huddleston, P. Le Lay, N. D. Brown, and R. G. Ratcliffe, *Phytochemistry*, 2007, **68**, 2176–2188.
31. R. Caspi, T. Altman, K. Dreher, C. A. Fulcher, P. Subhraveti, I. M. Keseler, A. Kothari, M. Krummenacker, M. Latendresse, L. A. Mueller, Q. Ong, S. Paley, A. Pujar, A. G. Shearer, M. Travers, D. Weerasinghe, P. Zhang, and P. D. Karp, *Nucleic Acids Research*, 2011.
32. V. V. Iyer, G. Sriram, D. B. Fulton, R. Zhou, M. E. Westgate, and J. V. Shanks, *Plant, Cell & Environment*, 2008, **31**, 506–517.
33. S. K. Masakapalli, P. L. Lay, J. E. Huddleston, N. L. Pollock, N. J. Kruger, and R. G. Ratcliffe, *Plant Physiology*, 2010, **152**, 602–609.
34. B. Wu, B. Zhang, X. Feng, J. R. Rubens, R. Huang, L. M. Hicks, H. B. Pakrasi, and Y. J. Tang, *Microbiology*, 2009, **156**, 596–602.
35. B. Zhu and G. D. Coleman, *Plant Physiology*, 2001, **126**, 342–351.
36. F. R. Cantón, M. F. Suárez, and F. M. Cánovas, *Photosynthesis Research*, 2005, **83**, 265–278.
37. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, Cambridge University Press, 3rd edn., 2007.

Accepted for publication in Molecular Biosystems, doi: 10.1039/C3MB70348G

Figures and captions

Fig. 1

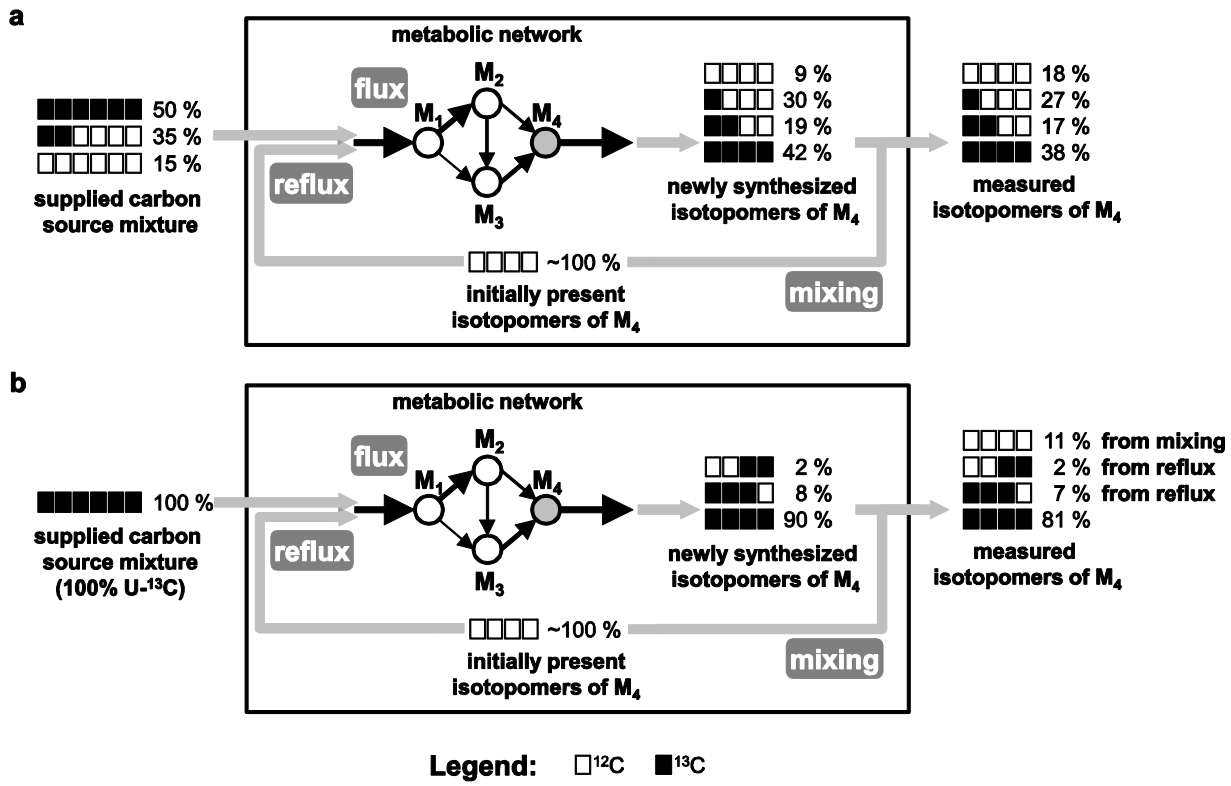


Fig. 1. Metabolic flux and reflux in isotope labeling experiments (ILEs). (a) In ILEs, a combination of isotopically labeled (^{13}C or ^{12}C) carbon sources is fed to a metabolic network. The metabolite molecules synthesized by the network from the ^{13}C or ^{12}C atoms of the carbon sources will consist of isotopically distinct isomers (isotopomers). Information on intracellular fluxes that is concealed in the isotopomer patterns and abundances can be unraveled by mathematical modeling. In batch ILEs, the biomass present at the beginning of the ILE may interfere with the ILE. This initial biomass, labeled at the natural abundance of ^{13}C (~1.1%), may either mix with the newly synthesized isotopomers ("mixing" arrow) or reflux into the metabolic network ("reflux" arrow). The isotopomers of a hypothetical four-carbon metabolite M_4 illustrate these interactions. **(b)** An ILE supplying a single, ~100% U-^{13}C carbon source will produce isotopomers that clearly distinguish mixing from reflux, as well as quantify the extent of these interactions. In such an ILE, biomass that is newly synthesized from the supplied carbon source will be represented by U-^{13}C isotopomers. Mixing of the initial biomass with the newly synthesized biomass will result in isotopomers that are almost entirely ^{12}C, whereas reflux of the initial biomass will result in isotopomers that are predominantly ^{13}C but contain a few ^{12}C atoms.

Small rectangles represent carbon atoms (black: ^{13}C, white: ^{12}C). Small circles represent entire metabolite molecules. The newly synthesized and initially present intracellular isotopomers correspond the metabolite M_4, shown shaded in gray.

Accepted for publication in Molecular Biosystems, doi: 10.1039/C3MB70348G

Fig. 2

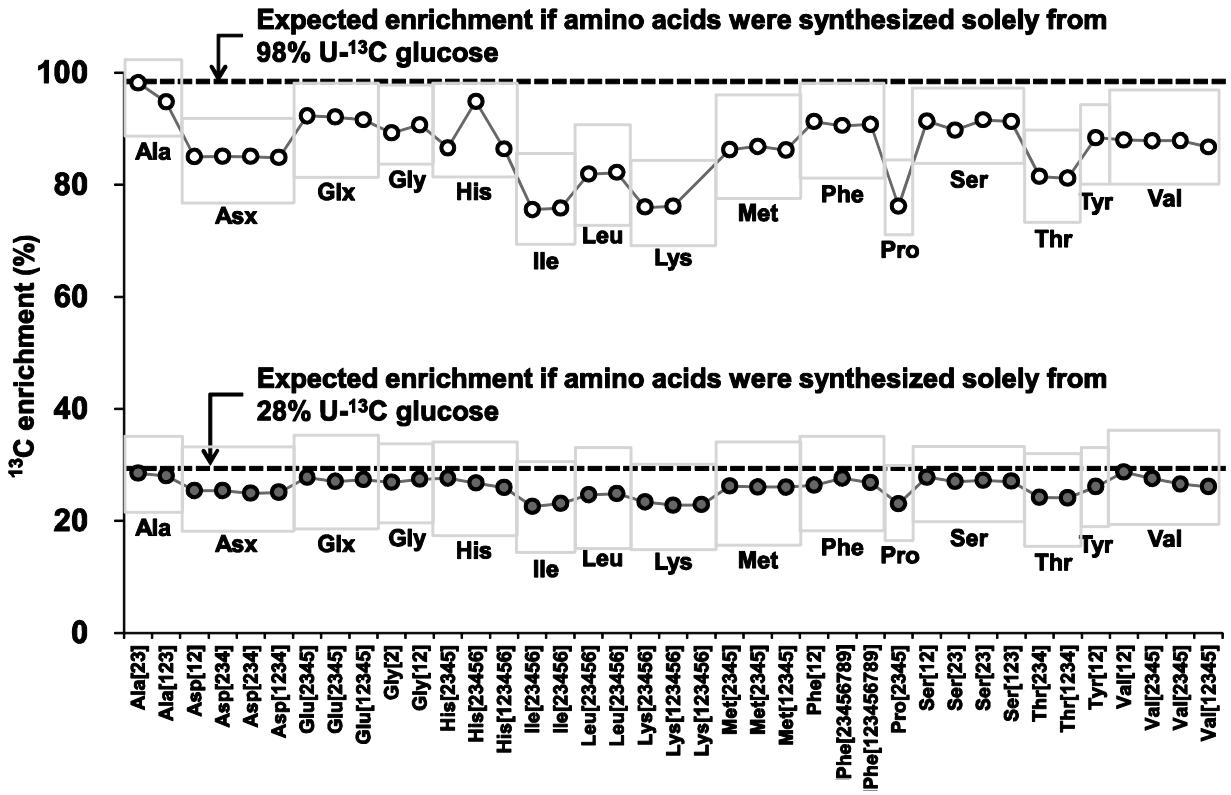


Fig. 2. Proteinogenic amino acid fragments from poplar cell suspensions grown on U-^{13}C glucose are less^{13}C-enriched than the supplied glucose.

For a fragment containing *n* metabolic carbon atoms, the ^{13}C enrichment is defined as:

$$^{13}C\ enrichment = \frac{0\times A_{m+0} + 1\times A_{m+1} + \cdots + n\times A_{m+n}}{n}$$

where A_{m+i} is the (relative) abundance of a mass isotopomer in which *i* of the *n* carbon atoms are ^{13}C. Thus, the ^{13}C enrichment simply represents the fraction of atoms in a fragment that are ^{13}C. The ^{13}C enrichments of various proteinogenic amino acid mass spectral fragments (horizontal axis) were calculated from the mass isotopomer distributions of the fragments. Fragments belonging to the same amino acid are grouped within boxes. Two ILEs were performed, using 98% U-^{13}C glucose (white circles) or 28% U-^{13}C glucose (gray circles) as the sole organic carbon source. The dashed lines denote the ^{13}C enrichments expected for each ILE if the amino acids were entirely synthesized from the supplied glucose. All data points have error bars, although most error bars are too small to be visible.

Accepted for publication in Molecular Biosystems, doi: 10.1039/C3MB70348G

Fig. 3

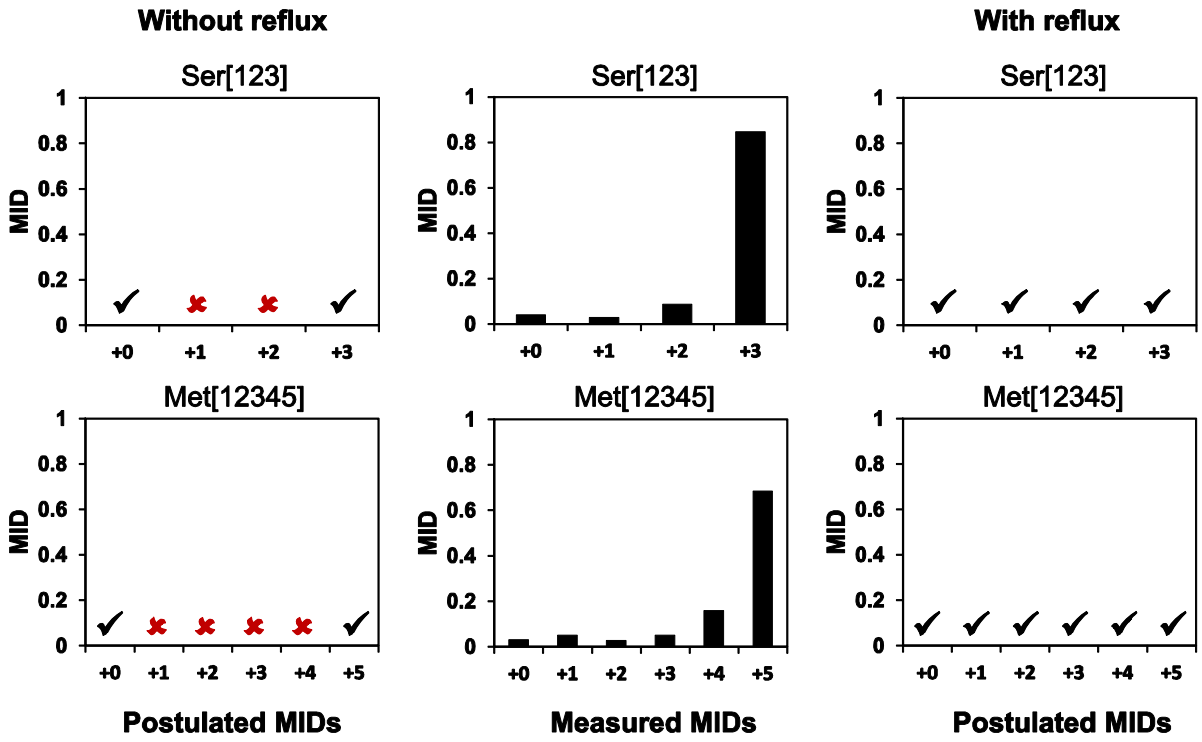
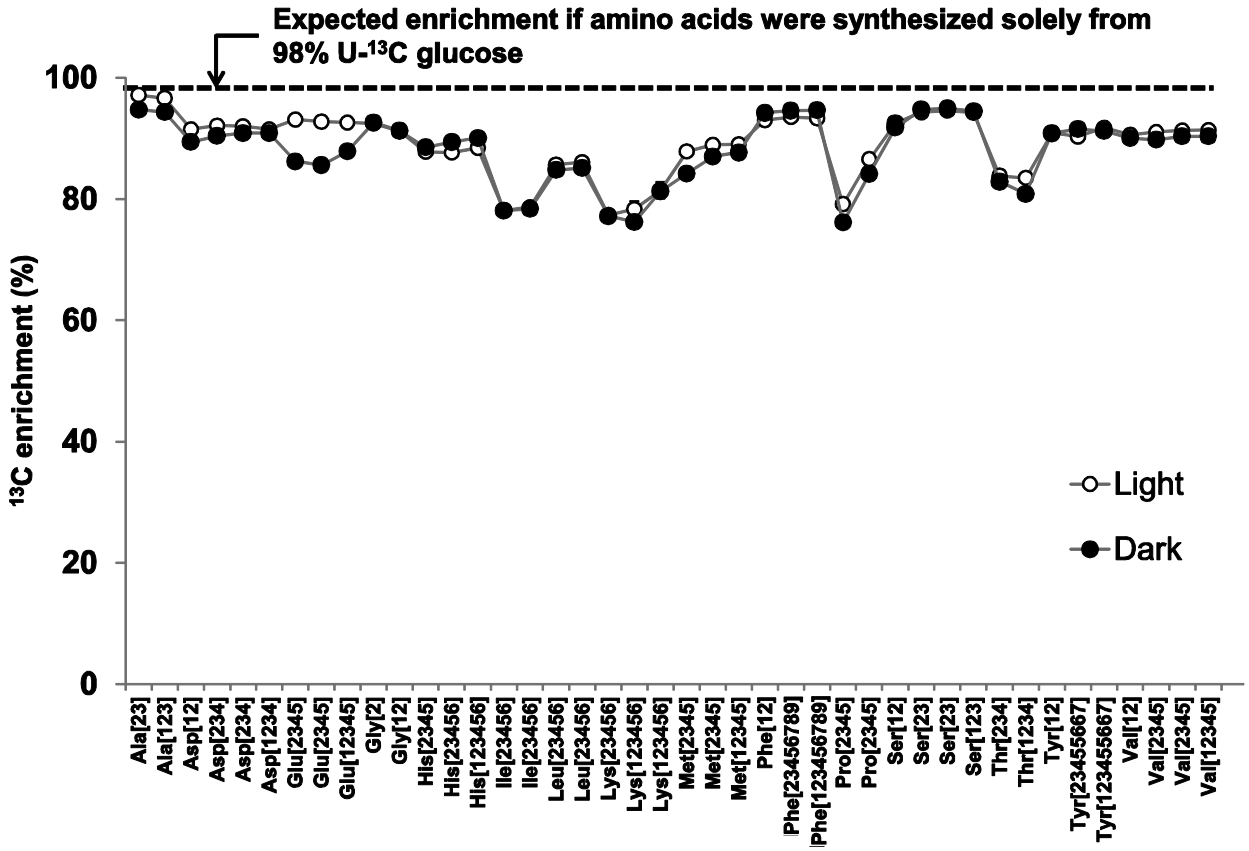


Fig. 3. Metabolites from a 100% U-^{13}C ILE contain anomalous mass isotopomers not explained by mixing of initial and newly synthesized biomass. This figure depicts mass isotopomer abundances (MIDs) of two illustrative fragments Ser[123] and His[123456] (more fragments are shown in Supplementary **Fig. S1**). The MIDs on the **left** were postulated by assuming mixing between the initial biomass and the newly synthesized biomass (at a ratio of 1:9 obtained from dry weight measurements) but no metabolite reflux. In this case, the initial biomass is naturally abundant (lowest mass isotopomers) and the newly synthesized biomass is fully ^{13}C-labeled (highest mass isotopomers). The MIDs in the **center** were experimentally measured in a 98% U-^{13}C glucose ILE. The MIDs on the **right** were postulated assuming metabolite reflux (**right**). The mass isotopomers between the lowest and highest masses are anomalous and cannot be explained by mixing between the initial and newly synthesized biomass.

The black "✓" marks indicate isotopomers postulated to be feasible, whereas the red "×" marks indicate isotopomers postulated to be infeasible under the relevant scenario. All data points have error bars, although most error bars are too small to be visible.

Accepted for publication in Molecular Biosystems, doi: 10.1039/C3MB70348G

662 **Fig. 4**



663

664 **Fig. 4. Photosynthetic $^{12}CO_2$ assimilation does not explain anomalous isotopomers.** The
665 ^{13}C enrichments (vertical axis) of various proteinogenic amino acid mass spectral fragments
666 (horizontal axis) were calculated from the mass isotopomer distributions of the fragments. The
667 dashed lines denote the ^{13}C enrichments expected for each ILE if the amino acids were entirely
668 synthesized from the supplied glucose.

669 Photosynthetic assimilation of atmospheric $^{12}CO_2$ may be one possible explanation for the
670 anomalous isotopomers observed in **Fig. 3**. To test this, we measured ^{13}C enrichments of
671 proteinogenic amino acid fragments between poplar cell suspensions grown on 98% U-^{13}C
672 glucose in continuous light (white circles) and continuous dark (black circles). The similarity of
673 ^{13}C enrichments between light and dark conditions indicates that photosynthetic carbon
674 assimilation did not occur; therefore photosynthesis cannot explain the anomalous isotopomers.
675 All data points have error bars, although most error bars are too small to be visible.

Fig. 5

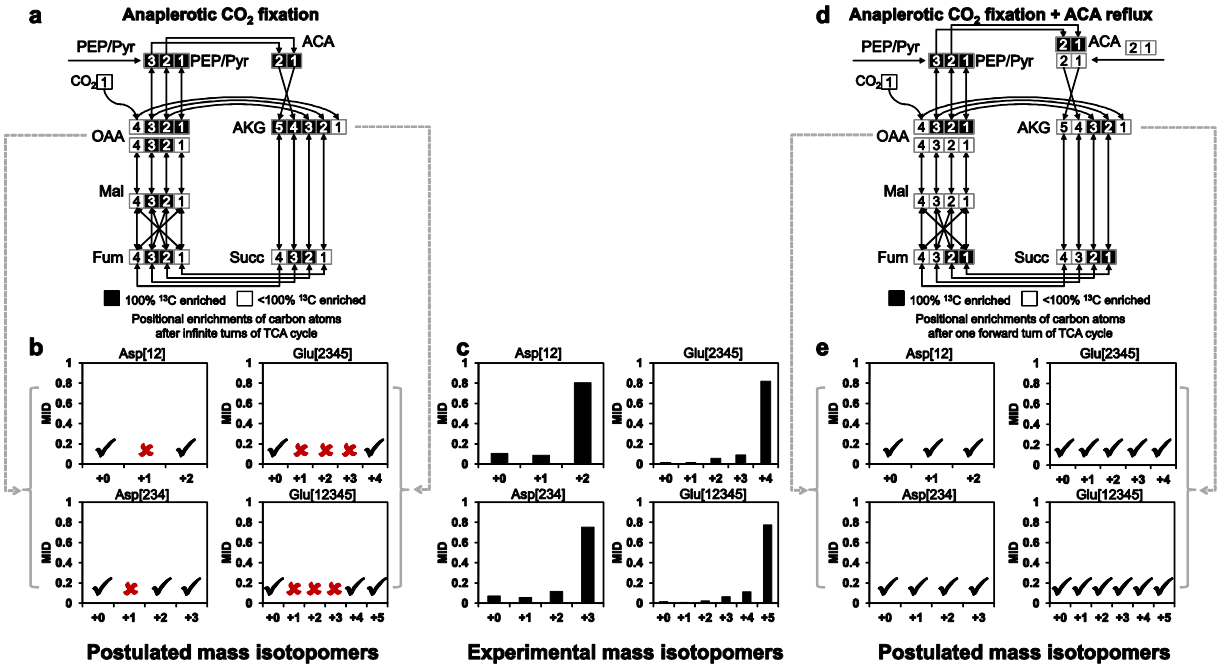


Fig. 5. Anomalous isotopomers of glutamate (Glu) and aspartate (Asp) are not fully explained by anaplerotic $^{12}CO_2$ assimilation, but are explained by acetyl-CoA (ACA) reflux. One possible explanation for the anomalous isotopomers in **Fig. 3** is that $^{12}CO_2$ was assimilated by anaplerotic reactions that interconvert three-carbon glycolytic metabolites (e.g. phosphoenolpyruvate or pyruvate) and four-carbon TCA cycle metabolites (e.g. oxaloacetate or malate). However, this possibility cannot explain all the anomalous isotopomers of Glu (derived from the metabolic precursor α-ketoglutarate [AKG]) and Asp (derived from the precursor oxaloacetate [OAA]). For example, in the 100% U-^{13}C ILE, anaplerotic $^{12}CO_2$ fixation **(a)** will result in AKG with ^{12}C on C-1 and OAA with ^{12}C on C-1 and/or C-4, even after infinite turns of the TCA cycle. This will result in mass isotopomer distributions of Glu and Asp **(b)** lacking containing only the isotopomers indicated with a black "✓" and lacking the mass isotopomers indicated with a red "×". The isotopomer measurements from the 100% U-^{13}C ILE **(c)** do not agree with this prediction. An alternative model **(d)** featuring anaplerotic $^{12}CO_2$ fixation and reflux of acetyl-CoA (ACA, obtainable from the degradation of several amino acids) predicts isotopomer abundances **(e)** consistent with measurements. Small rectangles represent carbon atoms (black: ^{13}C, white: ^{12}C). On the horizontal axis, the numbers "+0", "+1", etc. represent *M*+0, *M*+1, etc. mass isotopomers. All data points have error bars, although most error bars are too small to be visible.

Fig. 6

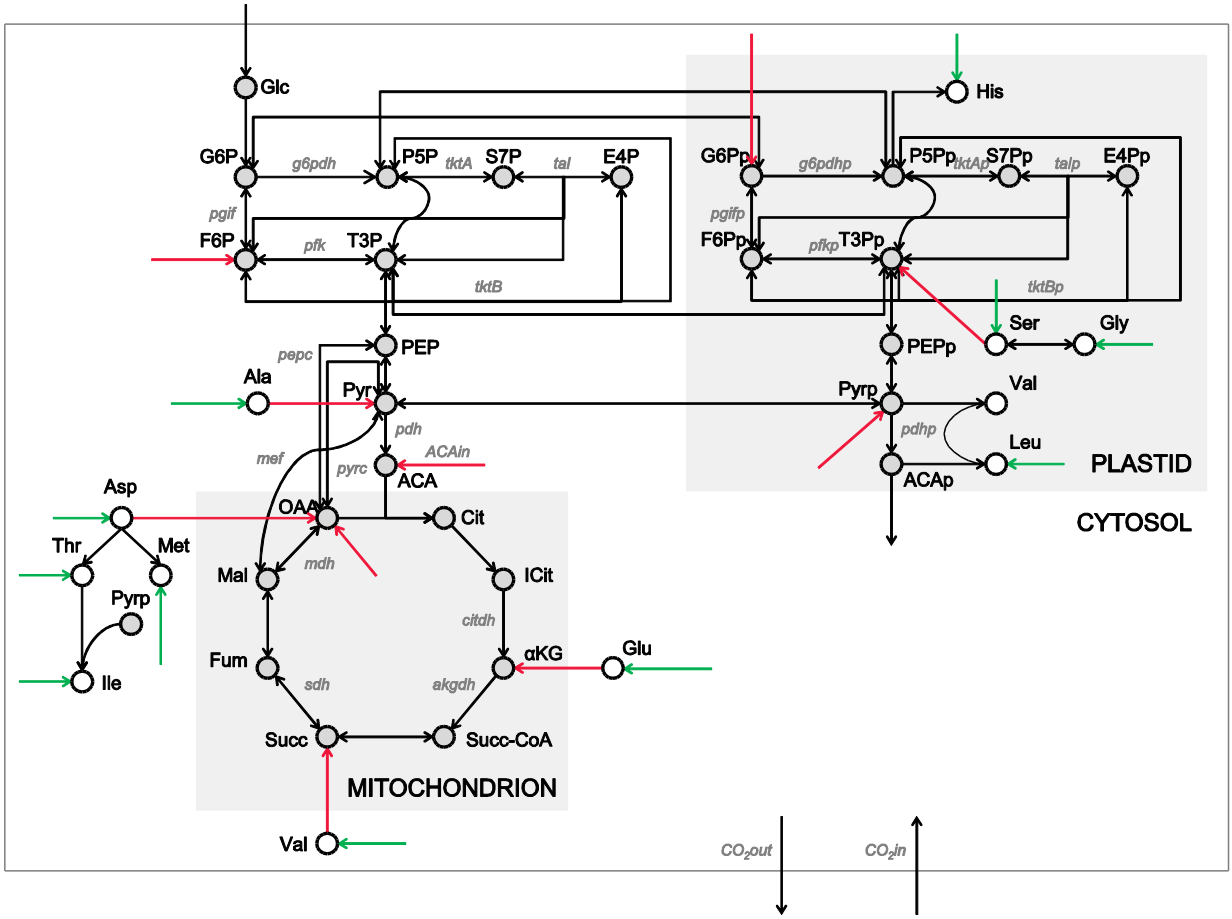
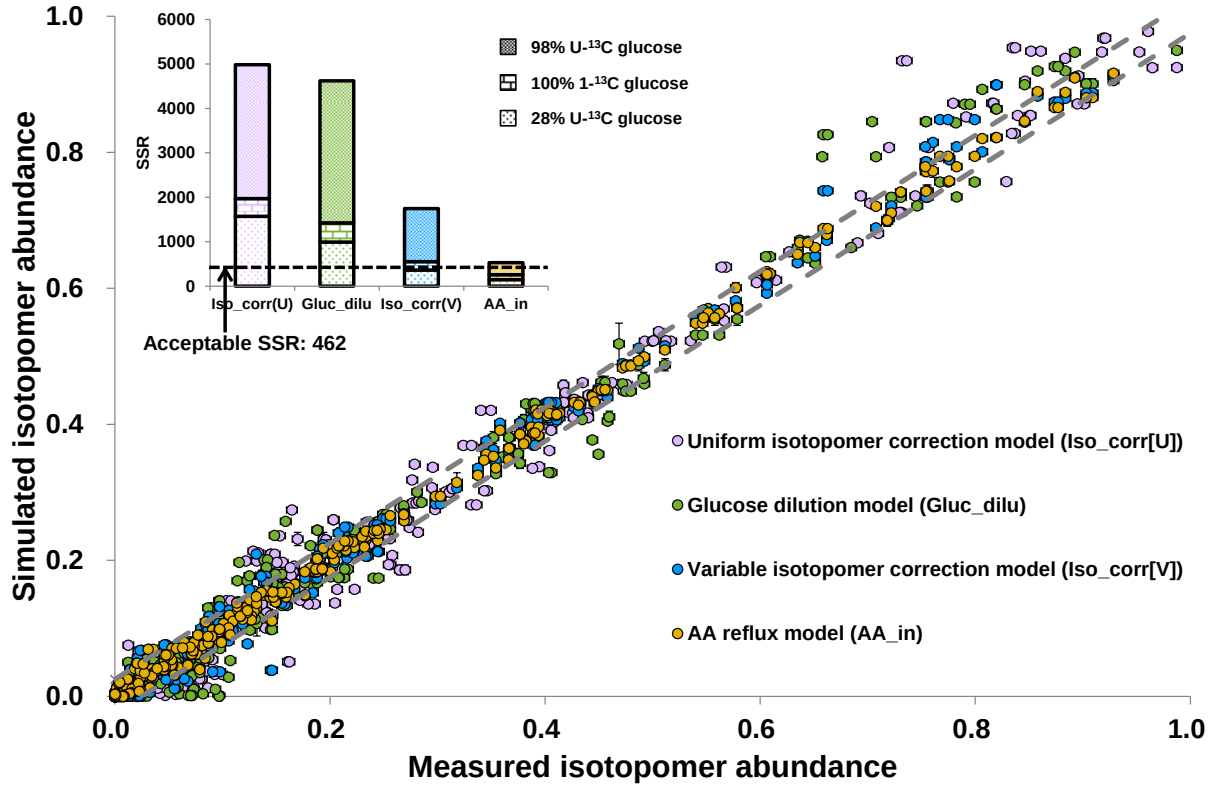


Fig. 6. Metabolic network model incorporating amino acid reflux. This three-compartment metabolic network is a model of primary metabolism in poplar cell suspension. It includes glycolysis and PPP in the cytosol and the plastid, TCA cycle in the mitochondrion as well as amino acids synthesis and reflux reactions. Gray circles represent intracellular metabolites and white circles represent amino acids derived from them. Arrows indicate biochemical reactions, with green arrows representing amino acid "dilution" or "mixing" and red arrows representing amino acid "reflux". Names of significant or interesting reactions (as designated by us) are shown adjacent to their arrows. The stoichiometries, carbon atom rearrangements and fluxes evaluated for all reactions are listed in Supplementary **Table S6** and **S7**. Standard three-letter abbreviations are used for amino acids. Abbreviations for other metabolite names: ACA, acetyl-CoA; AKG, α-ketoglutarate; Cit, citrate; E4P, erythrose-4-phosphate; F6P, fructose-6-phosphate; Fum, fumarate; G6P, glucose-6-phosphate; Glc, glucose; ICit, isocitrate; Mal, malate; OAA, oxaloacetate; P5P, pentose-5-phosphate (representing ribose-5-phosphate, ribulose-5-phosphate and xylulose-5-phosphate); PEP, phosphoenolpyruvate; Pyr, pyruvate; S7P, sedoheptulose-7-phosphate; Succ, succinate; Succ-CoA, succinyl-CoA; T3P, triose-3-phosphate (representing glyceraldehyde-3-phosphate, dihydroxyacetone phosphate and 3-phosphoglycerate). For metabolites duplicated in more than one compartment, compartmentalization is indicated by subscripts: c, cytosol; p, plastid.

Accepted for publication in Molecular Biosystems, doi: 10.1039/C3MB70348G

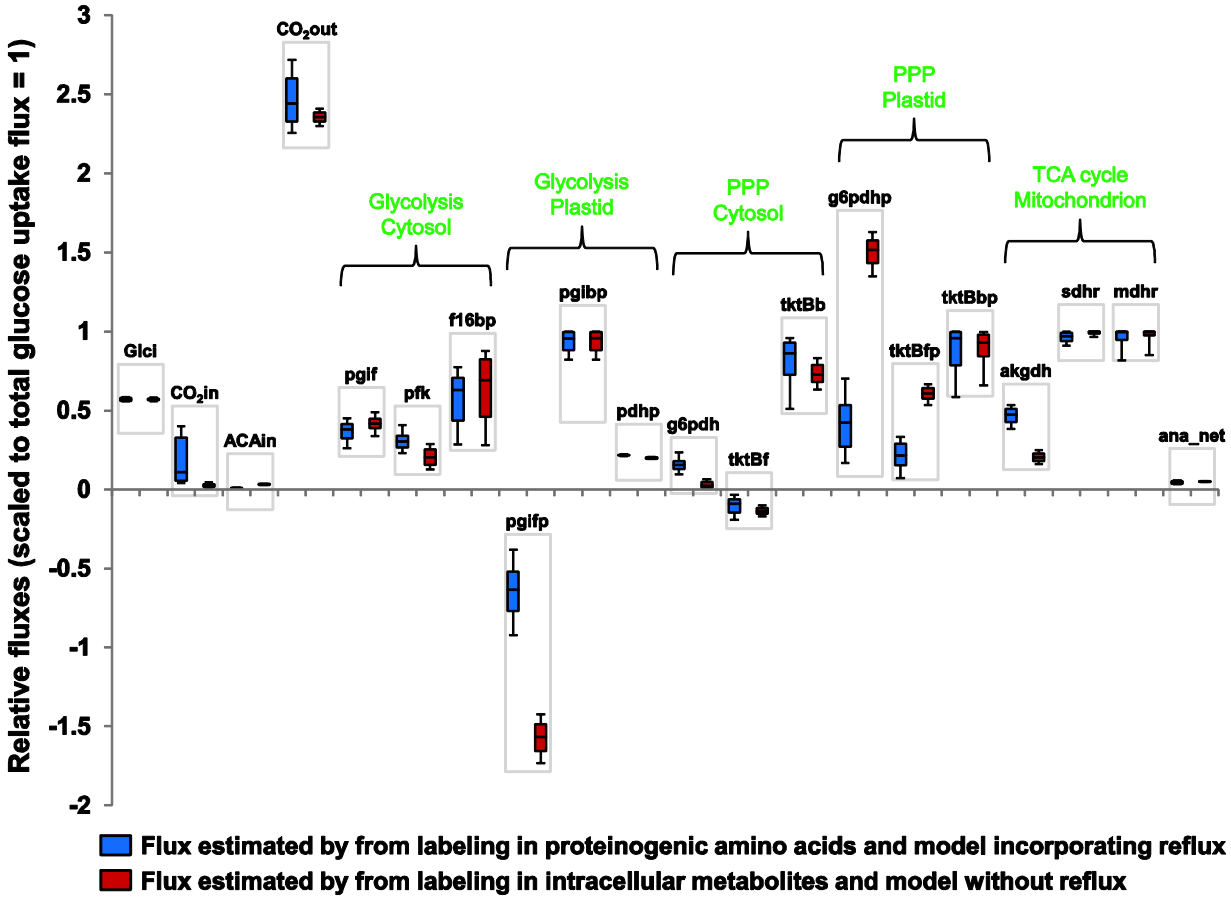
716 **Fig. 7**



717

718 **Fig. 7. Metabolic model incorporating amino acid reflux outperforms models**
719 **incorporating glucose dilution or isotopomer correction in explaining isotopomer**
720 **abundances measured from three ILEs.** In the parity (scatter) plot, 405 isotopomer
721 abundances measured from three ILEs (100% 1-^{13}C glucose, 28% U-^{13}C glucose and 98% U-
722 ^{13}C glucose) are plotted against isotopomer abundances simulated on the basis of the
723 estimated fluxes. Each ILE was performed in triplicate on batch cultures of poplar cell
724 suspensions. This plot illustrates the ability of four separate metabolic models respectively
725 employing uniform isotopomer correction (Iso_corr[U], lavender points and bars), glucose
726 dilution (Gluc_dilu, green points and bars), variable isotopomer correction (Iso_corr[V], blue
727 points and bars) and amino acid reflux (AA_in, gold points and bars) to model the batch culture.
728 The $y = x$ diagonal line corresponds to a perfect fit. The histogram in the inset depicts the sums
729 of squared residues (SSRs) for the three models. All data points have error bars, although most
730 error bars are too small to be visible.

731 **Fig. 8**



732

733 **Fig. 8. Comparison of metabolic fluxes estimated from two independent experiments and**
734 **models.** Fluxes shown here are identified by name in the metabolic network of **Fig. 7** and
735 Supplementary **Table S7**. Fluxes were estimated (i) from a one-week ILE by measuring labeling
736 in proteinogenic amino acids and employing a model incorporating amino acid reflux (blue
737 boxes) as well as (ii) from a three-week ILE by measuring labeling in intracellular metabolites
738 and employing a model without amino acid reflux (red boxes).. For each methodology, flux
739 distributions were obtained by performing 280 bootstrap Monte Carlo simulations[37] that
740 accounted for the larger of the measurement and biological errors in the measured isotopomer
741 abundances. For each flux, the central bar in the box denotes the median of the flux distribution,
742 the lower and upper ends of the box represent the 25th and 75th percentiles of the distribution,
743 and the lower and upper whiskers represent the 10th and 90th percentiles of the distributions.
744 Therefore, shorter boxes and whiskers indicate better flux identifiability. Gray boxes group flux
745 distributions of the same reaction obtained from the two methods. The canonical pathways
746 containing some of the fluxes are indicated above the data points. The flux ana_net is a net
747 anaplerotic CO_2 fixation flux that consolidates the forward and reverse directions of various
748 anaplerotic fluxes in the metabolic model.

Accepted for publication in Molecular Biosystems, doi: 10.1039/C3MB70348G

Supplementary Material

Fig. S1

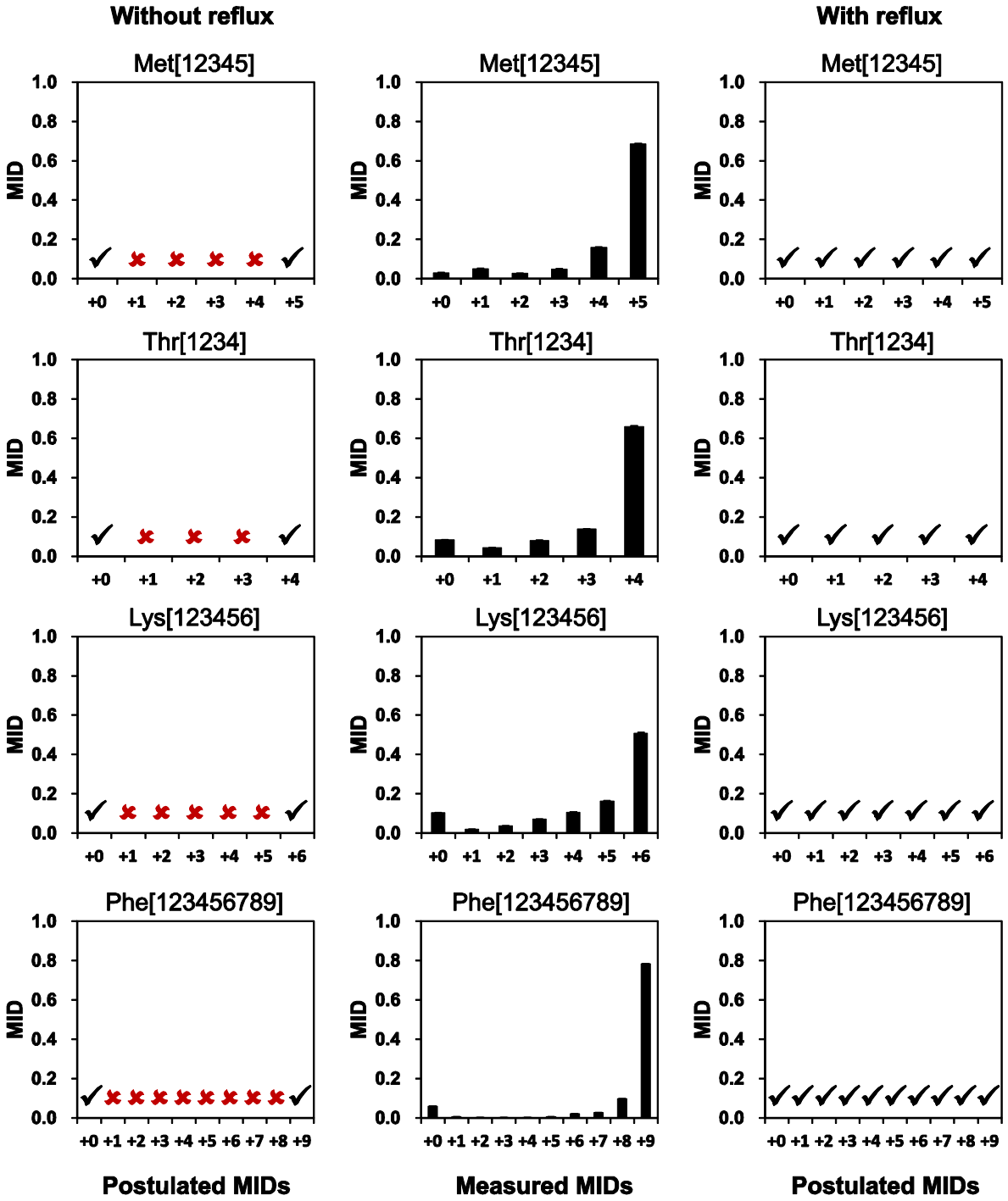


Fig. S1. Metabolites from a 100% U-^{13}C ILE contain anomalous mass isotopomers not explained by mixing of initial and newly synthesized biomass. This is a supplement to **Fig. 3** and depicts mass isotopomer abundances (MIDs) of four illustrative fragments Met[123], Thr[1234], Lys[123456] and Phe[123456789]. The MIDs on the **left** were postulated by

assuming mixing between the initial biomass and the newly synthesized biomass (at a ratio of 1:9 obtained from dry weight measurements) but no metabolite reflux. In this case, the initial biomass is naturally abundant (lowest mass isotopomers) and the newly synthesized biomass is fully ^{13}C-labeled (highest mass isotopomers). The MIDs in the **center** were experimentally measured in a 98% U-^{13}C glucose ILE. The MIDs on the **right** were postulated assuming metabolite reflux (**right**). The mass isotopomers between the lowest and highest masses are anomalous and cannot be explained by mixing between the initial and newly synthesized biomass. The black "✓" marks indicate isotopomers postulated to be feasible, whereas the red "×" marks indicate isotopomers postulated to be infeasible under the relevant scenario. All data points have error bars, although most error bars are too small to be visible.

Supplementary Tables were submitted together as a single PDF file, which appears at the end of this document.

Table S1. Mass isotopomer abundances of proteinogenic amino acids from three ILEs

MIDs of proteinogenic amino acids were measured in three ILEs performed for 7 d on poplar cell suspensions under light. The ILEs used 28% U-^{13}C, 100% 1-^{13}C and 98% U-^{13}C glucose, respectively, as the sole organic carbon source. MIDs are reported as percentages and were obtained by correcting GC-MS spectral data for the natural isotopic abundances of O, H, N, S, P, Si and non-metabolic C atoms[14]. For each ILE, we report individual MID measurements of three biological replicates, an average and a standard deviation (SD). ND, not determined.

Table S2. Mass isotopomer abundances of intracellular metabolites from one-week ILE

MIDs of intracellular metabolites were measured in ILEs performed for 7 d on poplar cell suspensions grown under light and dark. The ILEs used 98% U-^{13}C glucose as the sole organic carbon source. MIDs are reported as percentages and were obtained by correcting GC-MS spectral data for the natural isotopic abundances of O, H, N, S, P, Si and non-metabolic C atoms[14]. For each condition, we report individual MID measurements of three biological replicates, an average and a standard deviation (SD).

Table S3. Fluxes evaluated through the four metabolic models

This table reports averages and standard deviations of flux distributions obtained by performing 280 bootstrap Monte Carlo simulations[37] that accounted for the larger of the measurement and biological errors in the measured isotopomer abundances. The three models employed amino acid reflux (AA_in), glucose dilution (Gluc_dilu) and isotopomer correction (Iso_corr), respectively, to model the batch culture. **Tables S4, S5, S6** and **S7** list stoichiometries and carbon atom rearrangements for the reactions. SD, standard deviation.

Tables S4-S7. Metabolic network models: S4: uniform isotopomer correction model (Iso_corr[U]); S5: glucose dilution model (Gluc_dilu); S6: variable isotopomer correction model (Iso_corr[V]); S7: amino acid reflux model (AA_in)

In these three-compartment models, metabolites with a subscript "p" are plastidic, metabolites with a subscript "m" are mitochondrial and metabolites without a are cytosolic. Carbon atom rearrangements of reactions are indicated by the numbers in parenthesis; these numbers

correspond to carbon atoms numbered according to IUPAC rules. Estimated net flux values, standard deviations and reaction reversibilities are shown for each reaction. For a bidirectional reaction: $A \underset{v_b}{\overset{v_f}{\rightleftharpoons}} B$, where v_f and v_b are the fluxes of the forward and backward reactions, respectively ($v_f \geq v_b$) , the net flux is $v_f - v_b$ and the reaction reversibility is v_b/v_f (expressed as a percentage). SD, standard deviation.