SUPPLEMENTARY TEXT:

ESTIMATION OF METABOLIC FLUXES USING ISOTOPOMER BALANCING AND VALIDITY OF METABOLIC AND ISOTOPIC STEADY STATE ASSUMPTION

Estimation of Metabolic Fluxes Using Isotopomer Balancing

When a substrate is processed through the metabolic pathways of an organism, a variety of reactions take place to generate energy, reducing power and biomass precursors. These reactions may involve carbon-carbon bond breakage or formation. If a mixture of labeled and unlabeled substrate molecules is processed, these bond breakages and formations may lead to unique distributions of isotope isomers (isotopomers) for each intermediate metabolite, which in turn is determined by the *in vivo* fluxes (illustrated in Fig. 4). Therefore, *in vivo* fluxes may be deduced from the labeling pattern of the intermediate metabolites. Since biomass building blocks are formed from precursor metabolites distributed throughout the metabolic pathways of an organism, they serve as storage of the labeling information of the intermediates and hence the metabolic fluxes. Among building blocks, proteinogenic amino acids are commonly used for isotopomer analysis because of two reasons. First, and unlike precursor intermediates, proteinogenic amino acids are abundant and stable. Secondly, since they provide the isotopic labeling information of their central metabolic precursors, their labeling pattern can be used to constrain the central carbon metabolic network model to accurately estimate metabolic fluxes.

The process of estimating metabolic fluxes from isotopomer abundances of proteinogenic amino acids obtained via NMR experiments is illustrated in Supplementary Fig. 1. Cell biomass is harvested once metabolic and isotopic steady state is achieved (see section below for details), NMR samples are prepared, 2D HSQC NMR spectra acquired, the identity of peaks established and the multiplet intensities quantified. These intensities are then used to estimate the fraction of each isotopomer (denoted as isotopomer abundances, see Supplementary Table 2). Since analytical expressions to directly calculate fluxes from isotopomer data do not exist, the computer program NMR2Flux (1) was used to estimate the fluxes. An overview of the algorithm used by this software is also depicted in Supplementary Fig. 1. The software NMR2Flux is supplied with a metabolic network based on known *E. coli* biochemistry (see supplementary Table 1), isotopomer data (see Supplementary Table 2), and flux constraints (based on experimental measurement of extracellular metabolites). Based on the stoichiometry, the program starts with a set of guessed fluxes that satisfies the input constraints and simulates the resulting isotopomer distribution. The program then calculates the error between the simulated and experimental multiplet intensities: i.e. $\chi^2 = \sum_{j=1}^{P} \left[I_j - I_{xj} \right]^2$, where I_j and I_{xj} are the j-th simulated and experimental multiplet intensities, respectively, out of a total of P multiplet intensities. If this error is lower than the desired tolerance, the guessed flux set is accepted, otherwise another set of candidate fluxes is guessed and the process is repeated (via a simulated annealing approach). Standard deviations of the fluxes were computed by statistical analysis of the resulting fluxes from 250 simulation runs. A detailed description of the overall process for estimating metabolic fluxes via isotopomer balancing used in our study can be found in the Supplementary Materials IV and VI in Sriram et. al. (1).

Metabolic and Isotopic Steady State

Metabolic fluxes estimated using isotopomer balancing represent the time-averaged *in vivo* fluxes since the incorporation of labeling at any given time is dependent upon the fluxes at that moment. Therefore, it is important to ensure that the analysis is conducted during a phase of the culture where the metabolic fluxes are constant. The exponential phase of growth during batch cultures provides such a condition as the cells are in a pseudo-steady state: i.e. although the concentrations of substrate, products, and cells are changing with time, the *in vivo* fluxes remain constant (2). This study tested whether the fluxes were indeed constant during the exponential phase.

The flux of a metabolite is given by

$$\upsilon = \frac{1}{x}\frac{dM}{dt} \qquad (a)$$

where υ is the flux (mmol/g CDW/h) , x is the cell density (g CDW/L), M is the metabolite concentration (mmol/L), and t is time (h). The change in cell density due to growth is determined by the specific growth rate, μ (h^{-1}), which is defined as:

$$\mu = \frac{1}{x}\frac{dx}{dt} \qquad (b)$$

By solving equation *(b)* for x and substituting the resulting expression into *(a)*, the following equation is obtained:

$$\frac{dM}{dx} = \frac{\upsilon}{\mu} \qquad (c)$$

Since μ is constant throughout the exponential phase of growth (see Fig. 2), equation *(c)* indicates that, for a constant flux υ, the plot between M and x should be a straight line. The plots shown in supplementary Fig. 2 exhibit a very good fit to a straight line for the extracellular metabolites consumed and generated during the exponential phase by wild type strain MG1655 and mutant Pdh. These results indicate that the extracellular fluxes were indeed constant during exponential growth and thus cultures in this phase can be regarded at pseudo steady state.

The calculation of metabolic fluxes using isotopomer balancing also requires the cellular proteins to be in isotopic steady state. This requirement is met by cultivating cells in metabolic steady state (e.g. exponentially growing cells in a batch culture as discussed above) until the ^{13}C label is fully propagated throughout metabolism (2). In the batch cultures used in our study, cells precultured on unlabeled glucose were inoculated in a medium containing a mixture of labeled and unlabeled glucose (12.5% U-^{13}C, 25% 1-^{13}C, and 62.5% unlabeled). In order to ensure an isotopic steady state, cells were harvested after five generations of exponential growth (i.e. while in metabolic steady state). Although it is widely accepted that five generations are sufficient to achieve isotopic steady state (2), the validity of this criterion under our experimental conditions is verified in what follows.

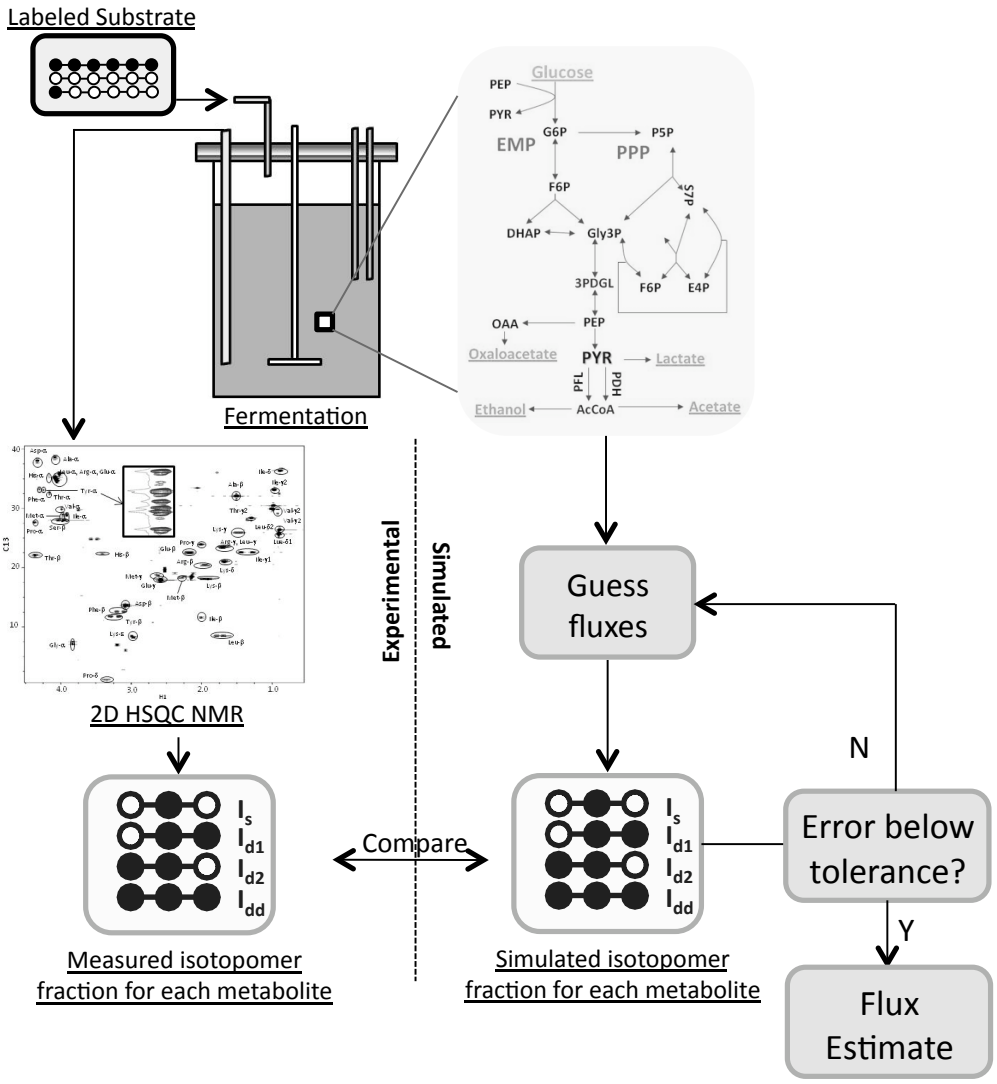
A batch experiment is typically started with an inoculum precultured on unlabeled glucose (i.e. with natural labeling of carbon atoms). These cells are then inoculated in a medium containing a mixture of labeled and unlabeled glucose to an initial cell concentration x_0 at time t_0. For analytical purposes, we will assume that the fraction of labeled glucose is y. As the cells start consuming glucose to synthesize biomass, labeled carbon atoms are incorporated in the same fraction as in the substrate. Since the amount of biomass at any given time (x) can be expressed in terms of the number of doublings (n) as $2^n x_0$, the fraction of labeled carbon atoms (f) in the biomass can be given by:

$$f = y\left(1 - \left(1/2^n\right)\right) \qquad (d)$$

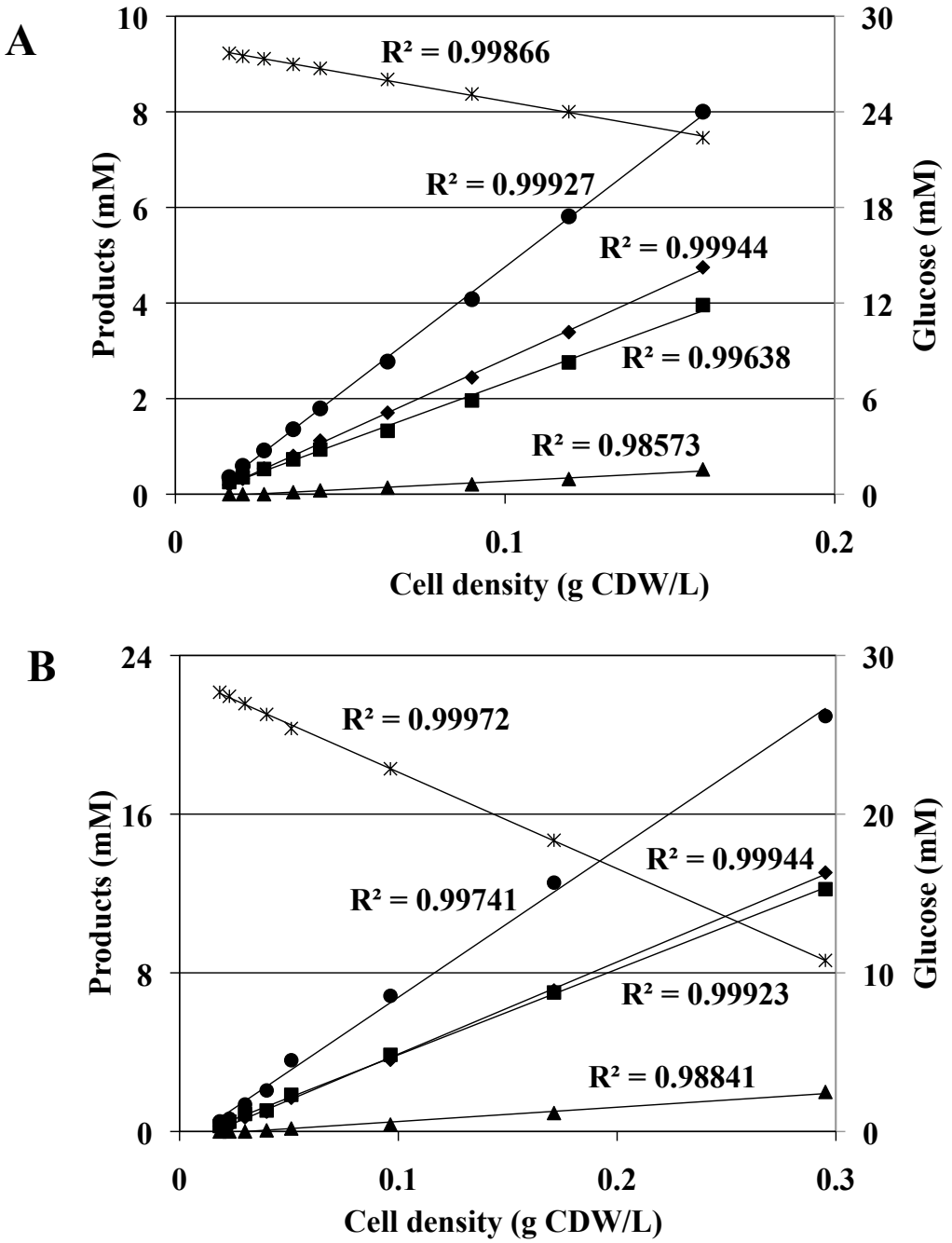
Using this expression it can be shown that as $n \to \infty$ then $f \to y$, and therefore y can be regarded as the steady state value of f. Using the above expression the calculated value of f after 5 doublings is 96.9% of its steady state value (y), and hence the assumption that five doublings are sufficient to achieve isotopic steady state. In our experiments, the cultures were harvested after five doublings. The computer program used to compute intracellular fluxes, NMR2Flux (1), accounts for the labeling pattern of the biomass before inoculation and hence estimates the true constant flux during exponential phase.

REFERENCES

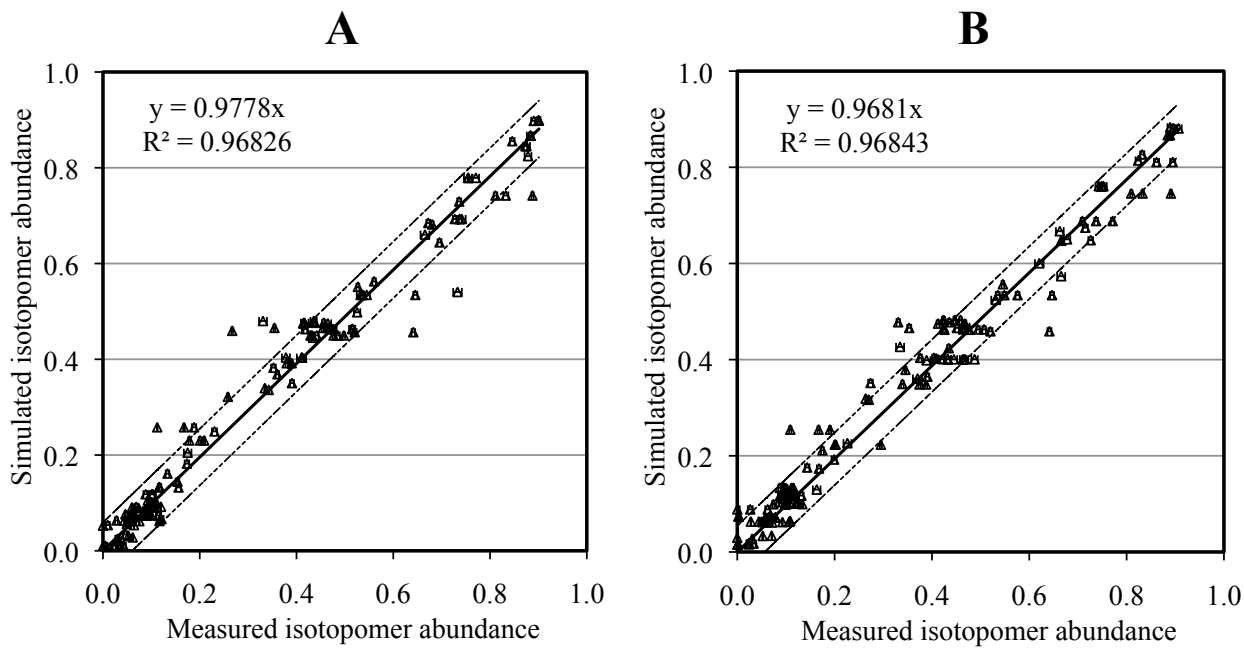
1. Sriram, G., Fulton, D. B., Iyer, V. V., Peterson, J. M., Zhou, R., Westgate, M. E., Spalding, M. H., and Shanks, J. V. **(**2004) *Plant Phys.* **136,**3043-3057.
2. Zamboni, N., Fendt, S-M., Ruhl, M., and Sauer, U. (2009) *Nat. Protoc.* **4,**878-892.



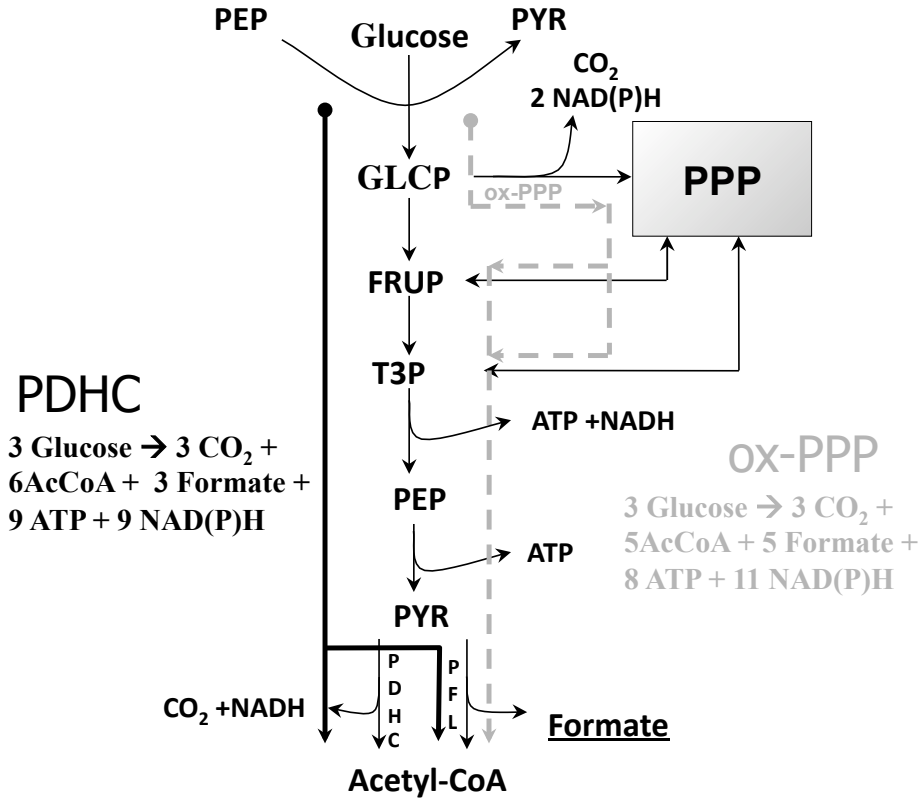
Supplementary Fig. 1. An overview of flux estimation using labeled substrate, NMR analysis, and isotopomer balancing. An experimenter cultures an organism on a a mixture of labeled (filled circle) and unlabeled (open circle) carbon source, harvest the biomass, obtains 2D [^{13}C, ^{1}H] HSQC spectra and estimates the fraction of each isotopomer, which constitutes experimental data as shown on the left hand side of the Figure. The computer program NMR2FLUX guesses fluxes satisfying the stoichiometry and other user input flux parameters and simulates isotopomer distributions based on the guessed fluxes and known biochemistry. This constitutes the simulated data shown on the right hand side of the Figure. The software then computes the error between simulated and experimental data and follows a simulated annealing protocol to identify the global minimum by adjusting the guessed fluxes. The final guessed set of fluxes at the identified global minimum is the output and represents the *in vivo* fluxes since they satisfy all the constraints as well as the NMR data.



Supplementary Fig. 2. Plot of the extracellular metabolite concentrations versus cell density during the exponential phase of growth showing a straight line fit between the two for all the measured metabolites in strains MG1655 (**A**) and Pdh (**B**). Symbols are as follows: glucose (asterisks), acetate (squares), ethanol (diamonds), formate (circles), and succinate (triangles).



Supplementary Fig. 3. Comparison of experimental and simulated isotopomer abundances for strains MG1655 (**A**) and Pdh (**B**). The *x*-axis represents experimental isotopomer abundances, measured from [^{13}C, ^{1}H] spectra, while the *y*-axis represents isotopomer abundances simulated by the computer program NMR2Flux, corresponding to the evaluated fluxes of Fig. 5. Isotopomer abundances are shown as fractions of the corresponding metabolite pool. The thick solid line represents a linear fitting as shown in the equation while the thin dotted lines illustrate the 95% confidence band.



Supplementary Fig. 4. Generation of CO_2 by the oxidative branch of the pentose phosphate pathway (ox-PPP) and the pyruvate dehydrogenase complex (PDHC). The stoichiometry of the overall reaction in each pathway is shown, assuming equal CO_2 yields (i.e., generation of the same amount of CO_2 upon consumption of the same amount of glucose).

Supplementary Table 1a. Reaction network model and metabolic fluxes calculated using metabolite balancing[a]

Rx#	Stoichiometry	Calculated fluxes (mmol/gCDW/h)				Calculated fluxes (normalized)			
		MG1655		Pdh		MG1655		Pdh	
		FLUX	SD	FLUX	SD	FLUX	SD	FLUX	SD
	Glucose uptake								
r1	GLUCOSE + PEP → G6P + PYR	5.225	0.364	4.620	0.423	100.000	6.966	100.000	9.152
	Embden-Meyerhof-Parnas								
r2	G6P → F6P	5.070	0.288	4.007	0.334	97.034	5.503	86.731	7.229
r3	F6P + ATP → 2 T3P + ADP	5.069	0.337	4.327	0.329	97.011	6.447	93.663	7.120
r4	T3P + NAD → NADH + 13P2DG	10.107	0.710	8.788	0.656	193.424	13.584	190.220	14.208
r5	13P2DG + ADP → ATP + 3PDGL	10.107	0.710	8.788	0.656	193.424	13.584	190.220	14.208
r6	3PDGL → PEP	9.938	0.710	8.643	0.656	190.202	13.584	187.070	14.208
r7	PEP + ADP → ATP + PYR	3.858	0.242	3.294	0.239	73.845	4.628	71.292	5.171
	Pentose phosphate pathway								
r8	G6P + 2 NADP → 2 NADPH + CO2 + RL5P	0.133	0.186	0.597	0.051	2.539	3.557	12.914	1.098
r9	RL5P → R5P	0.126	0.062	0.269	0.017	2.403	1.186	5.827	0.366
r10	RL5P → X5P	0.007	0.124	0.327	0.034	0.136	2.371	7.087	0.732
r11	R5P + X5P → T3P + S7P	0.024	0.062	0.181	0.017	0.453	1.186	3.920	0.366
r12	T3P + S7P → E4P + F6P	0.024	0.062	0.181	0.017	0.453	1.186	3.920	0.366
r13	X5P + E4P → F6P + T3P	-0.017	0.062	0.146	0.017	-0.318	1.186	3.167	0.366
	Glycogen metabolism								
r14	G6P → G1P	0.016	0.000	0.014	0.000	0.302	0.004	0.294	0.004
r15	G1P + ATP → ADP + GLYCOGEN	0.016	0.000	0.014	0.000	0.302	0.004	0.294	0.004
	Pyruvate dissimilation and fermentation								
r16	PYR + NADH → NAD + LAC	0.048	0.037	0.051	0.026	0.918	0.700	1.108	0.570
r17	PYR + COA + NAD → ACCOA + CO2 + NADH	1.093	0.155			20.928	2.957	0.000	0.000
r18	PYR + COA → FORMATE + ACCOA	7.622	0.666	7.592	0.550	145.864	12.741	164.330	11.914
r19	ACCOA + 2 NADH → 2 NAD + ETOH	4.564	0.320	3.990	0.230	87.345	6.132	86.370	4.971
r20	ACCOA + ADP → AC + ATP	3.701	0.282	3.560	0.328	70.839	5.393	77.056	7.097
r21	FORMATE → CO2 + H2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	TCA cycle - oxidative and reductive branches								
r22	PEP + CO2 → OA	0.774	0.115	0.656	0.097	14.813	2.205	14.208	2.101
r23	OA + ACCOA + NADP → AKG + COA + NADPH + CO2	0.114	0.002	0.098	0.017	2.177	0.030	2.127	0.362
r24	OA + 2 NADH → SUCC + 2 NAD	0.460	0.115	0.385	0.097	8.808	2.205	8.339	2.101
	Biosynthesis of amino acids								
r25	OA + NADPH → ASP + NADP	0.026	0.000	0.022	0.004	0.489	0.007	0.478	0.081
r26	OA + 3 ATP + NADPH → ASN + 3 ADP + NADP	0.026	0.000	0.022	0.004	0.489	0.007	0.478	0.081
r27	AKG + NADPH → GLU + NADP	0.031	0.000	0.027	0.005	0.594	0.008	0.580	0.099
r28	AKG + ATP + NADPH → GLN + NADP + ADP	0.028	0.000	0.024	0.004	0.534	0.007	0.522	0.089
r29	PYR + NADPH → ALA + NADP	0.061	0.001	0.052	0.009	1.161	0.016	1.134	0.193
r30	AKG + 7 ATP + NAD + 4 NADPH → ARG + 7 ADP + NADH + 4 NADP	0.031	0.000	0.027	0.005	0.600	0.008	0.587	0.100
r31	AKG + ATP + 3 NADPH → PRO + ADP + 3 NADP	0.023	0.000	0.020	0.003	0.448	0.006	0.438	0.075
r32	2 PYR + ACCOA + NAD + 2 NADPH → LEU + NADH + 2 NADP + 2 CO2	0.048	0.001	0.041	0.007	0.914	0.013	0.893	0.152
r33	2 PYR + 2 NADPH → VAL + 2 NADP + CO2	0.045	0.001	0.039	0.007	0.859	0.012	0.839	0.143
r35	2 PEP + E4P + ATP + 2 NADPH → PHE + 2 NADP + ADP + CO2	0.020	0.000	0.017	0.003	0.376	0.005	0.367	0.062
r36	2 PEP + E4P + ATP + NAD + 2 NADPH → TYR + ADP + NADH + 2 NADP + C02	0.015	0.000	0.013	0.002	0.280	0.004	0.273	0.046
r37	PEP + E4P + R5P + 5 ATP + 2 NAD + 3 NADPH → TRP + 5 ADP + 2 NADH + 3 NADP + CO2	0.006	0.000	0.005	0.001	0.115	0.002	0.113	0.019
r38	R5P + 6 ATP + 3 NAD + NADPH + 1C → HIS + 6 ADP + 3 NADH + NADP	0.010	0.000	0.009	0.001	0.192	0.003	0.188	0.032
r39	3 PDGL + NAD + NADPH → SER + NADH + NADP	0.038	0.001	0.032	0.006	0.718	0.010	0.701	0.119
r40	3 PDGL + NAD + NADPH → GLY + 1C + NADH + NADP	0.071	0.001	0.061	0.010	1.354	0.019	1.324	0.225
r41	3 PDGL + 4 ATP + NAD + NADPH → CYS + 4 ADP + NADH + NADP	0.010	0.000	0.008	0.001	0.186	0.003	0.182	0.031
r42	OA + 2 ATP + 3 NADPH → THR + 2 ADP + 3 NADP	0.027	0.000	0.023	0.004	0.515	0.007	0.503	0.086
r43	PYR + OA + 3 ATP + 4 NADPH → LYS + 3 ADP + 4 NADP + CO2	0.036	0.001	0.031	0.005	0.696	0.010	0.680	0.116
r44	OA + 7 ATP + 8 NADPH + 1C → MET + 7 ADP + 8 NADP	0.016	0.000	0.014	0.002	0.312	0.004	0.305	0.052
	Biosynthesis of purines								
r45	3 PDGL + R5P + 9 ATP + 3 NAD + NADPH + 1C → AMP + 9 ADP + 3 NADH + NADP	0.019	0.000	0.016	0.003	0.358	0.005	0.350	0.060
r46	3 PDGL + R5P + 11 ATP + 3 NAD + 1C → GMP + 11 ADP + 3 NADH	0.023	0.000	0.020	0.003	0.441	0.006	0.431	0.073
r47	3 PDGL + R5P + 9 ATP + 3 NAD + 2 NADPH + 1C → DAMP + 9 ADP + 3 NADH + 2 NADP	0.003	0.000	0.002	0.000	0.053	0.001	0.052	0.009
r48	3 PDGL + R5P + 11 ATP + 3 NAD + NADPH + 1C → DGMP + 11 ADP + 3 NADH + NADP	0.003	0.000	0.002	0.000	0.055	0.001	0.054	0.009
	Biosynthesis of pyrimidines								
r49	OA + P5P + 5 ATP + NADPH → UMP + 5 ADP + NADP	0.015	0.000	0.013	0.002	0.295	0.004	0.289	0.049
r50	OA + P5P + 7 ATP + NADPH → CMP + 7 ADP + NADP	0.014	0.000	0.012	0.002	0.274	0.004	0.268	0.045
r51	OA + P5P + 5 ATP + 3 NADPH + 1C + 2 NH3 → DUMP + 5 ADP + 3 NADP	0.003	0.000	0.002	0.000	0.053	0.001	0.052	0.009
r52	OA + P5P + 7 ATP + 2 NADPH → DCMP + 7 ADP + 2 NADP	0.003	0.000	0.002	0.000	0.055	0.001	0.054	0.009
	1Carbon metabolism (Formation of 1C)								
r53	GLY + NAD → 1C + CO2 + NADH	0.006	0.000	0.005	0.001	0.111	0.002	0.109	0.018
	Biosynthesis of lipid components								
r54	T3P + NADPH → GLY3P + NADP	0.015	0.000	0.013	0.002	0.280	0.004	0.273	0.046
r55	8.2 ACCOA + 7.2 ATP + 14 NADPH → AVGFAT + 7.2 ADP + 14 NADP	0.029	0.000	0.025	0.004	0.560	0.008	0.547	0.093
	Biosynthesis of LPS components								
r56	G6P + ATP → UDPG + ADP	0.002	0.000	0.002	0.000	0.038	0.001	0.038	0.006
r57	3 PDGL + 3 ATP + NAD + NADPH → CDPETN + 3 ADP + NADH + NADP	0.003	0.000	0.003	0.000	0.058	0.001	0.056	0.010
r58	7 ACCOA + 6 ATP + 11 NADPH → OHMA + 6 ADP + 11 NADP + 7 COA	0.003	0.000	0.003	0.000	0.058	0.001	0.056	0.010
r59	7 ACCOA + 6 ATP + 12 NADPH → C14:0 + 6A DP + 12 NADP +7 COA	0.003	0.000	0.003	0.000	0.058	0.001	0.056	0.010
r60	P5P + PEP + 2 ATP → CMPKDO + 2 ADP	0.003	0.000	0.003	0.000	0.058	0.001	0.056	0.010
r61	1.5 G6P + ATP + 4 NADP → NHEP + ADP + 4 NADPH	0.003	0.000	0.003	0.000	0.058	0.001	0.056	0.010
r62	F6P + 2 ATP + NH3 → TGSM + 2 ADP	0.002	0.000	0.002	0.000	0.038	0.001	0.038	0.006
	Biosynthesis of peptido components								
r63	F6P + ACCOA + 3 ATP + NH3 → UDPNAG + 3 ADP + COA	0.003	0.000	0.003	0.000	0.060	0.001	0.059	0.010
r64	F6P + PEP + ACCOA + 4 ATP + NADPH + NH3 → UDPNAM + 4 ADP + NADP + COA	0.003	0.000	0.003	0.000	0.060	0.001	0.059	0.010
r65	OA + PYR + 2 ATP + 3 NADPH + 2 NH3 → DAP + 2 ADP + 3 NADP	0.003	0.000	0.003	0.000	0.060	0.001	0.059	0.010
	ATP maintenance								

r66	ATP → ADP	7.963	0.881	6.962	0.891	152.389	16.855	150.702	19.289
	Transhydrogenase								
r67	NADH + NADP → NADPH + NAD	1.505	0.372	0.336	0.101	28.794	7.113	7.270	2.197
	Synthesis of macromolecules elemental units								
r68	0.105 UDPG + 0.1578 CDPETN + 0.1578 OHMA + 0.1578 C14:0 + 0.1578 CMPKDO + 0.1578 NHEP+ 0.105TGSM → LPScell	0.019	0.000	0.016	0.003	0.365	0.005	0.357	0.061
r69	0.25 GLY3P + 0.25 SER + 0.5 AVGFAT+2ATP → LIPIDcell + 2 ADP	0.059	0.001	0.051	0.009	1.120	0.016	1.093	0.186
r70	0.0960 ALA + 0.0553 ARG + 0.0451 ASN + 0.0451 ASP + 0.0171 CYS + 0.0492 GLN + 0.0492 GLU + 0.1145 GLY + 0.0177 HIS + 0.0543 ILE + 0.0842 LEU + 0.0641 LYS + 0.0287 MET + 0.0346 PHE + 0.0413 PRO + 0.0403 SER + 0.0474 THR + 0.0106 TRP + 0.0257 TYR + 0.0791 VAL + 4.3 ATP → PROTcell + 4.3ADP	0.567	0.008	0.490	0.083	10.854	0.152	10.597	1.802
r71	0.2465 DAMP + 0.2535 DCMP + 0.2535 DGMP + 0.2465 DUMP + 5.4 ATP → DNA + 5.4ADP	0.011	0.000	0.010	0.002	0.217	0.003	0.212	0.036
r72	0.2619 AMP + 0.2 CMP + 0.3222 GMP + 0.2159 UMP + 4.4 ATP → RNA + 4.4ADP	0.072	0.001	0.062	0.011	1.369	0.019	1.337	0.227
r73	0.167 UDPNAG + 0.167 UDPNAM + 0.33 ALA + 0.167 DAP + 0.167 GLU + 0.833 ATP → PEPTIDO + 0.833 ADP	0.019	0.000	0.016	0.003	0.359	0.005	0.351	0.060
	Synthesis of CELL								
r74	0.0149 DNA + 0.0201 GLYCOGEN + 0.0768 LIPIDcell + 0.0250 LPScell + 0.0246 PEPTIDO + 0.7441 PROTcell + 0.0939 RNA → CELL	0.434	0.006	3.750	0.638	8.306	0.118	81.169	13.799
	Transport reactions								
r75	GLUCOSEext → GLUCOSE	5.225	0.364	4.620	0.423	100.000	6.966	100.000	9.152
r76	LAC → LACext	0.048	0.037	0.051	0.026	0.918	0.700	1.108	0.570
r77	FORMATE → FORMATEext	7.622	0.666	7.592	0.550	145.864	12.741	164.330	11.914
r78	AC → ACext	3.701	0.282	3.560	0.328	70.839	5.393	77.056	7.097
r79	ETOH → ETOHext	4.564	0.320	3.990	0.230	87.345	6.132	86.370	4.971
r80	SUCC → SUCCext	0.460	0.115	0.385	0.097	8.808	0.000	8.339	0.000
r81	CO2 → CO2ext	0.805	0.087	0.245	0.104	15.405	1.657	5.304	2.259

Reaction r21 (underlined) was assumed not to be taking place at pH 7.4.

[a] *Normalized fluxes are expressed as the percentage of the glucose uptake flux.*

Supplementary Table 1b. Reaction network model and metabolic fluxes calculated using isotopomer balancing[a]

Rx#	Stoichiometry	Fluxes (isotopomer balancing, normalized)							
		MG1655				Pdh			
		FLUX	SD	REV	SD	FLUX	SD	REV	SD
	Glucose uptake								
r1	GLUCOSE + PEP → G6P + PYR	100.000	0.000			100.000	0.000		
	Embden-Meyerhof-Parnas								
r2	G6P → F6P	96.868	1.567	0.749	0.293	84.527	1.970	0.976	0.024
r3	F6P + ATP → 2 T3P + ADP	96.958	0.522	0.681	0.111	92.889	0.657	0.892	0.078
r4	T3P + NAD → NADH + 13P2DG								
r5	13P2DG + ADP → ATP + 3PDGL	190.234	0.522	IRREV	IRREV	186.309	0.657	IRREV	IRREV
r6	3PDGL → PEP								
r7	PEP + ADP → ATP + PYR	74.427	1.659			68.877	1.627		
	Pentose phosphate pathway								
r8	G6P + 2 NADP → 2 NADPH + CO2 + RL5P	2.705	1.567			15.056	1.970		
r9	RL5P → R5P	2.705	1.567			15.056	1.970		
r10	RL5P → X5P	2.705	1.567			15.056	1.970		
r11	R5P + X5P → T3P + S7P	0.509	0.522	0.987	0.004	4.635	0.657	0.899	0.049
r12	T3P + S7P → E4P + F6P	-0.261	0.522	0.648	0.309	3.881	0.657	0.118	0.178
r13	X5P + E4P → F6P + T3P	0.509	0.522	0.981	0.015	4.635	0.657	0.918	0.066
	Glycogen Metabolism								
r14	G6P → G1P	0.016	0.000			0.014	0.000		
r15	G1P + ATP → ADP + GLYCOGEN	0.016	0.000			0.014	0.000		
	Pyruvate dissimilation and fermentation								
r16	PYR + NADH → NAD + LAC	0.965	0.101			1.073	0.119		
r17	PYR + COA + NAD → ACCOA + CO2 + NADH	21.570	2.398			-	-		
r18	PYR + COA → FORMATE + ACCOA	145.846	2.423	0.436	0.012	161.889	1.696	0.427	0.014
r19	ACCOA + 2 NADH → 2 NAD + ETOH	158.816*	1.665			153.482*	1.696		
r20	ACCOA + ADP → AC + ATP								
r21	FORMATE → CO2 + H2	0.000	0.000			0.000	0.000		
	TCA cycle - oxidative and reductive branches								
r22	PEP + CO2 → OA	14.265	1.529	0.000	0.000	15.923	1.323	0.000	0.000
r23	OA + ACCOA + NADP → AKG + COA + NADPH + CO2	2.174	0.022	0.391	0.038	2.127	0.021	0.003	0.005
r24	OA + 2 NADH → SUCC + 2 NAD	8.267	1.529	0.313	0.267	10.056	1.323	0.431	0.308
	Biosynthesis of amino acids								
r25	OA + NADPH → ASP + NADP	0.490	0.005			0.480	0.005		
r26	OA + 3 ATP + NADPH → ASN + 3 ADP + NADP	0.490	0.005			0.480	0.005		
r27	AKG + NADPH → GLU + NADP	0.535	0.005			0.524	0.005		
r28	AKG + ATP + NADPH → GLN + NADP + ADP	0.535	0.005			0.524	0.005		
r29	PYR + NADPH → ALA + NADP	1.044	0.010			1.022	0.010		
r30	AKG + 7 ATP + NAD + 4 NADPH → ARG + 7 ADP + NADH + 4 NADP	0.601	0.006			0.589	0.006		
r31	AKG + ATP + 3 NADPH → PRO + ADP + 3 NADP	0.449	0.004			0.440	0.004		
r32	2 PYR + ACCOA + NAD + 2 NADPH → LEU + NADH + 2 NADP + 2 CO2	0.916	0.009			0.897	0.009		
r33	2 PYR + 2 NADPH → VAL + 2 NADP + CO2	0.860	0.008			0.842	0.008		
r34	OA + PYR + 2 ATP + 5 NADPH → ILE + 5 NADP + 2 ADP + CO2	0.591	0.006			0.578	0.006		
r35	2 PEP + E4P + ATP + 2 NADPH → PHE + 2 NADP + ADP + CO2	0.377	0.004			0.369	0.004		
r36	2 PEP + E4P + ATP + NAD + 2 NADPH → TYR + ADP + NADH + 2 NADP + C02	0.280	0.003			0.274	0.003		
r37	PEP + E4P + R5P + 5 ATP + 2 NAD + 3 NADPH → TRP + 5 ADP + 2 NADH + 3 NADP + CO2	0.116	0.001			0.113	0.001		
r38	R5P + 6 ATP + 3 NAD + NADPH + 1C → HIS + 6 ADP + 3 NADH + NADP	0.193	0.002			0.189	0.002		
r39	3 PDGL + NAD + NADPH → SER + NADH + NADP	0.439	0.004			0.429	0.004		
r40	3 PDGL + NAD + NADPH → GLY + 1C + NADH + NADP	1.245	0.012			1.219	0.012		
r41	3 PDGL + 4 ATP + NAD + NADPH → CYS + 4 ADP + NADH + NADP	0.186	0.002			0.182	0.002		
r42	OA + 2 ATP + 3 NADPH → THR + 2 ADP + 3 NADP	0.516	0.005			0.505	0.005		
r43	PYR + OA + 3 ATP + 4 NADPH → LYS + 3 ADP + 4 NADP + CO2	0.698	0.007			0.683	0.007		
r44	OA + 7 ATP + 8 NADPH + 1C → MET + 7 ADP + 8 NADP	0.312	0.003			0.306	0.003		
	Biosynthesis of purines								
r45	3 PDGL + R5P + 9 ATP + 3 NAD + NADPH + 1C → AMP + 9 ADP + 3 NADH + NADP	0.359	0.003			0.352	0.003		
r46	3 PDGL + R5P + 11 ATP + 3 NAD + 1C → GMP + 11 ADP + 3 NADH	0.442	0.004			0.433	0.004		
r47	3 PDGL + R5P + 9 ATP + 3 NAD + 2 NADPH + 1C → DAMP + 9 ADP + 3 NADH + 2 NADP	0.054	0.001			0.052	0.001		
r48	3 PDGL + R5P + 11 ATP + 3 NAD + NADPH + 1C → DGMP + 11 ADP + 3 NADH + NADP	0.055	0.001			0.054	0.001		
	Biosynthesis of pyrimidines								
r49	OA + P5P + 5 ATP + NADPH → UMP + 5 ADP + NADP	0.296	0.003			0.290	0.003		
r50	OA + P5P + 7 ATP + NADPH → CMP + 7 ADP + NADP	0.274	0.003			0.269	0.003		
r51	OA + P5P + 5 ATP + 3 NADPH + 1C + 2 NH3 → DUMP + 5 ADP + 3 NADP	0.054	0.001			0.052	0.001		
r52	OA + P5P + 7 ATP + 2 NADPH → DCMP + 7 ADP + 2 NADP	0.055	0.001			0.054	0.001		
	1Carbon Metabolism (Formation of 1C)								
r53	GLY + NAD → 1C + CO2 + NADH	0.100	0.000	0.277	0.267	0.100	0.000	0.989	0.248
	Biosynthesis of lipid components								
r54	T3P + NADPH → GLY3P + NADP	0.280	0.003			0.275	0.003		
r55	8.2 ACCOA + 7.2 ATP + 14 NADPH → AVGFAT + 7.2 ADP + 14 NADP	0.561	0.005			0.549	0.005		
	Biosynthesis of LPS components								
r56	G6P + ATP → UDPG + ADP	0.039	0.000			0.038	0.000		
r57	3 PDGL + 3 ATP + NAD + NADPH → CDPETN + 3 ADP + NADH + NADP	0.058	0.001			0.057	0.001		
r58	7 ACCOA + 6 ATP + 11 NADPH → OHMA + 6 ADP + 11 NADP + 7 COA	0.058	0.001			0.057	0.001		
r59	7 ACCOA + 6 ATP + 12 NADPH → C14:0 + 6A DP + 12 NADP +7 COA	0.058	0.001			0.057	0.001		
r60	P5P + PEP + 2 ATP → CMPKDO + 2 ADP	0.058	0.001			0.057	0.001		
r61	1.5 G6P + ATP + 4 NADP → NHEP + ADP + 4 NADPH	0.058	0.001			0.057	0.001		
r62	F6P + 2 ATP + NH3 → TGSM + 2 ADP	0.039	0.000			0.038	0.000		
	Biosynthesis of peptido components								
r63	F6P + ACCOA + 3 ATP + NH3 → UDPNAG + 3 ADP + COA	0.060	0.001			0.059	0.001		
r64	F6P + PEP + ACCOA + 4 ATP + NADPH + NH3 → UDPNAM + 4 ADP + NADP + COA	0.060	0.001			0.059	0.001		

r65	OA + PYR + 2 ATP + 3 NADPH + 2 NH3 → DAP + 2 ADP + 3 NADP	0.060	0.001	0.059	0.001
	ATP maintenance				
r66	ATP → ADP	ND	ND	ND	ND
	Transhydrogenase				
r67	NADH + NADP → NADPH + NAD	25.800		0.535	
	Synthesis of macromolecules elemental units				
r68	0.105 UDPG + 0.1578 CDPETN + 0.1578 OHMA + 0.1578 C14:0 + 0.1578 CMPKDO + 0.1578 NHEP+ 0.105TGSM → LPScell	0.366	0.004	0.358	0.003
r69	0.25 GLY3P + 0.25 SER + 0.5 AVGFAT+2ATP → LIPIDcell + 2 ADP	1.122	0.011	1.098	0.011
r70	0.0960 ALA + 0.0553 ARG + 0.0451 ASN + 0.0451 ASP + 0.0171 CYS + 0.0492 GLN + 0.0492 GLU + 0.1145 GLY + 0.0177 HIS + 0.0543 ILE + 0.0842 LEU + 0.0641 LYS + 0.0287 MET + 0.0346 PHE + 0.0413 PRO + 0.0403 SER + 0.0474 THR + 0.0106 TRP + 0.0257 TYR + 0.0791 VAL + 4.3 ATP → PROTcell + 4.3ADP	10.872	0.104	10.644	0.102
r71	0.2465 DAMP + 0.2535 DCMP + 0.2535 DGMP + 0.2465 DUMP + 5.4 ATP → DNA + 5.4ADP	0.217	0.002	0.213	0.002
r72	0.2619 AMP + 0.2 CMP + 0.3222 GMP + 0.2159 UMP + 4.4 ATP → RNA + 4.4ADP	1.372	0.013	1.343	0.013
r73	0.167 UDPNAG + 0.167 UDPNAM + 0.33 ALA + 0.167 DAP + 0.167 GLU + 0.833 ATP → PEPTIDO + 0.833 ADP	0.360	0.003	0.352	0.003
	Synthesis of CELL				
r74	0.0149 DNA + 0.0201 GLYCOGEN + 0.0768 LIPIDcell + 0.0250 LPScell + 0.0246 PEPTIDO + 0.7441 PROTcell + 0.0939 RNA → CELL	8.335	0.080	8.160	0.078
	Transport reactions				
r75	GLUCOSEext → GLUCOSE	100.000	0.000	100.000	0.000
r76	LAC → LACext	0.965	0.101	1.073	0.119
r77	FORMATE → FORMATEext	145.846	2.423	161.889	1.696
r78	AC → ACext	158.816**	1.665	153.482**	1.696
r79	ETOH → ETOHext				
r80	SUCC → SUCCext	8.267	2.226	10.056	2.325
r81	CO2 → CO2ext	12.285	4.027	1.361	2.059

Reaction r21 (underlined) was assumed not to be taking place at pH 7.4.

[a] Fluxes were normalized and expressed as the percentage of the glucose uptake flux. The transhydrogenase flux (r67) was calculated based on estimated fluxes for reactions consuming and generating NAD(P)H and assuming redox balanced conditions.

** Isotopomer balancing does not differentiate between ethanol and acetate since there are no measurable carbon backbone rearrangement between the two metabolites. Hence, the reported flux is a sum of flux generating these metabolites.*

Metabolites	
Abbreviation	**Description (Reference)**
1C	1 carbon carrier with methyl group
13P2DG	1,3-P-D glycerate
3PDGL	3-Phospho-D-Glycerate
AC	Acetate
ACext	Extracellular Acetate
ACCOA	Acetyl Coenzyme A
AKG	α-Ketoglutarate
ALA	Alanine
AMP	Adenosine monophosphate
ARG	Arginine
ASN	Asparagine
ASP	Aspartate
ATP	Adenosine triphosphate
AVGFAT	Average fatty acid molecule (3, 4, 6)
C14:0	Myristic acid
CDPETN	CDP-Ethanolamine
CELL	Average cell with composition corresponding to a molecular weight of 143 (3, 4, 6)
CMP	Cytidine monophosphate
CMPKDO	CMP-2-Keto-3-deoxyoctanoate
CO2	Carbon dioxide
CO2ext	Extracellular Carbon dioxide
CYS	Cysteine
DAMP	Deoxyadenosine monophosphate
DAP	Diaminopimelate
DCMP	Deoxycytidine monophosphate
DGMP	2-Deoxy-guanosine-5-phosphate
DNA	Average DNA-bound nucleic acid (mol. wt. 309.38) (3, 4, 6)
DUMP	Deoxyuridine monophosphate
E4P	Erythrose 4-phosphate
ETOH	Ethanol
ETOHext	Extracellular Ethanol
F6P	Fructose 6-phosphate
FORMATE	Formate
FORMATEext	Extracellular Formate
G1P	Glucose 1-phosphate
G6P	Glucose 6-phosphate
GLN	Glutamine
GLU	Glutamate
GLY	Glycine
GLUCOSE	Glucose
GLUCOSEext	Extracellular Glucose
GLYCOGEN	Glycogen
GLY3P	Glycerol-3-phosphate
GMP	Guanosine monophosphate
HIS	Histidine
ILE	Isoleucine
LAC	Lactate
LACext	Extracellular Lactate
LEU	Leucine
LIPIDcell	Average LIPID building block (MW 176) (3, 4, 6)
LPScell	Average Lipposaccharide building block (MW 202) (3, 4, 6)

LYS	Lysine
MET	Methionine
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NHEP	Heptulose
OA	Oxaloacetate
OHMA	Myristic acid
PEP	Phosphoenolpyruvate
PEPTIDO	Average Peptidoglycan building block (MW 151) (3, 4, 6)
PHE	Phenylalanine
PRO	Proline
PROTcell	Average proteinogenic amino acid (MW 110) (3, 4, 6)
PYR	Pyruvate
RL5P	Ribulose 5-phosphate
R5P	Ribose-5-phosphate
RNA	Average RNA bound nucleic acid (MW 325) (3, 4, 6)
S7P	Sedoheptulose-7-P
SER	Serine
SUCC	Succinate
SUCCext	Extracellular Succinate
T3P	Glyceraldehyde-3-phosphate/Dihydroxyacetone-3-phosphate
TGSM	Glucosamine
THR	Threonine
TRP	Tryptophan
TYR	Tyrosine
UDPG	UDP-Glucose
UDPNAG	UDP N-acetyl glucosamine
UDPNAM	UDP-N-acetyl-D-muramate
UMP	Uridine monophosphate
VAL	Valine
X5P	Xylose-5-Phosphate

Extracellular metabolites are shown in yellow

REFERENCES

1 **Neidhardt F C, Curtiss III R, Ingraham J L, Lin E C C, Low K B, Magasanik B, Reznikoff W S, Riley M, Schaechter M, Umbarger H E. , editors.** 1996. *Escherichia coli* and *Salmonella*: cellular and molecular biology. 2nd ed. ASM Press, Washington, D.C.

2 **Keseler, I. M., J. Collado-Vides, S. Gama-Castro, J. Ingraham, S. Paley, I. T. Paulsen, M. Peralta-Gil, and P. D. Karp.** 2005. EcoCyc: a comprehensive database resource for Escherichia coli. Nucleic Acids Res. **33**:(Database issue)D334-D337.

3 **Neidhardt, F. C., J. Ingraham, and M. Schaechter.** 1990. Physiology of the bacterial cell. Sinauer Associates, Sunderland, MA, USA.

5 **Pramanik, J. and J.D. Keasling.** 1997. A stoichiometric model of Escherichia coli metabolism: incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. Biotech. Bioeng. **56:**398-421.

6 **Stephanopoulos, G., A. A. Aristidou, and J. Nielsen.** 1998. Metabolic Engineering: principles and methodology. Academic Press, CA, USA.

7 **Stryer, L.** 1995. Biochemistry. Freeman, New York, USA.

Supplementary Table 2. Isotopomer abundances of intracellular precursor metabolites in MG1655 and Pdh grown on a mixture of labeled and unlabeled glucose (12.5% U-13C, 25%1-13C , 62.5% unlabeled)[a].

			MG1655			PDH-	
Nucleus	Precursor	Carbon backbone of the precursor leading to isotopomer fraction	Mean	SD		Mean	SD
Ala-a (s)	PYR	123	0.074	0.013		0.074	0.013
Ala-a (d1)	PYR	123	0.023	0.013		0.024	0.013
Ala-a (d2)	PYR	123	0.343	0.013		0.371	0.013
Ala-a (dd)	PYR	123	0.560	0.013		0.531	0.013
Ala-b (s)	PYR	123	0.533	0.010		0.576	0.006
Ala-b (d)	PYR	123	0.468	0.010		0.424	0.006
Arg-b (s)	AKG	x234x	0.058	0.013		0.062	0.014
Arg-b (d)	AKG	x234x + x234x	0.742	0.013		0.737	0.014
Arg-b (t)	AKG	x234x	0.201	0.013		0.200	0.014
Arg-d (s)	AKG	xx345	0.100	0.005		0.104	0.007
Arg-d (d)	AKG	xx345	0.900	0.005		0.896	0.007
Asx-a (s)	OAA	123x	0.102	0.012		0.121	0.014
Asx-a (d1)	OAA	123x	0.042	0.012		0.025	0.014
Asx-a (d2)	OAA	123x	0.101	0.012		0.112	0.014
Asx-a (dd)	OAA	123x	0.755	0.012		0.743	0.014
Asx-b (s)	OAA	x234	0.442	0.005		0.468	0.006
Asx-b (d1)	OAA	x234	0.380	0.005		0.339	0.006
Asx-b (d2)	OAA	x234	0.086	0.005		0.100	0.006
Asx-b (dd)	OAA	x234	0.092	0.005		0.093	0.006
Glx-a (s)	AKG	123xx	Not observed			0.432	0.015
Glx-a (d1)	AKG	123xx	Not observed			0.097	0.015
Glx-a (d2)	AKG	123xx	Not observed			0.374	0.015
Glx-a (dd)	AKG	123xx	Not observed			0.097	0.015
Glx-b (s)	AKG	x234x	0.094	0.010		0.027	0.013
Glx-b (d)	AKG	x234x + x234x	0.727	0.010		0.771	0.013
Glx-b (t)	AKG	x234x	0.179	0.010		0.202	0.013
Glx-g (s)	AKG	xx345	0.476	0.005		0.477	0.006
Glx-g (d1)	AKG	xx345	0.059	0.005		0.065	0.006
Glx-g (d2)	AKG	xx345	0.410	0.005		0.409	0.006
Glx-g (dd)	AKG	xx345	0.054	0.005		0.049	0.006
Gly-a (s)	T3P	12x	0.154	0.005		0.168	0.009
Gly-a (d)	T3P	12x	0.846	0.005		0.832	0.009
His-a (s)	R5P	xx345	Not observed			0.098	0.019
His-a (d1)	R5P	xx345	Not observed			0.078	0.019
His-a (d2)	R5P	xx345	Not observed			0.000	0.019
His-a (dd)	R5P	xx345	Not observed			0.824	0.019
His-b (s)	R5P	x234x	0.133	0.006		0.144	0.007
His-b (d1)	R5P	x234x	0.681	0.006		0.663	0.007
His-b (d2)	R5P	x234x	0.030	0.006		0.030	0.007
His-b (dd)	R5P	x234x	0.156	0.006		0.163	0.007
His-d2 (s)	R5P	12xxx	0.666	0.015		0.666	0.016
His-d2 (d)	R5P	12xxx	0.334	0.015		0.334	0.016
Ile-a (s)	OAA	12xx.x2x	0.174	0.024		0.199	0.026
Ile-a (d1)	OAA	12xx.x2x	0.672	0.024		0.715	0.026
Ile-a (d2)	OAA	12xx.x2x	0.061	0.024		0.000	0.026
Ile-a (dd)	OAA	12xx.x2x	0.093	0.024		0.086	0.026
Ile-g1 (s)	OAA	xx34.x2x	0.736	0.021		0.678	0.019
Ile-g1 (d)	OAA	xx34.x2x + xx34.x2x	0.230	0.021		0.270	0.019
Ile-g1 (t)	OAA	xx34.x2x	0.033	0.021		0.052	0.019
Ile-d (s)	OAA	xx34	0.833	0.032		0.833	0.021
Ile-d (d)	OAA	xx34	0.167	0.032		0.167	0.021
Ile-g2 (s)	PYR	x23	0.537	0.032		0.535	0.021
Ile-g2 (d)	PYR	x23	0.463	0.032		0.465	0.021
Leu-a (s)	ACCOA	12.x2x	0.473	0.022		0.464	0.033
Leu-a (d1)	ACCOA	12.x2x	0.412	0.022		0.404	0.033
Leu-a (d2)	ACCOA	12.x2x	0.065	0.022		0.079	0.033
Leu-a (dd)	ACCOA	12.x2x	0.051	0.022		0.053	0.033
Leu-b (s)	PYR	x2.x2x.x2x	0.696	0.022		0.666	0.022
Leu-b (d)	PYR	x2.x2x.x2x + x2.x2x.x2x	0.258	0.022		0.264	0.022
Leu-b (t)	PYR	x2.x2x.x2x	0.046	0.022		0.071	0.022
Leu-d1 (s)	PYR	x23	0.646	0.027		0.647	0.019
Leu-d1 (d)	PYR	x23	0.354	0.027		0.353	0.019
Leu-d2 (s)	PYR	xx3.x2x	0.884	0.027		0.889	0.019
Leu-d2 (d)	PYR	xx3.x2x	0.116	0.027		0.111	0.019
Lys-b (s)	PYR/OAA	0.5{x234 + x23.xxx4}	0.476	0.011		0.464	0.022
Lys-b (d)	PYR/OAA	0.5{x234 + x23.xxx4 + x234 + x23.xxx4}	0.438	0.011		0.447	0.011
Lys-b (t)	PYR/OAA	0.5{x234 + x23.xxx4}	0.086	0.011		0.089	0.011
Lys-g (s)	OAA	xx34.xx3	0.527	0.016		0.546	0.016
Lys-g (d)	OAA	xx34.xx3 + xx34.xx3	0.352	0.016		0.346	0.016
Lys-g (t)	OAA	xx34.xx3	0.121	0.016		0.109	0.016
Lys-d (s)	PYR/OAA	0.5{x234 + x23·xxx4}	0.429	0.011		0.422	0.012

Lys-d (d)	PYR/OAA	0.5{x234 + x234 + x23·xxx4 + x23·xxx4}	0.457	0.011		0.459	0.012
Lys-d (t)	PYR/OAA	0.5{x234 + x23·xxx4}	0.115	0.011		0.119	0.012
Lys-e (s)	PYR/OAA	0.5{x23+x23x}	0.110	0.008		0.111	0.011
Lys-e (d)	PYR/OAA	0.5{x23+x23x}	0.890	0.008		0.889	0.011
Met-a (s)	OAA	123x	0.105	0.022		0.113	0.022
Met-a (d1)	OAA	123x	0.036	0.022		0.034	0.022
Met-a (d2)	OAA	123x	0.088	0.022		0.105	0.022
Met-a (dd)	OAA	123x	0.771	0.022		0.748	0.022
Met-b (s)	OAA	x234	0.485	0.024		0.445	0.024
Met-b (d)	OAA	x234 + x234	0.415	0.024		0.423	0.024
Met-b (t)	OAA	x234	0.100	0.024		0.132	0.024
Met-g (s)	OAA	xx34	0.888	0.023		0.891	0.023
Met-g (d)	OAA	xx34	0.112	0.023		0.109	0.023
Phe-a (s)	PEP	123	0.060	0.017		0.102	0.017
Phe-a (d1)	PEP	123	0.000	0.017		0.001	0.017
Phe-a (d2)	PEP	123	0.064	0.017		0.002	0.017
Phe-a (dd)	PEP	123	0.876	0.017		0.895	0.017
Phe-b (s)	PEP	x23.x2x	0.515	0.022		0.507	0.022
Phe-b (d1)	PEP	x23.x2x	0.378	0.022		0.376	0.022
Phe-b (d2)	PEP	x23.x2x	0.054	0.022		0.058	0.022
Phe-b (dd)	PEP	x23.x2x	0.052	0.022		0.059	0.022
Phe-d (s)	PEP/E4P	0.5{x23.1xxx+x2x.xx34}	0.641	0.026		0.641	0.029
Phe-d (d)	PEP/E4P	0.5{x23.1xxx+x2x.xx34+x23.1xxx+x2x.xx34}	0.331	0.026		0.331	0.029
Phe-d (t)	PEP/E4P	0.5{x23.1xxx+x2x.xx34}	0.028	0.026		0.028	0.029
Pro-a (s)	AKG	123xx	0.432	0.005		0.432	0.008
Pro-a (d1)	AKG	123xx	0.088	0.005		0.087	0.008
Pro-a (d2)	AKG	123xx	0.390	0.005		0.388	0.008
Pro-a (dd)	AKG	123xx	0.089	0.005		0.092	0.008
Pro-b (s)	AKG	x234x	0.046	0.010		-0.004	0.031
Pro-b (d)	AKG	x234x + x234x	0.737	0.010		0.709	0.031
Pro-b (t)	AKG	x234x	0.209	0.010		0.295	0.031
Pro-g (s)	AKG	xx345	0.418	0.011		0.426	0.015
Pro-g (d)	AKG	xx345 + xx345	0.465	0.011		0.468	0.015
Pro-g (t)	AKG	xx345	0.117	0.011		0.107	0.015
Pro-d (s)	AKG	xxx45	0.099	0.008		0.093	0.008
Pro-d (d)	AKG	xxx45	0.901	0.008		0.907	0.008
Ser-a (s)	3PDGL	123	0.069	0.011		0.094	0.011
Ser-a (d1)	3PDGL	123	0.053	0.011		0.226	0.011
Ser-a (d2)	3PDGL	123	0.000	0.011		0.059	0.011
Ser-a (dd)	3PDGL	123	0.879	0.011		0.621	0.011
Ser-b (s)	3PDGL	x23	0.733	0.017		0.726	0.017
Ser-b (d)	3PDGL	x23	0.267	0.017		0.274	0.017
Thr-a (s)	OAA	123x	0.120	0.009		0.123	0.009
Thr-a (d1)	OAA	123x	0.022	0.009		0.023	0.009
Thr-a (d2)	OAA	123x	0.103	0.009		0.101	0.009
Thr-a (dd)	OAA	123x	0.754	0.009		0.752	0.009
Thr-b (s)	OAA	x234	0.498	0.010		0.488	0.014
Thr-b (d)	OAA	x234 + x234	0.417	0.010		0.424	0.014
Thr-b (t)	OAA	x234	0.085	0.010		0.088	0.014
Thr-g (s)	OAA	xx34	0.811	0.015		0.810	0.021
Thr-g (d)	OAA	xx34	0.188	0.015		0.190	0.021
Tyr-a (s)	PEP	123	0.112	0.013		0.134	0.013
Tyr-a (d1)	PEP	123	0.005	0.013		0.002	0.013
Tyr-a (d2)	PEP	123	0.010	0.013		0.003	0.013
Tyr-a (dd)	PEP	123	0.873	0.013		0.862	0.013
Tyr-b (s)	PEP	x23.x2x	0.516	0.011		0.495	0.011
Tyr-b (d)	PEP	x23.x2x + x23.x2x	0.414	0.011		0.413	0.011
Tyr-b (t)	PEP	x23.x2x	0.075	0.011		0.093	0.011
Tyr-d (s)	PEP/E4P	0.5{x23.1xxx+x2x.xx34}	0.520	0.012		0.520	0.012
Tyr-d (d)	PEP/E4P	0.5{x23.1xxx+x2x.xx34+x23.1xxx+x2x.xx34}	0.435	0.012		0.435	0.012
Tyr-d (t)	PEP/E4P	0.5{x23.1xxx+x2x.xx34}	0.044	0.012		0.044	0.012
Tyr-e (s)	E4P	0.5{12xx.xx3+ x234}	0.391	0.009		0.391	0.009
Tyr-e (d)	E4P	0.5{12xx.xx3+ x234 + x234 + 12xx.xx3}	0.175	0.009		0.175	0.009
Tyr-e (t)	E4P	0.5{12xx.xx3+ x234}	0.434	0.009		0.434	0.009
Val-a (s)	PYR	12x.x2x	0.361	0.022		0.389	0.024
Val-a (d1)	PYR	12x.x2x	0.525	0.022		0.469	0.024
Val-a (d2)	PYR	12x.x2x	0.049	0.022		0.072	0.024
Val-a (dd)	PYR	12x.x2x	0.065	0.022		0.070	0.024
Val-g1 (s)	PYR	x23	0.545	0.028		0.549	0.015
Val-g1 (d)	PYR	x23	0.455	0.028		0.451	0.015
Val-g2 (s)	PYR	xx3.x2x	0.883	0.028		0.884	0.015
Val-g2 (d)	PYR	xx3.x2x	0.117	0.028		0.116	0.015

[a]Isotopomer abundances were measured from multiplets of cross-peaks of proteinogenic amino acids on [13C, 1H] HSQC spectra and are reported as fractions of the total signal of the corresponding cross-peak. Standard deviations were derived from measured signal:noise ratios of the NMR spectra. Isotopomers are grouped and color-coded by the metabolic precursors from which they originate. Multiplets abbreviations are as described in the manuscript.

Supplementary Table 3. Probabilistic expressions developed to calculate the relative intensities of ^{13}C multiplet components in biosynthetically directed ^{13}C-labeled amino acids.

K	Expression for different carbon atoms	
	Terminal carbon atom	
$\mathbf{K^{(1)}}$	$K_S^{(1)} = \left[(1-x)(1-P_n)\right]$	$K_d^{(1)} = \left[x + (1-x)P_n\right]$
$\mathbf{K^{(2)}}$	$K_S^{(2)} = \left[\frac{y(1-P_n) + (1-x-y)P_n(1-P_n)}{x+y+(1-x-y)P_n}\right]$	$K_d^{(2)} = \left[\frac{x + yP_n + (1-x-y)P_n^2}{x+y+(1-x-y)P_n}\right]$
	Central carbon atoms	
$\mathbf{K^{(1)}}$	$K_S^{(1)} = \left[\frac{D\{(1-x-y)(1-P_n)\}(1-D)}{D}\right]$	$K_{da}^{(1)} = \left[\frac{D\{x+y+(1-x-y)P_n\}(1-D)}{D}\right]$
	$K_{db}^{(1)} = \left[\frac{D^2\{(1-x-y)(1-P_n)\}}{D}\right]$	$K_{dd}^{(1)} = \left[\frac{D^2\{x+y+(1-x-y)P_n\}}{D}\right]$
$\mathbf{K^{(2a)}}$	$K_S^{(2a)} = \left[\frac{(1-x-y)(1-P_n)P_n(1-D)}{D}\right]$	$K_{da}^{(2a)} = \left[\frac{\{x+yP_n+(1-x-y)P_n^2\}(1-D)}{D}\right]$
	$K_{db}^{(2a)} = \left[\frac{(1-x-y)(1-P_n)P_nD}{D}\right]$	$K_{dd}^{(2a)} = \left[\frac{\{x+yP_n+(1-x-y)P_n^2\}D}{D}\right]$
$\mathbf{K^{(2b)}}$	$K_s^{(2b)} = \left[\frac{(1-x-y)(1-P_n)P_n(1-D)}{D}\right]$	$K_{da}^{(2b)} = \left[\frac{\{x+y+(1-x-y)P_n\}(1-D)P_n}{D}\right]$
	$K_{db}^{(2b)} = \left[\frac{\{x+(1-x)P_n^2\}(1-x-y)(1-P_n)}{D}\right]$	$K_{dd}^{(2b)} = \left[\frac{\{x+(1-x)P_n^2\}\{x+y+(1-x-y)P_n\}}{D}D\right]$
$\mathbf{K^{(3)}}$	$K_3^{(S)} = \left[\frac{(1-x-y)P_n(1-P_n)^2}{D}\right]$	$K_{da}^{(3)} = \left[\frac{P_n(1-P_n)\{y+(1-x-y)P_n\}}{D}\right]$
	$K_{db}^{(3)} = \left[\frac{(1-x-y)P_n^2(1-P_n)}{D}\right]$	$K_{dd}^{3} = \left[\frac{x+yP_n^2+(1-x-y)P_n^3}{D}\right]$
	with $D = x + (1-x)P_n$	

The fractions of glucose isotopomers are represented as x (fraction of U-^{13}C labeled glucose), y (fraction of 1-^{13}C labeled glucose), and 1-x-y (fraction of unlabeled glucose). The unlabeled carbon atoms exhibit natural ^{13}C labeling with a probability P_n (Szypreski, 1995). A central carbon atom (*cent*) can exhibit nine peaks corresponding to four relative isotopomer intensities I_s, I_{da}, I_{db} and I_{dd} (Szypreski, 1995). These arise from a singlet (*s*), a doublet split by a small coupling constant (*da*), a doublet split by a larger coupling constant (*db*), and a doublet of doublets (*dd*). A terminal carbon atom (*term*) can exhibit two relative intensities, I_s and I_d. A vector **I** can be defined such that $\mathbf{I}_{term}$ = (I_s, I_d) and $\mathbf{I}_{cent}$ = (I_s, I_{da}, I_{db}, I_{dd}). The vector $\mathbf{K^i}$, on the other hand, denotes the relative intensities of multiplet components (with *i* denoting the number of intensities). $\mathbf{K^i}$ is defined as $\mathbf{K^i} = (K^i_s, K^i_s)\langle i = 1, 2\rangle$ for a terminal carbon and $\mathbf{K^i} = (K^i_s, K^i_{da}, K^i_{db}, K^i_{dd})\langle i = 1, 2a, 2b, 3\rangle$ for a central carbon (Szypreski, 1995).

- **Szypreski, T**. 1995. Biosynthetically directed fractional 13 C labeling of proteinogenic amino acids. Eur. J. Biochem. 232:433-448.

Biotechnology Progress

Design of 13C Labeling Experiments to Estimate Metabolic Fluxes in Escherichia coli During the Anaerobic Fermentation of Glucose

Journal:	*Biotechnology Progress*
Manuscript ID:	BTPR-10-0131
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	04-Apr-2010
Complete List of Authors:	Choudhary, Madhuresh; Iowa State University, Chemical and Biological Engineering Yoon, Jong Moon; Iowa State University, Chemical and Biological Engineering Gonzalez, Ramon; Rice University, Chemical and Biomolecular Engineering Shanks, Jacqueline; Iowa State University, Chemical and Biological Engineering
Keywords:	Escherichia coli, metabolic flux analysis, anaerobic, identifiability

Design of ^{13}C Labeling Experiments to Estimate Metabolic Fluxes in *Escherichia coli* During the Anaerobic Fermentation of Glucose

Madhuresh K. Choudhary and Jong Moon Yoon

Dept. Chemical and Biological Engineering, Iowa State University, Ames, IA 50011

Ramon Gonzalez

Dept. of Chemical and Biomolecular Engineering, Rice University, Houston, TX 77251

Dept. of Bioengineering, Rice University, Houston, TX 77251

Jacqueline V. Shanks

Dept. Chemical and Biological Engineering, Iowa State University, Ames, IA 50011

Corresponding author: Jacqueline V. Shanks

3031 Sweeney Hall, Iowa State University, Ames, Iowa 50011

E-mail: jshanks@iastate.edu

Abstract

Metabolic fluxes are an important physiological characteristic, providing a global perspective of the integrated functioning between levels of transcripts, proteins, and metabolites to cellular phenotype. Comprehensive metabolic flux maps for *Escherichia coli* under anaerobic conditions were determined by using a mixture of differently labeled glucose and compared to conventional flux maps and comprehensive metabolic flux maps obtained with using only U-^{13}C glucose as the substrate. As expected, conventional flux analysis performs poorly in comparison to ^{13}C-MFA, especially in the elucidation of carbon partitioning between the Embden-Meyerhof-Parnas (EMP) and pentose phosphate (PP) pathways. An identifiability analysis indicated that a mixture of 10% U-^{13}C glucose, 25 % 1-^{13}C glucose, and 65% naturally labeled glucose would significantly improve the statistical quality of calculated fluxes over other labeling schemes. Indeed, experimentally the statistical quality of all fluxes in the PP pathway, the EMP pathway, the anaplerotic reactions, and the tricarboxylic acid cycle were improved. The effect of network topology was studied by investigating the distribution of metabolic fluxes in the presence and absence of the Entner-Doudoroff pathway, the malic enzyme, and the glyoxylate shunt. These changes did not affect the value or quality of estimated fluxes in a significant way. Another aspect investigated was the possibility of estimating intracellular fluxes from labeling data alone. While the combined acetate-ethanol flux can be estimated from the labeling information, the fluxes around the formate node cannot be estimated in the absence of a formate measurement.

Keywords: *Escherichia coli*, identifiability, anaerobic, metabolic flux analysis,

Introduction

Metabolic flux analysis (MFA) has become an important tool in cellular physiology and metabolic engineering as it allows the quantification of steady state intracellular fluxes in a metabolic network. [1-3] Flux measurements and changes in the distribution of metabolic fluxes in response to genetic and environmental perturbations contribute to elucidating the contribution of various pathways in cellular metabolism and can support the design of metabolic engineering strategies.[4]

The classical approach of analyzing intracellular carbon fluxes is metabolite balancing [5] and we refer to it here as conventional metabolic flux analysis (c-MFA). c-MFA is based on mass balances around intracellular metabolites (which are considered in pseudo steady state) with the measurements of extracellular fluxes acting as constraints for flux calculation. Frequently, the lack of enough measurements requires assumptions about redox (NADH/NADPH) or energy (ATP) balances. However, incomplete knowledge about pathways involving NADH/NADPH or ATP (which is very common as these cofactors are involved in a very large number of reactions) can lead to incorrect flux estimation. Moreover, c-MFA cannot account for parallel metabolic pathways, metabolic +cycles, and reversible or bidirectional reactions.[6]

The use of ^{13}C-labeled substrates provides additional constraints to the stoichiometric equations used in c-MFA, avoiding assumptions about redox and energy balances and potentially accounting for parallel pathways, cycles, and reversibility.[5] In this approach, a mixture of a specifically ^{13}C-labeled substrate and a naturally abundant version of the same substrate are fed to the organism of interest and the ^{13}C enrichments or isotopomer distributions in the carbon

atoms of different metabolites are measured. The most common metabolites used for this purpose are proteinogenic amino acids because they are abundant, stable, and their labeling pattern reflects that of precursors metabolites generated in central metabolism. In one approach introduced by Szyperski,[7] the biosynthetically directed fractional ^{13}C labeling of proteinogenic amino acids is measured through 2-D [^{13}C, ^{1}H] HSQC (Heteronuclear Single-Quantum Coherence) or COSY (Correlation Spectroscopy) NMR experiments. Probabilistic equations are then used to relate the observed multiplet intensities of the ^{13}C fine structures to the relative abundance of the intact carbon fragments and are very useful in the quantitative study of intermediary metabolism.[7,8] An extended version of this approach, called metabolic flux ratio (MetaFoR) analysis, was later introduced by Szyperski and co-workers.[9] MetaFoR aims at obtaining relative local fluxes around a node from the abundances of intact carbon fragments in metabolites (calculated from the aforementioned 2-D [^{13}C, ^{1}H] NMR data).[10,11] Comprehensive ^{13}C-based metabolic flux analysis (^{13}C-MFA), which cannot be performed with MetaFoR, is achieved through a modeling approach that requires information on the metabolic network, labeling patterns of amino acids, and extracellular fluxes. This information is combined in an error function that accounts for the average difference between measured and simulated labeling patterns. An iterative, optimization procedure is followed to solve for intracellular fluxes that minimize this error function.[5,12]

While ^{13}C-MFA has been extensively used to study the metabolism of wild-type and recombinant *E. coli* growing under aerobic conditions,[13-22] only a handful of studies has examined fermentative metabolism.[7,17,23] Among the latter, Szyperski[7] used biosynthetically directed fractional ^{13}C labeling of proteinogenic amino acids to analyze important metabolic branch points under anaerobic conditions. The NMR data were used to calculate the bond

integrity of precursor molecules that were in turn used to estimate flux ratios. Using the same NMR data, Schmidt *et al.*[23] carried out the first comprehensive ^{13}C-MFA where isotopomer balances were used in conjunction with constrains from extracellular fluxes. However, the extracellular measurements used in the flux estimation criterion were obtained from a different study conducted with unlabeled glucose.[24] Using the MetaFoR approach described above, Sauer *et al.*[17] reported the analysis of flux ratios at a few nodes under anaerobic conditions.

None of the aforementioned studies examined the appropriate labeling required to identify metabolic fluxes during the anaerobic fermentation of glucose in *E. coli*. In this work, we report the design of ^{13}C labeling experiments, including identifiability analysis, to estimate metabolic fluxes in *E. coli* during the fermentative metabolism of glucose. We found that the use of 1-^{13}C- and U-^{13}C-labeled glucose in combination with extracellular measurements yielded the most reliable estimate of intracellular fluxes.

Materials and Methods

Strain, medium and culture conditions

Escherichia coli K12 strain W3110 (ATCC 27325) was used throughout the study. The minimal media[25] with 1% glucose was used. Ten-time-concentrated media solution was prepared and stored at -20 ^{o}C after filter sterilization. Chemicals were obtained from Fisher Scientific (Pittsburgh, PA) and Sigma-Aldrich Co. (St Louis, MO), except 13C-labeled glucose, which was obtained from Cambridge Isotope Ltd (Andover, MA).

Fermentations were conducted in a 1L bioreactor (Bioflow 110, New Brunswick Scientific, Edison, NJ) with a working volume of 0.9 L and independent control of temperature

(37 °C), pH (6.8), and stirrer speed (200 r.p.m.). The system is fully equipped and computer controlled using manufacturer BioCommand software. A condenser was installed and operated at 4 ^{o}C to minimize evaporation out of the bioreactor vessel. The pH was controlled by adding 4M KOH and anaerobic conditions were maintained by flushing the headspace with high purity nitrogen (Airgas North Central, Des Moines, IA).

Pre-cultures used to inoculate the above fermenters were prepared as follows. A single colony was used to inoculate two 50 mL conical tubes (BD Biosciences, San Jose, CA) completely filled with minimal medium supplemented with 10 g/L of glucose. The tubes were placed in a rotator and kept at 37 °C until a cell density of 0.6 OD_{550} (Genesys 20, Thermo Scientific, MA, USA) was reached. This actively growing pre-culture was centrifuged at 5000g for 15 min at 4 ^{o}C and the pellet resuspended in minimal medium to inoculate the fermenter with a target starting OD_{550} of 0.05.

Analytical methods

Optical density was measured at 550 nm in a Genesys 20 spectrophotometer (Thermo Scientific, MA, USA) and used as an estimate of cell mass (1 $O.D._{550}$ = 0.36 g dry weight/L). After centrifugation, the supernatants were stored at -20 °C for HPLC (High Performance Liquid Chromatography) analysis (glucose and fermentation products) using a Waters HPLC system (Milford, MA) with a 410 refractive index (RI) detector. The Aminex column (HPX-87H, Bio-Rad, Hercules, CA, USA) was maintained at 42 ^{o}C and 5 mM H_2SO_4 was used as the mobile phase at a flow rate of 0.3 mL/min.

Sample preparation for NMR experiments

Experiments with ^{13}C-labeled glucose were conducted to assess the incorporation of this carbon source into proteinogenic biomass, information that was then used to calculate isotopomer abundance and estimate intracellular fluxes (see next sections). Two experiments were carried out, one with 10% U-^{13}C glucose and 90% naturally labeled glucose and another one with 10% U-^{13}C glucose, 25% 1-^{13}C glucose and 65% naturally labeled glucose. Cultures grown as described in the previous section were harvested when the OD_{550} was 0.6 (mid exponential phase) and kept in an ice-water bath. The cells were centrifuged at 5000g for 15 min at 4˚C and the pellets washed with 0.9% saline water. An appropriate amount of the pellet was transferred to hydrolysis tubes (Pierce Endogen, Rockford, IL), to which 6 N hydrochloric acid was added (1 mL of HCl per 4 mg of biomass). The hydrolysis was performed at 110°C for 12 hours after flushing the tubes with nitrogen. The acid in the protein hydrolysates was evaporated in a Rapidvap evaporator (Labconco, Kansas City, MO). The residue was reconstituted in 2 mL of deionized water, lyophilized for 72 h, and dissolved in 500 µL D_2O in an NMR tube. The pH of the NMR sample was adjusted to 1 using DCl.

NMR experiments and calculation of isotopomer fractions

Samples prepared as described above were analyzed via NMR spectroscopy to determine the labeling pattern of proteinogenic amino acids. Two-dimensional Heteronuclear Single-Quantum Coherence [^{13}C,^{1}H]-correlation (2D ^{1}H-^{13}C HSQC) NMR spectra[9,26,30] were acquired on a Bruker Avance DRX 500 MHz spectrometer (Bruker Instruments, Billerica, MA) at 298 K. The reference to 0 ppm was set using the methyl signal of dimethylsilapentane sulfonate (Sigma, St. Louis) as an internal standard. The resonance frequency of ^{13}C and ^{1}H were 125.7 MHz and 499.9 MHz, respectively. The spectral width was 5,482.26 Hz along the ^{1}H (F2) dimension and

5,028.05 Hz along the ^{13}C (F1) dimension. Peak aliasing was used in order to minimize the sweep width along the F1 dimension. The number of complex data points was 1,024 (^{1}H) x 900 (^{13}C). A modification of the INEPT (insensitive nuclei enhanced by polarization transfer) pulse sequence was used for acquiring HSQC spectra. The number of scans was generally set to 16.

The software Xwinnmr (Bruker Instruments, Billerica, MA) was used to acquire all spectra, and the software NMRView[27] was used to quantify nonoverlapping multiplets on the HSQC spectrum. Overlapping multiplets (α-amino acids), which could not be processed with NMRView, were quantified by a previously developed peak deconvolution software[28] that is based on the spectral processing algorithm proposed by Van Winden *et al.*[29] The standard deviations associated with the NMR intensity measurements were estimated from the noise to peak intensity ratio with minimum set to 1%. The resulting intensities were used to calculate the isotopomer fractions shown in Supplementary Table 4. Isotopomer fractions, in turn, represent the key input used in the calculation of metabolic fluxes as described in the next section.

Metabolic flux analysis (MFA)

The metabolic network used in the Metabolic Flux Analysis (MFA) is shown in Figure 1 and Supplementary Table 1. MFA was conducted using two different approaches. First, intracellular fluxes were estimated based on network stoichiometry and extracellular measurements using the technique of metabolite balancing,[1] which we refer to here as "conventional" MFA (c-MFA). The metabolic network consists of reactions involved in the transport and phosphorylation of glucose via the (PEP)-dependent phosphotransferase system (PTS), the Embden-Meyerhof-Parnas (EMP) and Pentose Phosphate (PP) pathways, anaplerotic

reactions, reductive and oxidative branches of the tricarboxylic acid (TCA) cycle, along with fermentative and biosynthetic reactions (Figure 1) (see also Supplementary Table 1). The stoichiometric model used for c-MFA consisted of 37 reactions (fluxes) and 20 balanceable metabolites (rows in Supplementary Table 2), thus resulting in a system with 17 degrees of freedom. The fermentation data was used to calculate 17 extracellular fluxes (i.e. specific rates) associated with biomass synthesis, consumption of glucose, and synthesis of fermentation products (columns in blue color in Supplemenary Table 2). The availability of these 17 extracellular fluxes made the system determined and allowed the calculation of intracellular fluxes without the use of an optimization routine. Three independent pairs of measurements were used in the calculations, thus allowing the estimation of standard deviations.

A second MFA technique, based on the use ^{13}C labeled substrate(s), NMR analysis and isotopomer balancing, was also employed to estimate intracellular fluxes (referred to here as ^{13}C-based MFA or ^{13}C-MFA). The metabolic network was similar to that described above for c-MFA but now accounted for reversibility of reactions, cyclic nature of the TCA cycle, and included the glyoxylate shunt and other reactions as described below (see Supplementary Table 3 and Figure 1). The carbon fate of precursors leading to proteinogenic amino acids was established based on the work of Szyperski.[7] The synthesis of serine from 3-phospho-glycerate and the one carbon metabolism of serine to glycine were also included in the model. Fermentation reactions leading to acetate and ethanol from acetyl-CoA (Figure 1) were combined as they lead to similar carbon rearrangement. Triose phosphates were considered as a single metabolite pool (G3P). Since a high exchange between ribose-5-phosphate and xylose-5-phosphate was observed, a single pentose phosphate pool (R5P) was assumed. The reactions leading to the oxidative pentose pathway (ox-PPP, glucose-6-phosphate dehydrogenase) and the TCA cycle (citrate synthase)

were considered irreversible with no negative flux allowed through them. The reactions through phosphoglucose isomerase and enolase in the EMP pathway and transketolases and transaldolases in the PPP (Figure 1) were considered reversible. The flux though pyruvate formate lyase was also assumed to be reversible. Since succinate is a symmetric molecule, the scrambling reaction was also included in the model. The forward (V_1) and backward (V_{-1}) fluxes associated with each reversible reaction step were transformed into a net flux (V_{net}) and extent of reversibility (r)

$$V_{net} = V_1 - V_{-1} \qquad r = \frac{\min(V_1, V_{-1})}{\max(V_1, V_{-1})} \qquad (1)$$

In order to avoid numerical problems, the extents of reversibility were constrained between 0 and 0.99.

The flux of precursor metabolites toward biomass was estimated based on measurements of biomass yield and literature data on biomass composition (Table 1). The following were chosen as free fluxes: glucose uptake, production of lactate and succinate, oxidative pentose pathway (ox-PPP), glyoxylate shunt pathway, and biomass synthesis. The extents of reversibility were also considered as free parameters. The pools of intracellular metabolites were assumed to be in isotopic steady state and the dilution effect in HSQC labeling measurement due to initial unlabeled biomass was considered negligible.

Estimation of intracellular fluxes via ^{13}C-MFA required modifying the computer program NMR2Flux (originally developed to estimate metabolic fluxes in plants[28,30]) to a generic form that uses as input metabolic network information from any system (i.e. reaction stoichiometries and carbon skeleton rearrangements). The other input to the software are isotopomer abundances and extracellular flux and biomass composition data, which were obtained as described in

previous sections. NMR2Flux estimates fluxes by minimizing the difference between simulated and experiment NMR intensities. Free fluxes are first guessed and used to calculate all intracellular fluxes using stoichiometric balances. The calculated set of fluxes allows estimating labeling patterns for proteinogenic amino acids (i.e. simulated intensities, I_{sim}). A chi-square (χ^2) is then calculated for the difference between simulated and experimental intensities (I_{sim} and I_{exp}, respectively). The extracellular flux measurements (F_{mes}) of acetate-ethanol and formate were also included in the χ^2.

$$\chi^2 = \frac{(I_{sim} - I_{\exp})^2}{N_{\exp}^{\ 2}} + \frac{(F_{mes} - F_{sim})^2}{N^2} \qquad (2)$$

The set of fluxes that gives minimum χ^2 is taken as the best estimate of the metabolic fluxes. To verify the global error minimum, multiple simulations were carried out from different starting points. A statistical error analysis was performed by using a Monte Carlo simulations approach[31] in which synthetic NMR intensities were used as surrogate for experimental data. Finally, the set of 100 flux distributions obtained by Monte Carlo simulation were used to calculate standard deviations for the fluxes.

Identifiability analysis

An identifiability analysis was conducted to determine the effect of substrate labeling on the statistical identifiability of the fluxes[32,33]. This analysis used the information content (*IC*) as the objective criterion for identifiability. *IC* is defined as the reciprocal of the geometric mean of the standard deviation (*SD*) of the fluxes. The geometric mean of the *SD* is the nth root of *D* ($\sqrt[n]{D}$), where n is number of flux parameters and *D* is the D-criterion that measures the volume of the confidence ellipsoid of the evaluated flux parameters and which is equal to the

determinant of the covariance matrix of the flux parameters. Expressions for *IC* and *D* are as follows:

$$IC = \frac{1}{\sqrt[n]{D}} \qquad D = \det\left(Cov\left(P\right)\right) \qquad (3)$$

The free flux parameters (*P*) and the NMR intensities (*I*) are related through a non-linear relationship of the form $I = h(P)$

The NMR2Flux software previously developed in our group evaluates flux parameters iteratively from the labeling data. All computations of *IC* are reported with respect to a reference experiment with 10% U-^{13}C glucose as the only labeled substrate. Various combinations of U-^{13}C glucose, 1-^{13}C glucose and naturally labeled glucose were examined for their ability to provide an improved labeling data set.

Results and Discussion

Metabolic fluxes calculated using conventional metabolic flux analysis

The fermentation data for the anaerobic growth of strain W3110 on glucose was used to obtain the measured fluxes shown in Table 1. Acetate, ethanol, succinate, and lactate were the major fermentation products. Ethanol and acetate production was high compared to other fermentation products, as this partition of carbon is known to be favorable to support redox balance and generation of ATP in the absence of external electron acceptors.[34] Metabolite balancing, referred to here as conventional metabolic flux analysis (c-MFA), was then used to calculate the intracellular fluxes by making use of the measured fluxes (Table 1) and the stoichiometric model described in Supplementary Table 3. The following assumptions were made in constructing the model. The TCA cycle was assumed to be incomplete, and thus

operating as two (oxidative and reductive) branches.[35] The glyoxylate shunt pathway has also been reported inactive during glucose metabolism due to repression by this carbon source [13,36] and hence was not considered in the model. Fluxes are considered to be net fluxes because c-MFA cannot account for reaction reversibility. These assumptions yielded an exactly determined metabolic model. The condition number of the stoichiometric matrix consisting of mass balances of intracellular metabolites was found to be 8 (Supplementary Table 2), indicating a well-conditioned matrix.[1] Measurements of glucose utilization, synthesis of fermentation products, and biomass formation (Table 1) were used to calculate the intracellular fluxes, with a selected group of them shown in Table 2 (see Supplementary Table 3 for all calculated fluxes). The fluxes are reported relative to 100 moles of glucose and standard deviations were estimated using a Monte Carlo simulation approach as described in *Materials and Methods*.

The flux through pyruvate formate lyase (*pfl*), an enzyme that catalyzes the conversion of pyruvate to acetyl-CoA and formate, was found to be 144.8±10.5 (Table 2) suggesting very high activity under anaerobic conditions. Most of the produced formate was secreted to the extracellular medium (135.9±3.8), with only about 6% converted to carbon dioxide and hydrogen (8.9±11.2) by the action of formate hydrogenlyase (*fhl*) (Table 2). The combined flux for the conversion of acetyl-CoA-to-ethanol and acetyl-CoA-to-acetate was found to be 133.9±10.5. The flux through the other major pathway consuming pyruvate (i.e. conversion to lactate through *ldh*) was much lower (11.6±4.6). The higher flux through *pfl* is advantageous as the production of equimolar amounts of ethanol and acetate from glucose is the most ATP efficient anaerobic mode producing three molecules of ATP per molecule of glucose fermented in a redox balanced manner.[34] The succinate flux was found to be very small (5.5±0.8).

While most fluxes in the Embden-Meyerhof-Parnas (EMP) pathway exhibited small standard deviations (11.2), the standard deviation of the first step (catalyzed by the enzyme phosphoglucose isomerase) was about three times larger (33.7) (Table 2). This resulted in a very large coefficient of variation (CV = standard deviation/average × 100) of 61% for the phosphoglucose isomerase flux (*pgi*). A substantial flux was calculated for the pentose phosphate pathway (PPP) (Table 2), suggesting that this pathway may be active during anaerobic fermentation of glucose, which is consistent with previous findings.[7,17,23] However, all fluxes in the PPP, including that through its oxidative branch (ox-PPP, flux *zwf*: conversion of glucose-6-phosphate to 6-phosphogluconolactone, catalyzed by glucose-6-phosphate dehydrogenase), exhibited large standard deviations (Table 2) with average CV of 78.7 % (Supplementary Table 3). When taken together, the above results indicate a poor resolution of metabolic fluxes at glucose-6-phosphate, a key metabolic node that determines the partition of carbon flux between EMP and PPP.

The other group of reactions with large error in estimated fluxes were associated with fermentative pathways, primarily those involved in the conversion of formate to carbon dioxide and hydrogen (catalyzed by the enzyme FHL, *fhl* flux) and the carbon dioxide evolution flux (Table 2 and Supplementary Table 3). As in the case of glucose-6-phosphate, this result indicates poor resolution of flux partition at the formate node.

Metabolic flux analysis using uniformly (U)-^{13}C-labeled glucose

^{13}C labeling data provide additional measurements that can be used to extend the analysis conducted with c-MFA, thus obtaining a more comprehensive characterization of metabolic fluxes and network topology.[5,12] To this end, a ^{13}C labeling experiment was carried out using

10% uniformly (U)-labeled-^{13}C glucose and 90% naturally labeled glucose (referred to as U-^{13}C experiment). Exponentially growing cells were harvested and the protein extracted and hydrolyzed. A two-dimensional [^{1}H, ^{13}C] Heteronuclear Single-Quantum Coherence (HSQC) spectrum of the proteinogenic amino acids thus obtained was acquired (Figure 2A). Carbon atoms of 13 amino acids were identified by their unique ^{13}C/^{1}H chemical shifts, distinctive coupling patterns, and J-coupling constants (J_{CC}).[30, 37] Upon quantification of peak integrals 91 peak intensities from 31 sets of relative isotopomer abundances were obtained, corresponding to 31 observed carbon atoms of proteinogenic amino acids, as shown in Supplementary Table 4.

The reactions corresponding to the TCA cycle and the glyoxylate shunt were now included in the stoichiometric model and reaction reversibility was also accounted for. The new model contains 40 net reactions, 11 reversible reactions and 1 scrambling reaction: i.e. a total of 52 fluxes to be estimated (Supplementary Table 3). Assuming pseudo steady state, 21 intracellular metabolite balances contribute to 21 linear constraints. Hence, the model has 31 parameters including 11 reversibility parameters, 1 scrambling parameter, and 19 independent flux parameters. The fluxes corresponding to the incorporation of 12 precursor metabolites into biomass were estimated using data on biomass yield and biomass composition reported in the literature (Table 1). The extracellular fluxes of glucose, lactate and succinate were obtained from the experimental measurements. The extracellular fluxes of formate and the combined flux of acetate-ethanol were included in the chi-square (χ^2) criterion for the estimation of fluxes.

The software NMR2Flux [28,30] was used to obtain a new set of intracellular fluxes (Table 2, U-^{13}C-MFA column) based on the above-described model, isotopomer abundances, and extracellular and biomass flux measurements. In general, there was a good fit between the simulated and experimental measurements of isotopomer abundances (Figure 2B) and calculated

and measured formate and acetate-ethanol fluxes (Tables 1 and 2). The total χ^2 was 650 (Table 2) with an average difference between simulated and experimental intensities of 0.019. Out of 91 NMR measurements, 27 measurements from asp-α, ile- α, phe- α, leu- α, tyr-β and tyr-δ contributed to 60% of the total χ^2. Since most of these peaks were analyzed by spectral deconvolution, their high contribution to χ^2 is most likely due to the low standard deviations assumed for these peaks rather than an inappropriate metabolic model.

The most salient features of the calculated fluxes are as follows. A very low flux through 2-oxoglutarate dehydrogenase (0.57 ± 0.96; CV = 170: Supplementary Table 3) indicates that the TCA cycle operates as two branches to fulfill demand for precursor metabolites for biomass synthesis, in agreement with previous studies.[7,17,23] The relative flux through the oxidative branch of the PPP (ox-PPP) was found to be 40±30 (Table 2). Using semi-quantitative NMR analysis, Szyperski *et al.*[7] estimated that 20 to 30% of glucose is converted to PEP via the PPP in *E. coli* B, which appears to be similar to the value calculated here via U-^{13}C-MFA. Additionally, they found that less than 20% of R5P originates from G6P via ox-PPP. However, it is not possible to compare these results directly with the net fluxes obtained by c-MFA or ^{13}C-MFA because rapid equilibration of pentose pool in addition to rapid exchange via transketolase and/or transaldose can lead to similar carbon labeling pattern for various intracellular flux distributions. Using the same NMR data and a more comprehensive ^{13}C flux analysis, Schmidt *et al.* [23] found the flux through ox-PPP to be 77%. However, the extracellular flux measurements for the fermentation products were taken from a different study in the literature.[24] Moreover, the demand for NADPH and precursor metabolites generated by PPP (and used in biomass synthesis) should be low under fermentative conditions due limited cell growth (about 90% of

the carbon is recovered as fermentation products: see Table 1). Hence, a low flux through PPP appears more reasonable.

The glyoxylate shunt is often assumed to be inactive in glucose grown cultures as this pathway is subjected to catabolite repression by glucose.[36] However, the glyoxylate shunt was found to be active in wild-type *E. coli* under conditions of glucose hunger in a slow-growing continuous culture[22] and under glucose-excess batch conditions in a phosphoglucose isomerase mutant.[10,18] Moreover, while repressed by glucose in K12 strains, the glyoxylate cycle appears to be expressed in *E. coli* B during growth on glucose.[38,39] Although comprehensive, all the aforementioned studies investigated the metabolism of glucose under aerobic conditions. Our U-^{13}C-MFA indicates that the flux through the glyoxylate shunt during the anaerobic fermentation of glucose is very low (2.34) with a large standard deviation (1.48) that reaches 63% of the calculated value (Supplementary Table 3). This result appears to indicate that the glyoxylate shunt is unlikely to be active during fermentative growth of *E. coli* W3110 on glucose.

Identifiability analysis

While the results of U-^{13}C-MFA were superior to those obtained with c-MFA (see previous sections), the estimated fluxes for PPP and some fermentative reactions still exhibited large standard deviations (Table 2). Consequently, the flux split ratio at several nodes such as the glucose 6-phosphate and formate were not well resolved. These findings indicate that the labeling measurements from the U-^{13}C experiment do not provide enough information to estimate the above-mentioned fluxes. The large standard deviations associated with the ox-PPP flux are in agreement with previous observations by Dauner *et al.*[40] who found large confidence interval for ox-PPP flux in their MFA of aerobic metabolism of *B. subtilis* using U-^{13}C glucose.

However, Schmidt *et al.*[41] were able to accurately estimate the ox-PPP flux using a mixture of 1-^{13}C glucose and 6-^{13}C glucose to characterize the aerobic metabolism of glucose in a glucoamylase-producing strain of *Aspergillus niger*. In general, it has been well documented in the literature that a statistical analysis is required to identify the best mixture of labeled carbon that supports good estimates of intracellular fluxes.[12,21,32,42] Specifically, the fluxes in the PPP and the flux split ratio at the glucose-6-phosphate node can be well resolved using 1st position labeled glucose.[21,32,41,42] Therefore, we conducted an identifiability analysis to determine the impact of using 1-^{13}C-labeled glucose in combination with U-^{13}C-labeled and naturally labeled glucose on flux identifiability compared to the flux values estimated from the 10% U-^{13}C labeling experiment. To this end, we used the fluxes obtained via U-^{13}C-MFA and employed identifiability analysis based on linear statistics[32] to obtain synthetic measurement data sets as surrogates for labeling experiments.[23,40,41] Figure 3 depicts the information content (*IC*, which indicates the statistical quality of the experiment) for various combinations of 1-^{13}C- and U-^{13}C-labeled glucose relative to the reference experiment with 10% U-^{13}C-labeled glucose. In computation of *IC*, the statistical quality of all the flux parameters was taken into account. The maximum improvement in *IC* (2.2-fold) was observed with the combination of 5% U-^{13}C glucose and 95% 1-^{13}C glucose. This overall improvement is primarily due to a very large improvement in statistical quality of the ox-PPP flux when 1-^{13}C glucose is used in combination with U-^{13}C glucose as labeled substrate (Figure 4). Using these simulations we determined that a labeled substrate mixture containing 25% of 1-^{13}C and 10% of U-^{13}C glucose (balance naturally labeled glucose) supports adequate estimation of the PPP fluxes along with high overall statistical quality while maintaining the use of labeled substrates at a reasonable level (e.g. less than 40% of the total substrate).

Metabolic flux analysis using a complex mixture of 1-^{13}C and U-^{13}C-labeled glucose

Based on the results of the identifiability analysis presented above, a labeling experiment was carried out with a mixture of 10% U-^{13}C glucose, 25 % 1-^{13}C glucose, and 65% naturally labeled glucose (referred to as 1-U-^{13}C experiment). The flux map thus obtained is shown in Figure 5 (see also Table 2 and Supplementary Table 3). In agreement with the *in silico* analysis, the statistical quality of calculated fluxes was significantly improved. All fluxes in the PPP, including the ox-PPP and backward fluxes in reversible reactions, showed improvements in the 1-U-^{13}C-MFA. The same was observed for the EMP, anaplerotic reactions, and the TCA cycle (Figure 5; see also Table 2 and Supplementary Table 3). The largest improvement in estimated fluxes was observed at the glucose-6-phosphate and formate nodes, which are discussed in detail below.

Glucose-6-phosphate is a very important metabolic node as it determines the distribution of carbon between the PP and EMP pathways by converting glucose-6-phosphate to fructose-6-P (enzyme phosphoglucose isomerase, *pgi* flux) or 6-phosphogluconolactone (enzyme glucose-6-phosphate dehydrogenase, ox-PPP, *zwf* flux). The standard deviation of the flux through phosphoglucose isomerase decreased by 10 times, with an estimated value 1.3 higher in the 1-U-^{13}C experiment (Figure 5 and Table 2). When both changes were taken into account, the coefficient of variation in the 1-U-^{13}C experiment was only 7% of that observed in the U-^{13}C experiment (Table 2). The same decrease in standard deviation was observed for the estimated flux through glucose-6-phosphate dehydrogenase, the committed step of the ox-PPP (Figure 5). In this case the CV decrease to around 80% of its value in the U-^{13}C experiment. When the results for phosphoglucose isomerase and glucose-6-phosphate dehydrogenase are combined, a

very large improvement in flux resolution at the glucose-6-phosphate node is realized: i.e. an average decrease in standard deviation and coefficient of variation of 10- and 2-fold, respectively.

Formate is another important metabolic node under fermentative conditions as this metabolite can be either exported to the extracellular medium via transporters FocA and FocB or disproportionated to carbon dioxide and hydrogen by the action of the enzyme fomate hydrogenlyase (FHL).[34,35] While the net efflux of formate can be estimated based on the measurement of formate accumulated in the extracellular medium, the *fhl* flux cannot be calculated based on measurements of carbon dioxide or hydrogen evolution; the latter due to the involvement of carbon dioxide in many other metabolic pathways and the recycling of hydrogen by the action of hydrogenases.[35] As previously discussed, the estimated *fhl* flux was very poor in c-MFA, with a standard deviation that exceeded the calculated value of the flux (Table 2). While the U-^{13}C-MFA improved the quality of the estimated flux, its coefficient of variation was still very large representing 80% of the calculated flux (Table 2). As in the case of glucose-6-phosphate, the use of a complex mixture of 1-^{13}C- and U-^{13}C-labeled glucose allowed a better resolution of fluxes at the formate node (Figure 5 and Table 2). For example, the coefficient of variation for the *fhl* flux decreased from 79% to 37 %.

The estimated flux distribution was used to calculate the overall redox balance by considering generation and consumption of reducing equivalents in biomass formation, oxidative pentose phosphate pathway, isocitrate dehydrogenase, glyceraldehydes 3-phosphate dehydrogenase, acetaldehyde/alcohol dehydrogenase and lactate dehydrogenase. According to our calculations, the net flux from NADPH to NADP was negative (-14.7), but this is probably

compensated by the action of transhydrogenases, which interconvert NADH and NADPH.[18] The transhydrogenase flux converting NADH into NADPH was therefore estimated to be 14.7.

Unlike the case of aerobic conditions, under fermentative conditions it is possible to estimate metabolic production of ATP without assuming a P/O ratio. The total ATP production in central carbon metabolism was found to be 152.9 mole of ATP per 100 mole of glucose consumed. Several cellular processes require the consumption of ATP for maintenance such as constant electrochemical gradients across membranes, futile cycles, and turnover of macromolecules without net generation of cell biomass. The ATP consumption for maintenance was found to be 128.2 mole ATP per 100 mole glucose consumption.

Topology of the metabolic network and identifiability of extracellular fluxes from labeling data

Since the generic nature of our flux evaluation methodology allows easy modification of the metabolic network, the topology of the network was further investigated. To this end, two general areas of the original network were modified: i) the glyoxylate shunt was excluded and the TCA cycle modified to operate as two independent branches ii) the Entner-Doudoroff (ED) pathway (*ed*) and the malic enzyme reaction (*me*) were added to the network. The exclusion of the glyoxylate shunt and operation of the TCA cycle as two branches did not affect the χ^2 and no significant changes were observed in the resolution of fluxes in the EMP, PPP or fermentative pathways (Table 2). The only significant changes were observed in anaplerotic and TCA cycle fluxes, which were better estimated in this scenario: i.e. the average CV decreased by 40% (data not shown).

When the ED pathway and the malic enzyme reaction were added to the metabolic network, a significant decrease in χ^2 was observed, although the quality of most estimated fluxes remained almost unchanged, as can be judged by their standard deviations (and coefficients of variation) (Table 2 and Supplementary Table 3). In this scenario, the calculated flux for the ED pathway was only 1.3±1.5, which is clearly negligible and statistically unidentifiable (CV = 121%). The flux through the malic enzyme reaction (*me* flux), however, was found to be considerably large (65.5±21.2) and brought about significant changes around the phosphoenolpyruvate-pyruvate node. First, there was a 10-fold increase in the conversion of PEP to OAA to MAL, which is supported by 10-fold increase in the *ppc* and *mdh* fluxes by PEP carboxylase and malate dehydrogenase (Table 2). Along with this, there was a 5-fold decrease in the pyruvate kinase flux (*pyk*). Taken together, this scenario indicates that a large fraction of the pyruvate is generated by the combined action of PEP carboxylase, malate dehydrogenase, and malic enzyme, thus by-passing pyruvate kinase. This three-step conversion of PEP to PYR involves carboxylation and decarboxylation reactions, thus affecting the labeling pattern of amino acids and explaining the better χ^2 (Table 2). However, the high-energy bond of PEP is wasted and no ATP is generated. In contrast, the conversion of PEP to PYR catalyzed by pyruvate kinase generates one molecule of ATP (Figure 1). In conclusion, while the inclusion of *me* flux provides a better fit of the experimental data (i.e. lower χ^2) the calculated fluxes around the phosphoenolpyruvate-pyruvate node do not reflect a metabolically feasible scenario due to its low energy efficiency. That is, the ATP generated without the malic enzyme reaction was 152.9 mole per 100 mole glucose, but upon inclusion of the *me* flux it significantly decreased to 83.9 mole per 100 mole of glucose.

Another aspect investigated was the possibility of estimating intracellular fluxes from labeling data alone (i.e. without including extracellular flux measurements in the χ^2 criterion). In comprehensive ^{13}C-MFA, intracellular fluxes are estimated by using the NMR measurements along with extracellular measurements, both being included in the χ^2 criterion (see *Materials and Methods* for details). The labeling patterns *per se* are dependent on intracellular fluxes only and do not contain information about extracellular fluxes. However, since intracellular fluxes are related to extracellular fluxes through stoichiometric constrains, the accurate determination of intracellular fluxes would result in accurate determination of extracellular fluxes. To test the hypothesis of whether extracellular fluxes can be estimated from labeling measurements alone, formate and ethanol-acetate measurements were not included in the χ^2 criterion and the results are shown in Table 2 (see also Supplementary Table 3). The calculated flux values in ethanol-acetate production are very similar in both cases of inclusion and exclusion of two measurements in the χ^2 criterion (Equation "(2)"). The acetate-ethanol flux was estimated to be 144.4±5 with inclusion of both ethanol-acetate and formate extracellular measurements and it was 146.3±6.9 without inclusion of them. This result indicates that the acetate-ethanol combined flux can indeed be estimated from the labeling information alone.

Unlike the acetate-ethanol flux, the fluxes around the formate node (i.e. conversion of formate to carbon dioxide and hydrogen by FHL, formate export, and carbon dioxide evolution) were all associated with large standard deviations and coefficients of variation (Table 2 and Supplementary Table 3). Moreover, when compared to experimental measurements, the net formate (export) flux was poorly estimated: 78.5±45.6 (Table 2) compared to a measured value of 135.9±3.8 (Table 1). We then conclude that the fluxes around the formate node cannot be estimated in the absence of formate measurement.

To investigate why the fluxes around the formate node are associated with large standard deviations (and why formate production can not be estimated from labeling data alone), a linearized method was used to estimate the effect of the ox-PPP, PEP carboxylase and TCA cycle fluxes on the SD of the formate-related fluxes (i.e. *fhl* flux, formate export, and carbon dioxide evolution) (Figure 6). These reactions/pathways were chosen because of their involvement in carbon dioxide metabolism, which link them to the reactions involved in the formate node. The simulations predict that the TCA and PEP carboxylase fluxes do not have a significant impact on the standard deviation of the formate flux but the ox-PPP flux significantly affected it (Figure 6A).

Theoretically, the fluxes of *fhl* and formate production in the branch pathways from the formate node could be estimated if the enrichment of CO_2 and formate is different. When 10% U-^{13}C glucose and 90% unlabeled glucose were used, the enrichment of the carbon in formate and CO_2 would be the same to 11% (10% from U-^{13}C and 1% from natural abundance), and thus the fluxes in extracellular formate production and *fhl* flux are unidentifiable. On the other hand, using a mixture of 10% U-^{13}C glucose and 25% 1-^{13}C glucose, the carbon enrichment in formate and CO_2 will be different through metabolic pathways. Of the fluxes involved with the formate node, the ox-PPP flux (*zwf*) affected the most the *fhl* flux and the carbon dioxide evolution flux (Figure 6B). Since the flux through ox-PPP affects the fraction of carbon dioxide originating from the first carbon of glucose, higher ox-PPP fluxes result in higher enrichment of carbon dioxide. However, the small flux in *zwf* (3.4±1.3) from 1-U-^{13}C-MFA results in similar enrichment of formate and carbon dioxide as shown in Figure 6B.

Conclusions

Comprehensive metabolic flux analysis using a mixture of differently labeled glucose led to a superior estimation of metabolic fluxes during the fermentation of glucose by *Escherichia coli* when compared to the use of conventional flux analysis or only U-^{13}C glucose as the substrate. An identifiability analysis indicated that a mixture of 10% U-^{13}C glucose, 25 % 1-^{13}C glucose, and 65% naturally labeled glucose would significantly improve the statistical quality of calculated fluxes over other labeling schemes. The most significant improvements were observed for fluxes involved in two metabolic nodes: the glucose-6-P node, which determines carbon partitioning between the Embden-Meyerhof-Parnas and pentose phosphate pathways, and the formate node, which determines the fate of formate between export and oxidation to CO_2 and hydrogen. The study of network topology indicated that the inclusion of the Entner-Doudoroff pathway, the malic enzyme, or the glyoxylate shunt does not significantly affect the value or quality of estimated fluxes. It was also concluded that while the combined acetate-ethanol flux can be estimated from the labeling information alone, the fluxes around the formate node couldn't be estimated in the absence of a formate measurement.

Acknowledgements

This work was supported by grants from the U.S. National Science Foundation (EEC-0813570 and BES-0331388/BES-0601549).

Supplementary Material Statement

Table S1. Metabolic pathways and involved enzymes, EC numbers, and gene names.

Table S2. Stiochiometric matrix (G) of metabolic network of *E.coli* used in conventional metabolic flux analysis

Table S3. Metabolic reactions and flux values quantified by c-MFA and 13C based MFA

Table S4. Simulated and experimentally meausred NMR intensity from U-^{13}C-MFA and 1-^{13}C-MFA

Abbreviations

AcCoA, acetyl coenzyme A; ACK, acetate kinase; ADH, alcohol dehydrogenase; AKG, α-ketoglutarate; ED, Entner-Doudoroff pathway; EMP, Embden-Meyerhof-Parnas; E4P, erythrose-4-phosphate; FHL, formate hydrogen-lyase; FUM, fumarate; F6P, fructose-6-phosphate; G3P, combined pool of triose-3-phosphate; G6P; glucose-6-phosphate; HPLC, high performance liquid chromatography; GOX, glyoxylate; HSQC, Heteronuclear Single-Quantum Coherence; ICIT, isocitrate; MAL, malate; MFA, metabolic flux analysis; non-ox-PPP, non-oxidative branch of the pentose phosphate pathway; OAA, oxaloacetate; ox-PPP, oxidative branch of the pentose phosphate pathway; PEP, phosphoenolpyruvate; PFL, pyruvate formate-lyase; PGLU, 6-phospho-D-gluconate; PGLUL, D-glucono-δ-lactone-6-phosphate; PP, pentose phosphate; PPP, pentose phosphate pathway; PYK, pyruvate kinase; PYR, pyruvate; RL5P, ribulose-5-phosphate; R5P, ribose-5-phosphate; S7P, sedoheptulose-7-phosphate; SUCC, succinate; TCA, tricarboxylic acid; X5P, xylose-5-phosphate

Figure Captions

Figure 1. Pathways involved in the synthesis of fermentation products, precursor metabolites, ATP, and reducing equivalents during the fermentative utilization of glucose by *E. coli*. Enzyme(s) catalyzing shown reaction(s) are as follows. Glucose transport and phosphorylation: [1], phosphoenolpyruvate-dependent phosphotrasnferase system (PTS). EMP: [2], phosphoglucose isomerase; [3], combined reactions by 6-phosphofructokinase, fructose bisphosphate aldolase and triose phosphate isomerase; [4], glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase, phosphoglycerate mutases and enolase; and [5], pyruvate kinase. ox-PPP: [6], glucose 6-phosphate dehydrogenase; [7], 6-phosphogluconolactonase; and [8], 6-phosphogluconate dehydrogenase. non-ox-PPP: [9], ribulose phosphate 3-epimerase; [10], ribose-5-phosphate isomerases; [11], transketolases; and [12], transaldolases. Oxidative and reductive branches of the TCA cycle: [13], citrate synthase and acomitases; [14], isocitrate dehydrogenase; [15], malate dehydrogenase and fumarases; and [16], fumarate reductase. Anaplerotic reaction: [17], phosphoenolpyruvate carboxylase. Pyruvate dissimilation: [18], pyruvate formate-lyase. Fermentation: [19], lactate dehydrogenase; [20], formate hydrogen-lyase; [21], phosphate acetyltransferase; [22], acetate kinase; [23], alcohol/acetaldehyde dehydrogenase. Cell growth: [24], synthesis of cell mass from precursor metabolites (*), ATP, and reducing equivalents. See the detail description in Supplementary Table S1 and S3 and the list of abbreviations.

Figure 2. Two-dimensional [^{13}C, ^{1}H] HSQC spectra of hydrolyzed extracts of wild type *E. coli* K12 strain W3110 grown on a mixture of 10% U-^{13}C and 90% naturally labeled glucose (A) and comparison of experimental and simulated NMR intensities of proteinogenic amino acids (B). The solid line represents a linear fitting as shown in the equation while the short dashed lines illustrate the 95% confidence intervals.

Figure 3. Effect of the use of 1-^{13}C-labeled glucose in combination with U-^{13}C-labeled on flux identifiability as represented by the information content (*IC*). The *IC*, which takes into account the statistical quality of all the flux parameters, is shown relative to the reference experiment [10% U-^{13}C].

Figure 4. Identifiability of the ox-PPP flux for various combinations of 1-^{13}C- and U-^{13}C-labeled glucose, expressed relative to the reference experiment with 10% U-^{13}C-labeled glucose.

Figure 5. *In vivo* distribution of metabolic fluxes for wild-type *E. coli* K12 strain W3110 calculated using 10% U-^{13}C-labeled, 25% 1-^{13}C-labeled, and 65% naturally labeled glucose (lower values) compared to fluxes using 10% U-^{13}C glucose (upper values). The values in parentheses represent standard deviations. Estimated fluxes represent the molar percentages of an average specific glucose uptake rates of 30.7 ± 2.8 mmol/gCDW/h. Arrowheads indicate the direction of fluxes shown as positive (negative fluxes are in opposite direction to that of arrowheads). See Supplementary Table S1 for details about these pathways and nomenclature for abbreviations.

Figure 6. Formate flux identifiability. (A) Effect of *zwf*, *akgdh* and extent of reversibility of *ppc* on *fhl* flux standard deviations in 1-U-^{13}C experiment. Linearized method was used to estimate standard deviations of *fhl*. (B) Effect of *zwf* flux on CO_2 and formate enrichment, and the SD of *fhl* flux in 1-^{13}C experiment. Linearized method was used to estimate standard deviations of *fhl*.

Tables

Table 1. Metabolic fluxes for the synthesis of fermentation products and biomass formation obtained from experimental data in c-MFA, U-^{13}C-MFA, and 1-U-^{13}C-MFA.

		c-MFA/U-^{13}C-MFA	1-U-^{13}C-MFA
Biomass synthesis			
	G6P	0.46 ± 0.05	0.44 ± 0.04
	F6P	0.16 ± 0.02	0.15 ± 0.02
	R5P	2.02 ± 0.20	1.91 ± 0.19
	E4P	0.81 ± 0.08	0.77 ± 0.08
	T3P	0.29 ± 0.03	0.28 ± 0.03
	3PG	1.59 ± 0.16	1.51 ± 0.15
	PEP	1.17 ± 0.12	1.11 ± 0.11
	PYR	6.37 ± 0.64	6.04 ± 0.60
	AcCoA	8.43 ± 0.84	7.99 ± 0.80
	AKG	2.43 ± 0.24	2.30 ± 0.23
	OAA	4.02 ± 0.40	3.81 ± 0.38
	Serine	0.46 ± 0.05	0.44 ± 0.04
	Glycine	1.31 ± 0.13	1.24 ± 0.12
Fermentation productions			
	Glucose	100.00 ± 0.00	100.00 ± 0.00
	Lactate	11.78 ± 4.84	11.28 ± 4.25
	Succinate	5.49 ± 0.81	5.22 ± 1.08
	Formate	136.13 ± 4.07	136.32 ± 5.01
	Acetate	70.90 ± 10.02	72.37 ± 13.36
	Ethanol	62.86 ± 0.99	60.13 ± 0.18
	Acetate/Ethanol*	133.77 ± 11.01	132.51 ± 13.54

*combined flux of acetate and ethanol production

Table 2. Selected intracellular fluxes calculated using conventional Metabolic Flux Analysis (c-MFA) and ^{13}C-based Metabolic Flux Analysis (^{13}C-MFA). ^{13}C-MFA was conducted using either uniformly (U)-^{13}C-labeled glucose (U-^{13}C-MFA) or a combination of U-^{13}C- and 1-^{13}C-labeled glucose (1-U-^{13}C-MFA).

MFA technique / Pathway/Reaction	c-MFA	^{13}C-MFA				
		U-^{13}C-MFA	1-U-^{13}C-MFA	1-U-^{13}C-MFA*	1-U-^{13}C-MFA**	1-U-^{13}C-MFA***
Glucose transport and phosphorylation						
GLU + PEP → G6P + PYR	100.0±0.0	100.00±0.00	100.00±0.00	100.0±0.0	100.0±0.0	100.0±0.00
Embden-Meyerhof-Parnas						
G6P → F6P	55.1±33.7	75.7±13.5	96.9±1.3	98.4±0.6	97.0±1.7	96.2±1.3
F6P + ATP → G3P + G3P	83.0+11.2	89.8±4.5	97.2±0.4	97.6±0.2	96.2±1.6	96.9±0.5
G3P → PEP + ATP + NADH	175.9±11.2	182.7±4.5	192.4±0.5	192.7±0.5	191.2±1.7	192.0±0.6
PEP → PYR + ATP	62.8±11.2	71.6±6.2	82.8±1.7	78.2±1.8	17.8±21.2	82.1±1.5
Oxidative pentose phosphate						
G6P → R5P + CO_2 + 2NAD(P)H	44.4±33.7	23.9±13.5	2.7±1.3	1.2±0.6	1.3±0.7	3.4±1.3
PEP carboxylation						
PEP + CO_2 → OAA	11.9±0.8	10.0±2.0	6.9±1.5	11.9±1.2	70.6±20.9	7.2±1.6
Pyruvate dissimilation						
PYR → AcCoA + Formate	144.8±10.5	151.2±8.8	167.6±6.3	160.6±6.7	165.±6.8	165.0±4.1
Fermentation						
PYR + NADH → Lactate	11.6±4.6	14.0±5.3	9.4±4.9	11.5±5.2	13.3±5.0	11.3±4.0
AcCoA → Acetate + ATP AcCoA + 2NADH → Ethanol	133.9±10.5	135.1±7.2	146.3±6.9	149.2±6.9	136.5±7.7	144.4±4.4
Formate → CO_2 + H_2	8.91±11.21	16.5±13.1	33.7±12.5	26.5±12.0	28.3±12.4	86.6±44.8
Formate→ $Formate_{ext}$	135.9±3.8	134.7±11.3	133.9±12.4	134.1±12.6	137.2±12.8	78.5±45.6
Carbon dioxide evolution						
CO_2 → CO_{2ext}	45.1±24.9	35.3±13.9	36.4±13.3	19.4±12.4	34.4±13.2	91.1±44.8
Chi square (χ^2)	646	652	729.5	681	544.5	651

Estimated fluxes represent the molar percentages of the average specific glucose uptake rates, which were 29.0 ± 2.4 mmol/gCDW/h for c-MFA and U-^{13}C-MFA, and 30.7 ± 2.8 mmol/gCDW/h for 1-U-^{13}C-MFA. See nomenclature for abbreviations and Figure 1 for details about pathways. Asterisks indicate the following modifications: *) the glyoxylate shunt was excluded and the TCA cycle modified to operate as two independent branches, **) the Entner-Doudoroff (ED) pathway and the malic enzyme (ME) were added to the network; and ***) extracellular measurements were not included in the χ^2 criterion.

Literature Cited

1. Stephanopoulos G, Aristidou AA, Nielsen J. *Metabolic Engineering: Principles and Methodology*. San Diego, CA: Academic Press, CA; 1998.
2. Romeo T, Snoep JL. Glycolysis and flux control. In: Böck A, Curtiss R III, Kaper JB, Karp PD, Neidhardt FC, Nyström T, Slauch JM, Squires CL, Ussery D. EcoSal—*Escherichia coli* and *Salmonella*: cellular and molecular biology. Washington, DC: ASM Press; 2005.
3. Nielsen J. It is all about metabolic fluxes. *J Bacteriol*. 2003;185:7031-7035.
4. Stephanopoulos G. Metabolic fluxes and metabolic engineering. *Metab Eng*. 1999;1:1-11.
5. Klapa MI, Park SM, Sinskey AJ, Stephanopoulos G. Metabolite and isotopomer balancing in the analysis of metabolic cycles: I. Theory. *Biotechnol Bioeng*. 1999;62:375–391.
6. Wiechert W, Möllney M, Petersen S, de Graaf AA. A universal framework for ^{13}C metabolic flux analysis. *Metab Eng*. 2001;3:265–283.
7. Szyperski T. Biosynthetically directed fractional ^{13}C-labeling of proteinogenic amino acids: An efficient analytical tool to investigate intermediary metabolism. *Eur J Biochem*. 1995;232:433-448.
8. Grivet JP. NMR for microbiology: In vivo and in situ applications. *Prog Nucl Mag Res Sp*. 2009;54:1-53.
9. Szypreski T, Bailey JE, Wuthrich K. Detecting and dissecting metabolic fluxes using biosynthetic fractional ^{13}C labeling and two-dimensional NMR spectroscopy. *Trends Biotech*. 1996;14:453-458.

10. Fischer E, Sauer U. Metabolic flux profiling of *Escherichia coli* mutants in central carbon metabolism using GC-MS. *Eur J Biochem*. 2003a;270:880-891.

11. Fischer E, Sauer U. Large-scale in vivo flux analysis shows rigidity and suboptimal performance of *Bacillus subtilis* metabolism. *Nat Genet*. 2005;37:636–640.

12. Tang YJ, Martin HG, Myers S, Rodriguez S, Baidoo EK, Keasling JD. Advances in analysis of microbial metabolic fluxes via ^{13}C isotopic labeling. *Mass Spectrom Rev*. 2009;28:362– 375.

13. Emmerling, M., Dauner M, Ponti A, Fiaux J, Hochuli M, Szyperski T, Wuthrich K, Bailey JE, Sauer U. Metabolic flux responses to pyruvate kinase knockout in *Escherichia coli*. *J Bacteriol*. 2002;184:152-164.

14. Al Zaid Siddiquee K, Arauzo-Bravo MJ, Shimizu K. Metabolic flux analysis of *pykF* gene knockout *Escherichia coli* based on ^{13}C-labeling experiments together with measurements of enzyme activities and intracellular metabolite concentrations. *Appl Microbiol Biotechnol*. 2004;63:407–417.

15. Zhao J, Shimizu K. Metabolic flux analysis of *Escherichia coli* K12 grown on ^{13}C-labeled acetate and glucose using GC-MS and powerful flux calculation method. *J Biotechnol*. 2003;101:101–117.

16. Schuetz R, Kuepfer L, Sauer U. Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol Syst Biol*. 2007;3:119.

17. Sauer U, Lasko DR, Fiaux J, Hochuli M, Glaser R, Szyperski T, Wuthrich K, Bailey JE. Metabolic flux ratio analysis of genetic and environmental modulations of *Escherichia coli* central carbon metabolism. *J Bacteriol*. 1999;181:6679–6688.

18. Sauer U, Canonaco F, Heri S, Perrenoud A, Fischer E. The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of *Escherichia coli*. *J Biol Chem*. 2004;279:6613–6619.

19. Nanchen A, Schicker A, Sauer U. Nonlinear dependency of intracellular fluxes on growth rate in miniaturized continuous cultures of *Escherichia coli*. *Appl Environ Microbiol*. 2006; 72:1164–1172.

20. Fong SS, Nanchen A, Palsson BO, Sauer U. Latent pathway activation and increased pathway capacity enable *Escherichia coli* adaptation to loss of key metabolic enzymes. *J Biol Chem*. 2006;281:8024–8033.

21. Suthers PF, Burgard AP, Dasika MS, Nowroozi F, Van Dien S, Keasling JD, Maranas CD. Metabolic flux elucidation for large-scale models using ^{13}C labeled isotopes. *Metab Eng*. 2007;9:387–405.

22. Fischer E, Sauer U. A novel metabolic cycle catalyzes glucose oxidation and anaplerosis in hungry *Escherichia coli*. *J Biol Chem*. 2003b;278:46446-46451.

23. Schmidt K, Nielsen J, Villadsen J. Quantitative analysis of metabolite fluxes in *Escherichia coli*, using two-dimensional NMR spectroscopy and complete isotopomer models. *J Biotechnol*. 1999a;71:175-190.

24. Ingraham JL, Maaloe O. and Neidhardt FC. *Growth of the Bacterial Cell*. Sunderland, MA: Sinauer Associates; 1983.

25. Neidhardt FC, Bloch PL, Smith DF. Culture Medium for Enterobacteria. *J Bacteriol*. 1974;119:736-747.

26. Szypreski T. ^{13}C-NMR, MS and metabolic flux balancing in biotechnology research. *Q Rev Biophys*. 1998;31:41-106.

27. Johnson BA, Blevins RA. NMRView: a computer program for the visualization and analysis of NMR data. *J Biomol NMR*. 1994;4:603-614.

28. Sriram G, Fulton DB, Shanks JV. Flux quantification in central carbon metabolism of *Catharanthus roseus* hairy roots by C-13 labeling and comprehensive bondomer balancing. *Phytochemistry*. 2007;68:2243-2257.

29. Van Winden W, Schipper D, Verheijen P, Heijnen J. Innovations in generation and analysis of 2D [^{13}C,^{1}H] COSY NMR spectra for metabolic flux analysis purposes. *Metab Eng*. 2001;3:322-343.

30. Sriram G, Fulton DB, Iyer VV, Peterson JM, Zhou R, Westgate ME, Spalding MH, Shanks JV. Quantification of compartmented metabolic fluxes in developing soybean embryos by employing biosynthetically directed fractional ^{13}C labeling, two-dimensional [^{13}C, ^{1}H] nuclear magnetic resonance, and comprehensive isotopomer balancing. *Plant Physiol*. 2004;136:3043-3057.

31. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical Recipes in C: The Art of Scientific Computing* (2nd edition). Cambridge, UK: Cambridge University Press; 1992.

32. Arauzo-Bravo MJ, Shimizu K. An improved method for statistical analysis of metabolic flux analysis using isotopomer mapping matrices with analytical expressions. *J Biotechnol*. 2003;105:117–133.

33. Mollney M, Wiechert W, Kownatzki D, de Graaf AA. Bidirectional reaction steps in metabolic networks: IV. Optimal design of isotopomer labeling experiments. *Biotechnol Bioeng*. 2006;66:86-103

34. Sawers RG, Clark DP. Fermentative pyruvate and acetyl-coenzyme A metabolism. In: Böck A, Curtiss R III, Kaper JB, Karp PD, Neidhardt FC, Nyström T, Slauch JM, Squires CL, Ussery D. EcoSal—*Escherichia coli* and *Salmonella*: cellular and molecular biology. Washington, D.C: ASM press; 2004.

35. Sawers RG, Blokesch M, Bock A. Anaerobic formate and hydrogen metabolism. In: Böck A, Curtiss R III, Kaper JB, Karp PD, Neidhardt FC, Nyström T, Slauch JM, Squires CL, Ussery D. EcoSal—*Escherichia coli* and *Salmonella*: cellular and molecular biology. Washington, D.C: ASM press; 2004.

36. Cozzone AJ. Regulation of acetate metabolism by protein phosphorylation in enteric bacteria. *Annu Rev Microbiol*. 1998;52:127–164.

37. Harris RK. *Nuclear Magnetic Resonance Spectroscopy: A Physicochemical View*. London: Pitman Books; 1983.

38. Phue JN, Shiloach J. Transcription levels of key metabolic genes are the cause for different glucose utilization pathways in *E. coli* B (BL21) and *E. coli* K (JM109). *J Biotechnol*. 2004;109:21–30.

39. Van de Walle M, Shiloach J. Proposed mechanism of acetate accumulation in two recombinant *Escherichia coli* strains during high density fermentation. *Biotechnol Bioeng*. 1998;57:71-78.

40. Dauner M, Bailey JE, Sauer U. Metabolic flux analysis with a comprehensice isotopomer model in *Bacillus subtilis*. *Biotechnol Bioeng*. 2000;76:144-156.

41. Schmidt K, Norregaard LC, Pedersen B, Meissner A, Duus J, Nielsen J, Villadsen J. Quantification of intracellular metabolic fluxes from fractional enrichment ^{13}C-^{13}C

coupling constraints on the isotopomer distribution in labeled biomass components. *Metab Eng*. 1999b;1:166-179.

42. Fischer E, Zamboni N, Sauer U. High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived ^{13}C constraints. *Anal Biochem*. 2004;325:308-316.

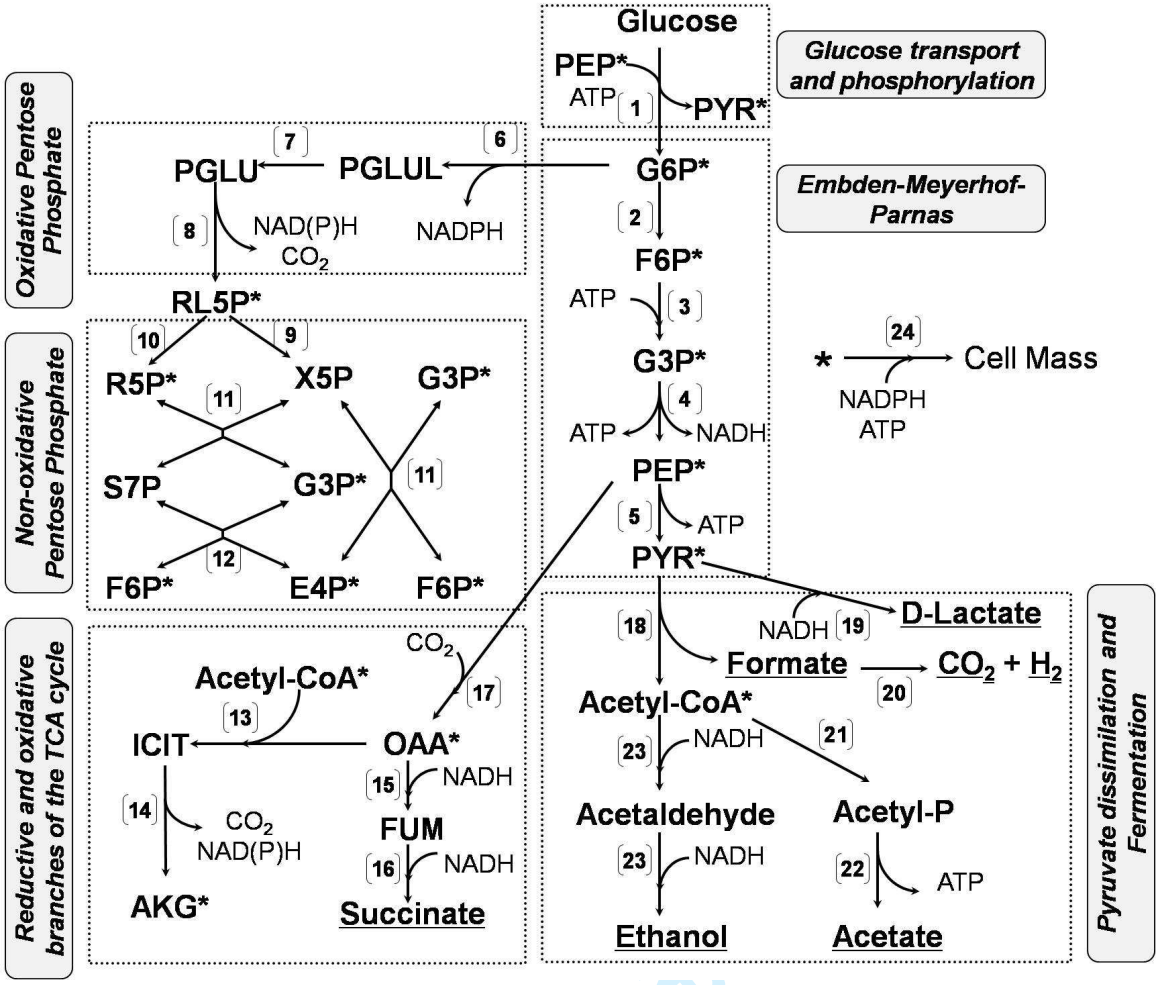
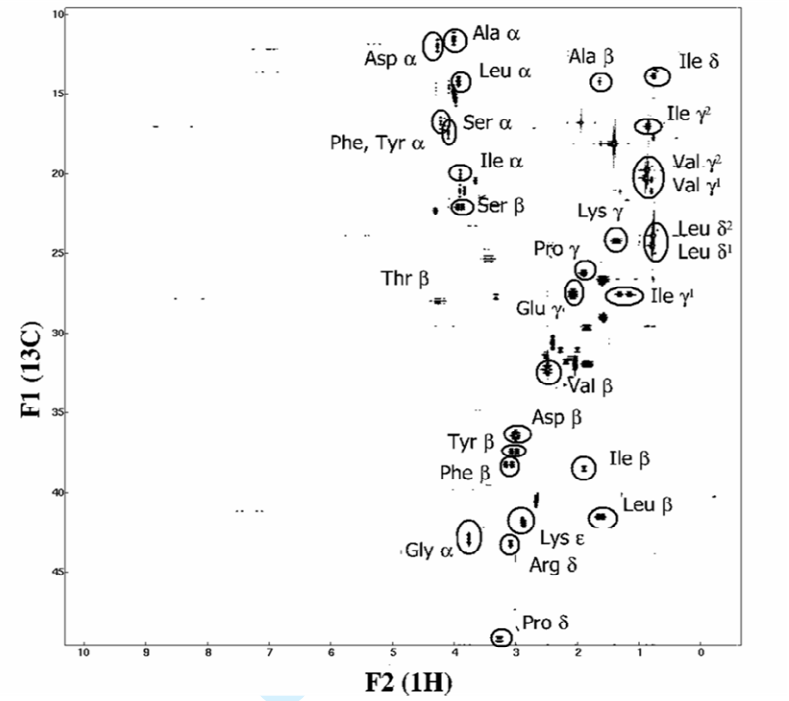
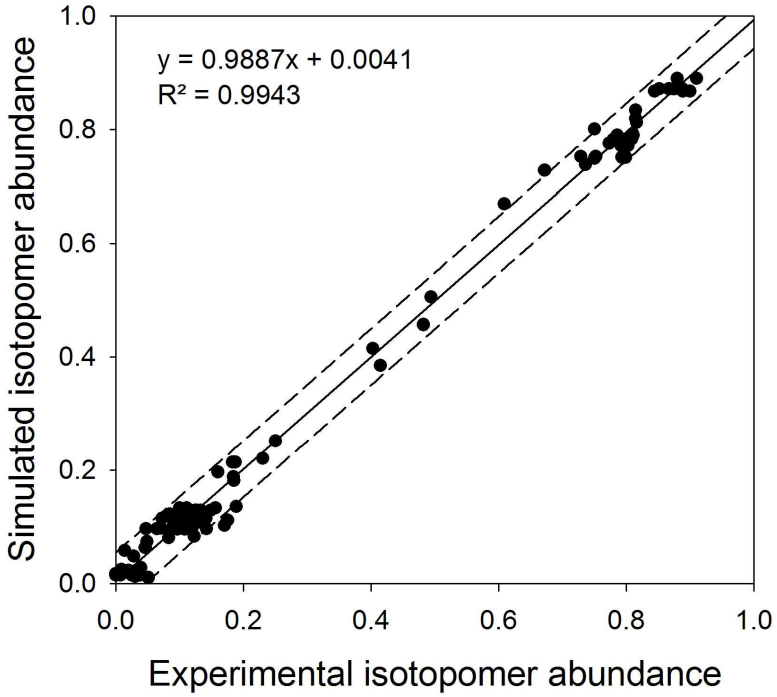


Figure 1.


Ala α
Asp α
Leu α
Ala β
Ile δ
Ser α
Ile γ²
Phe, Tyr α
Ile α
Val γ²
Val γ¹
Ser β
Lys γ
Leu δ²
Leu δ¹
Pro γ
Thr β
Ile γ¹
Glu γ
Val β
Asp β
Tyr β
Phe β
Ile β
Leu β
Lys ε
Gly α
Arg δ
Pro δ
F1 (13C)
F2 (1H)

(A)


y = 0.9887x + 0.0041
R² = 0.9943
Simulated isotopomer abundance
Experimental isotopomer abundance
0.0
0.2
0.4
0.6
0.8
1.0

(B)

Figure 2.

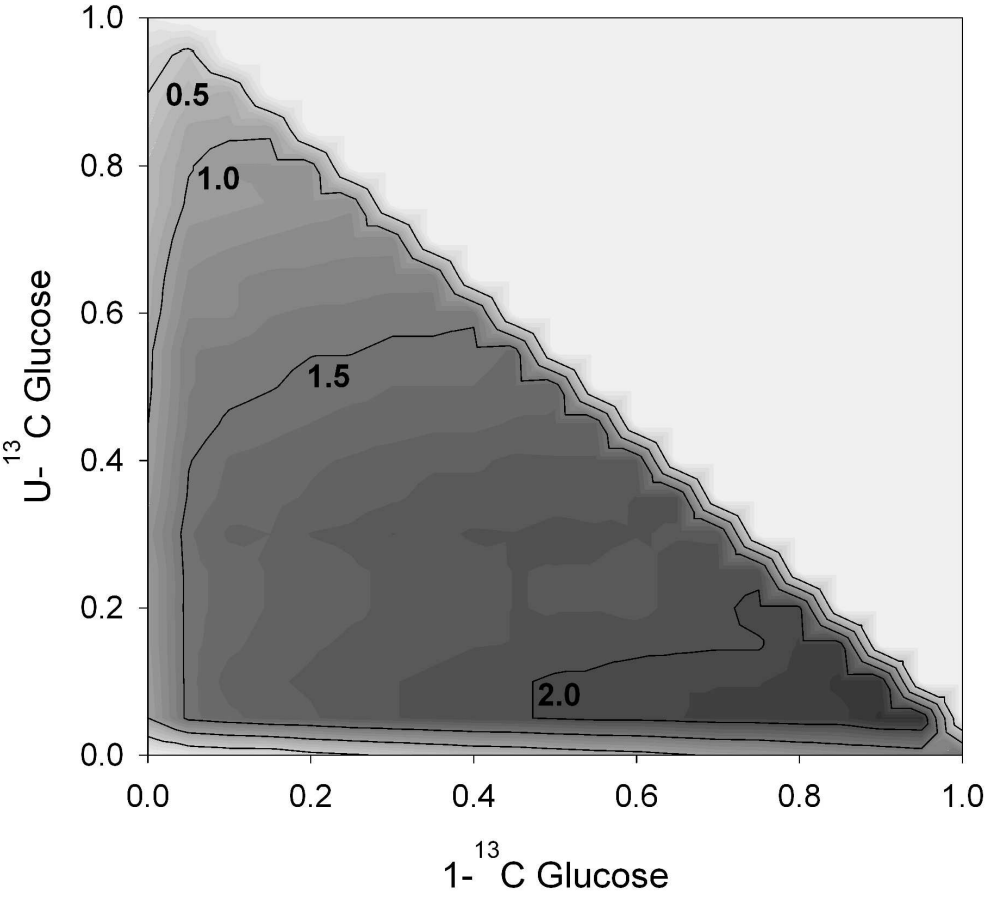


Figure 3.

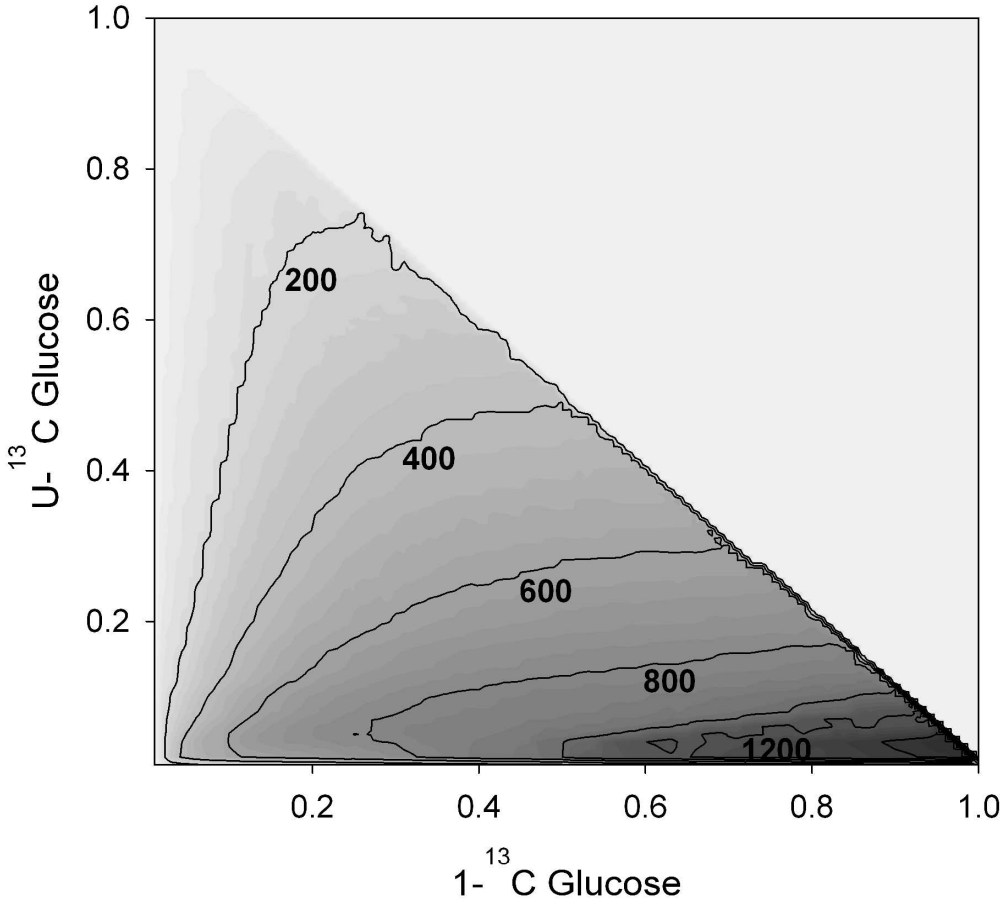


Figure 4.

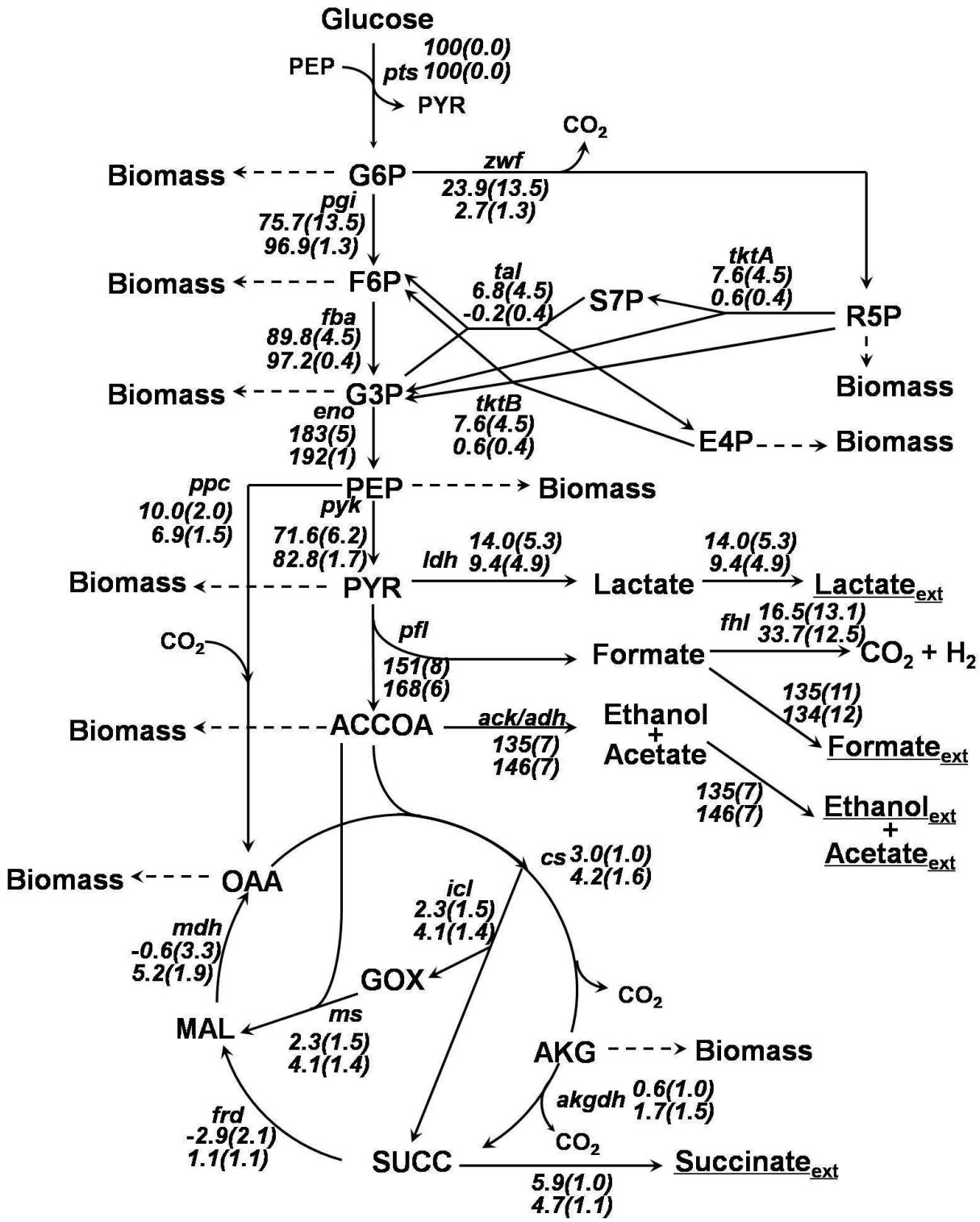
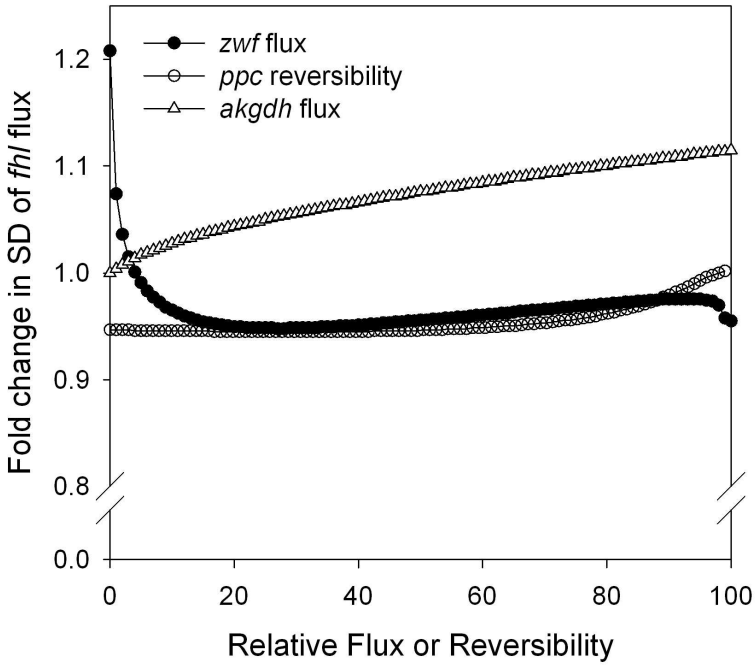
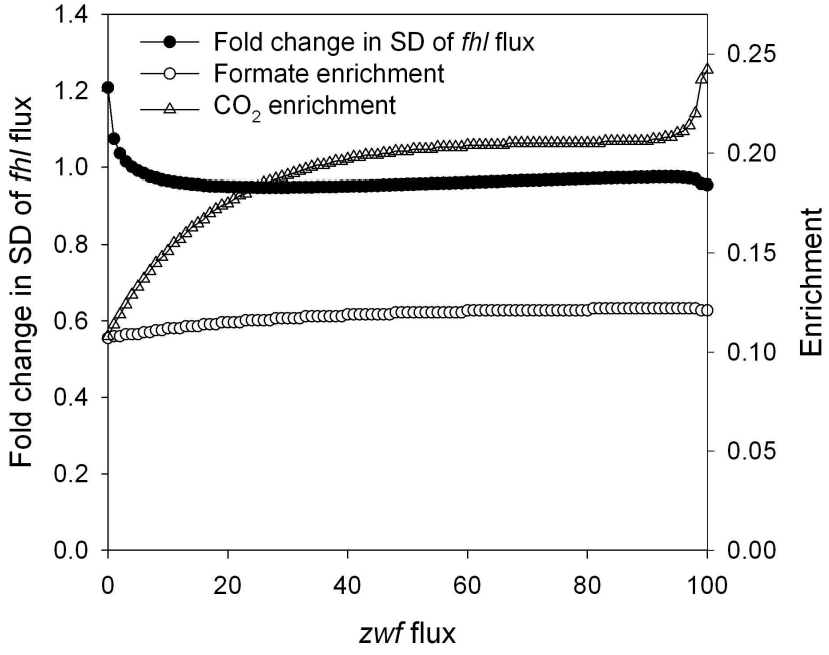

Glucose
PEP
100(0.0)
pts 100(0.0)
PYR
CO2
Biomass
G6P
zwf
23.9(13.5)
2.7(1.3)
pgi
75.7(13.5)
96.9(1.3)
F6P
tal
6.8(4.5)
-0.2(0.4)
S7P
tktA
7.6(4.5)
0.6(0.4)
R5P
fba
89.8(4.5)
97.2(0.4)
G3P
tktB
7.6(4.5)
0.6(0.4)
E4P
eno
183(5)
192(1)
PEP
ppc
10.0(2.0)
6.9(1.5)
pyk
71.6(6.2)
82.8(1.7)
PYR
ldh 14.0(5.3)
9.4(4.9)
Lactate
14.0(5.3)
9.4(4.9)
Lactate_ext
pfl
151(8)
168(6)
Formate
fhl 16.5(13.1)
33.7(12.5)
CO2 + H2
135(11)
134(12)
Formate_ext
ACCOA
ack/adh
135(7)
146(7)
Ethanol
+
Acetate
135(7)
146(7)
Ethanol_ext
+
Acetate_ext
OAA
cs 3.0(1.0)
4.2(1.6)
icl
2.3(1.5)
4.1(1.4)
mdh
-0.6(3.3)
5.2(1.9)
GOX
ms
2.3(1.5)
4.1(1.4)
MAL
AKG
akgdh 0.6(1.0)
1.7(1.5)
frd
-2.9(2.1)
1.1(1.1)
SUCC
5.9(1.0)
4.7(1.1)
Succinate_ext

Figure 5.


zwf flux
ppc reversibility
akgdh flux
Fold change in SD of fhl flux
Relative Flux or Reversibility
1.2
1.1
1.0
0.9
0.8
0.0
0
20
40
60
80
100

(A)


Fold change in SD of fhl flux
Formate enrichment
CO2 enrichment
Fold change in SD of fhl flux
Enrichment
zwf flux
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.25
0.20
0.15
0.10
0.05
0.00
0
20
40
60
80
100

(B)

Figure 6.