# Final Technical Report
# Quantitative analysis of metabolic regulation by integration of metabolomics, proteomics, and fluxomics

**The present project involved two major areas of research achievement: quantitative analysis of yeast metabolic regulation and elucidation of the reversibility of glycolysis in cellulolytic bacteria. In addition, it involved the development of a diversity of new tracer methodologies and advancements in metabolomics.**

## Part I: Systems-level analysis of yeast metabolic regulation

Despite extensive knowledge of metabolic pathways, the means by which metabolic reaction rates (fluxes) are controlled remains incompletely understood. As one of the best-studied and most industrially and biofuel important organisms, Baker's yeast (*S. cerevisiae*) provides an ideal testbed for examining metabolic flux control.

Most metabolic regulatory mechanisms were discovered by isolating enzymes and studying their kinetics *in vitro*. While powerful for proving specific regulatory interactions, this reductionist approach has been less effective at revealing regulation in the intact cell. Metabolic regulation *in vivo* depends not only on enzyme kinetics, but also on how flux control is distributed across enzymes in a pathway and how much different metabolite and enzyme levels change across physiological states.

Recent advances in metabolomics open new avenues to understanding metabolic regulation. For example, the dynamic metabolome response to a biological perturbation can be measured and the resulting data fit with differential equations. Such fitting can in principle identify both kinetic parameters and regulatory interactions, but requires global non-linear search, which scales poorly with network size. Therefore, there is a need for new ways of deducing physiological metabolic regulation from systems-level measurements. To this end, we developed an approach which uses measurements both of yeast metabolic flux and of enzyme and metabolite concentrations, across multiple different nutrient conditions, to identify enzyme kinetic parameters and regulators on a reaction-by-reaction basis.

We analyzed the metabolism of 25 steady-state yeast cultures spanning diverse growth rates and nutrient limitations—carbon (glucose), nitrogen (ammonia), phosphorous (phosphate), leucine (in a

leucine auxotroph), and uracil (in a pyrimidine auxotroph). Fluxes were determined by flux balance analysis constrained by experimental measurements of nutrient uptake, biomass production, and waste excretion. Metabolite and enzyme levels were measured by liquid chromatography-mass spectrometry. For each of 56 individual reactions with adequate experimental data, we evaluated whether the observed fluxes across conditions could be explained by Michaelis-Menten kinetics and whether inclusion of measured metabolites as regulators could better explain the fluxes. The metabolites that we "computationally tested" as potential regulators were chosen based on their regulating the reaction of interest in at least one other organism, and were prioritized based on Bayesian analysis including both our results and prior literature data.

Through this strategy, which we term Systematic Identification of Meaningful Metabolic Enzyme Regulation (SIMMER), we identified 35 best-supported candidate regulators. Biochemical testing of 10 novel predictions validated three instances of yeast metabolic regulation: inhibition of ornithine transcarbamylase (Arg3) by alanine, pyruvate decarboxylase (Pdc) by phenylpyruvate, and pyruvate kinase (Cdc19) by citrate. The latter two interactions serve to coordinate yeast carbon and nitrogen metabolism by slowing fermentation when nitrogen is scarce. Each of these three new physiological regulatory interactions spans pathways. This contrasts with the nine instances of known yeast regulation that were reconfirmed by SIMMER, each of which was within a pathway. Thus, cross-pathway regulation may be more common and important than currently appreciated.

We also assessed the extent to which changes in flux across the studied environmental conditions were determined by concentration changes in individual reaction species (enzyme, substrate, product, and regulator). For reversible reactions, substrate concentrations largely explain flux, whereas for strongly forward driven reactions, enzymes, substrates, and allosteric regulators each play a major role. Overall, the combined impact of metabolites was more than double that of enzymes.

In summary, systems-level measurements of metabolic fluxes, metabolite concentrations, and enzyme concentrations can be dissected at the individual reaction level to reveal metabolic regulation. Such analysis in yeast showed that changes in flux across nutrient conditions are predominantly due to metabolite, not enzyme, levels. Thus, metabolism is substantially self-regulating. The SIMMER methodology should be broadly useful for revealing metabolic regulation also in less studied biofuel-relevant microbes.

Please see: Hackett, S. R., Zanotelli, V. R., Xu, W., Goya, J., Park, J. O., Perlman, D.H., Gibney, P. A., Botstein, D., Storey, J.D., Rabinowitz, J. D. (2016) Systems-level analysis of mechanisms regulating yeast metabolic flux. *Science*, 354: aaf2786.

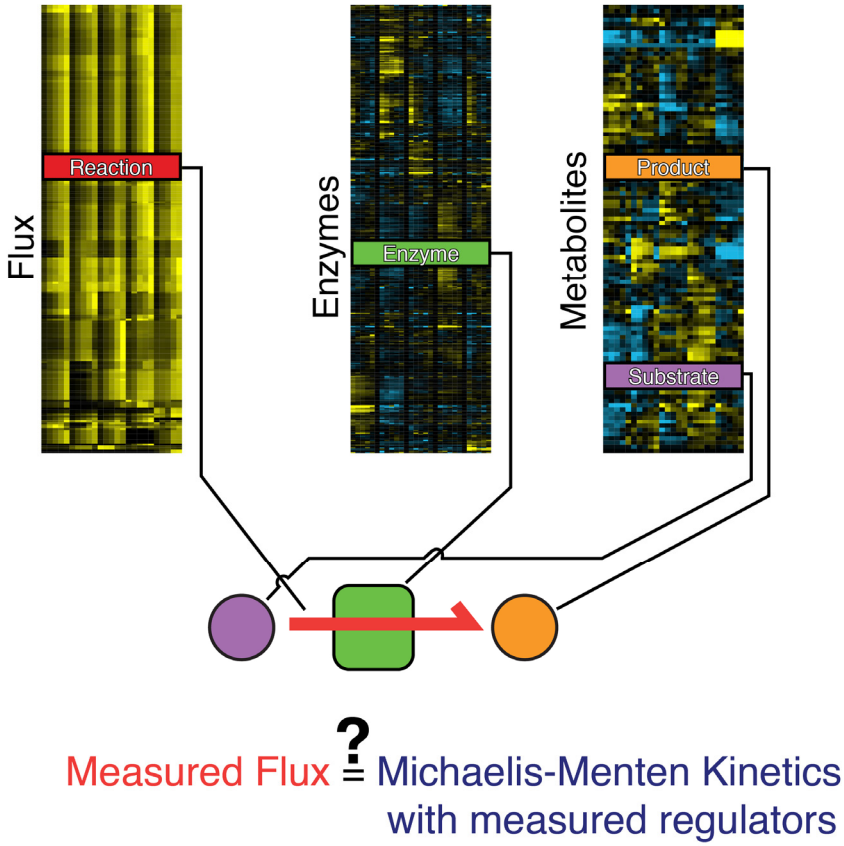


**Figure 1. Systematic Identification of Meaningful Metabolic Enzyme Regulation (SIMMER).** Systems-level data – including metabolic fluxes, metabolite concentrations, and enzyme concentrations – is analyzed at the level of individual reactions to identify metabolic regulation. Specifically, SIMMER evaluates whether observed fluxes across biological conditions are explained by a Michaelis-Menten relationship between substrate, product, and enzyme concentrations and whether inclusion of additional metabolites as regulators better explains the fluxes.

## Part II: Glycolytic reversibility in the cellulolytic bacterium *C. cellulolyticum*

The steps of glycolysis have long been known but the pathway's design principles and regulatory mechanisms remain incompletely understood. A better understanding would benefit metabolic engineering.

A key aspect of any metabolic pathway is its energetics. According to the second law of thermodynamics, every pathway step must have a negative Gibbs free energy of reaction (ΔG). Steps with large absolute values of ΔG are effectively irreversible and preferred sites for regulation of flux based on enzyme activity. Steps with small ΔG are near equilibrium, energy efficient, and enable large changes in net flux with small changes in substrate and product levels. Steps with large absolute

values of ΔG consume more than their share of the finite available pathway driving force, while steps with small ΔG require "extra" enzyme to maintain the reaction near to equilibrium.

While advances in analytical capabilities have rendered the measurement of metabolite concentrations and fluxes increasingly straightforward, methods for measuring metabolic energetics remain underdeveloped. Here we focused on using reaction reversibility to probe thermodynamics:

$$\Delta G = \Delta G'^{\circ} + RT \ln Q = RT \ln \frac{Q}{K_{eq}} = RT \ln \frac{J^{-}}{J^{+}} \qquad (2)$$

where $J^{-}/J+$ is the backward-to-forward flux ratio, as determined from isotope tracer studies. We selected [5-$^{2}$H]- and [1,2-$^{13}$C]-glucose as tracers, as they illuminate reverse flux in glycolysis by either loss of $^{2}$H or scrambling of $^{13}$C.

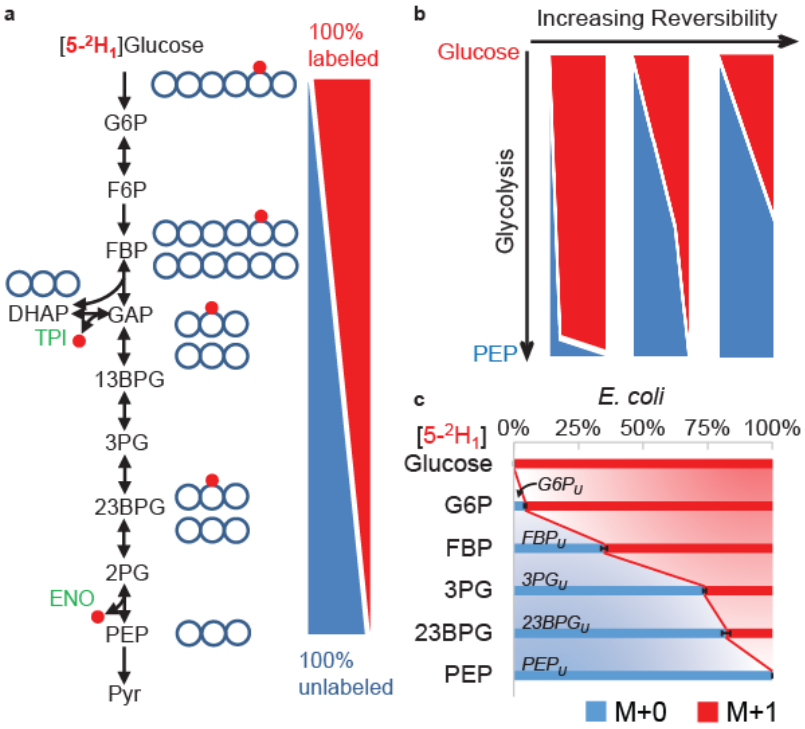


**Figure 2. Visualizing the reversibility of glycolysis using [5-$^{2}$H$_1$]glucose.** (**a**) The deuterium of [5-$^{2}$H$_1$]glucose is removed from glycolytic intermediates in the reverse triose phosphate isomerase (TPI: DHAP ⇌ GAP) and forward enolase (ENO: 2PG ⇌ PEP) reactions, leading to a descending gradient of $^{2}$H enrichment along glycolysis. (**b**) Increasing reversibility of glycolysis results in $^{2}$H depleting earlier in the pathway. (**c**) Labeling data from [5-$^{2}$H$_1$]glucose in *E. coli*: metabolite names with subscript U indicate unlabeled fractions. Error bars represent s.e.m. (n=3).

Specifically, in glycolysis, the deuterium of [5-$^{2}$H]glucose can be lost at the following two steps: reverse triose phosphate isomerase (TPI) and forward enolase (ENO). If ENO were strongly forward-driven ($\Delta G \ll 0$), the labeling of 1,3-bisphosphoglycerate (13BPG), 3-phosphoglycerate (3PG), 2,3-bisphosphoglycerate (23BPG), and 2-phosphoglycerate (2PG) would be identical to that of their precursor GAP. With increasing reversibility of glycolytic reactions including ENO, unlabeled intermediates travel upstream. Therefore, [5-$^{2}$H]glucose results in a monotonic decrease of $^{2}$H labeling down glycolysis that reflects the reversibility of the constituent reactions (**Figure 2**).

To validate the $^{2}$H gradient, we cultured *E. coli* as an example microbe. LC-MS data showed the predicted labeling gradient, but a flux model revealed that various combinations of TPI and lower glycolysis reversibility may result in indistinguishable $^{2}$H labeling patterns. To address this, we added [1,2-$^{13}C_2$]glucose tracing, which more directly probes TPI reversibility, and thereby allowed us to algebraically express the reverse flux (and thus $\Delta G$) of each glycolytic reaction step in terms of measured unlabeled (non-deuterated) metabolite fractions from [5-$^{2}H_1$]glucose and reverse TPI flux.

We then applied this technology to *C. cellulolyticum* (an obligate anaerobe capable of hydrolyzing cellulosic biomass into simple sugars and fermenting them into ethanol and other products and, for comparison, a non-cellulolyticum Clostridium species of historical importance as a solvent producer, *C. acetobutylicum.* Interestingly, the $^{2}$H labeling of all glycolytic intermediates in *C. cellulolyticum* was much less than in *E. coli* or its faster fermenting relative *C. acetobutylicum* (**Figure 3**)**.** We found that the total $\Delta G$ drop of *C. cellulolyticum* glycolysis from G6P to PEP is only –3 kJ/mol, roughly 10-fold less than in *E. coli*, or *C. acetobutylicum*.

A major cause for the small free energy drop in *C. cellulolyticum* glycolysis was the small drop in the kinase step that converts F6P into FBP. In *C. thermocellum*, this reaction step is catalyzed by the PFK enzyme that uses pyrophosphate (PPi) instead of ATP. We found that the same occurs in *C. cellulolyticum*. The use of PPi instead of ATP at this step increases the ATP yield of glycolysis from 2 to 3 ATP-equivalents per glucose. This extra ATP generation would appear to be an important benefit for this anaerobe, in which oxidative phosphorylation is unavailable for ATP synthesis and glucose is scarce, coming mainly from cellulose degradation. The downside is a weaker driving force for glucose catabolism and thus slower net glycolysis.

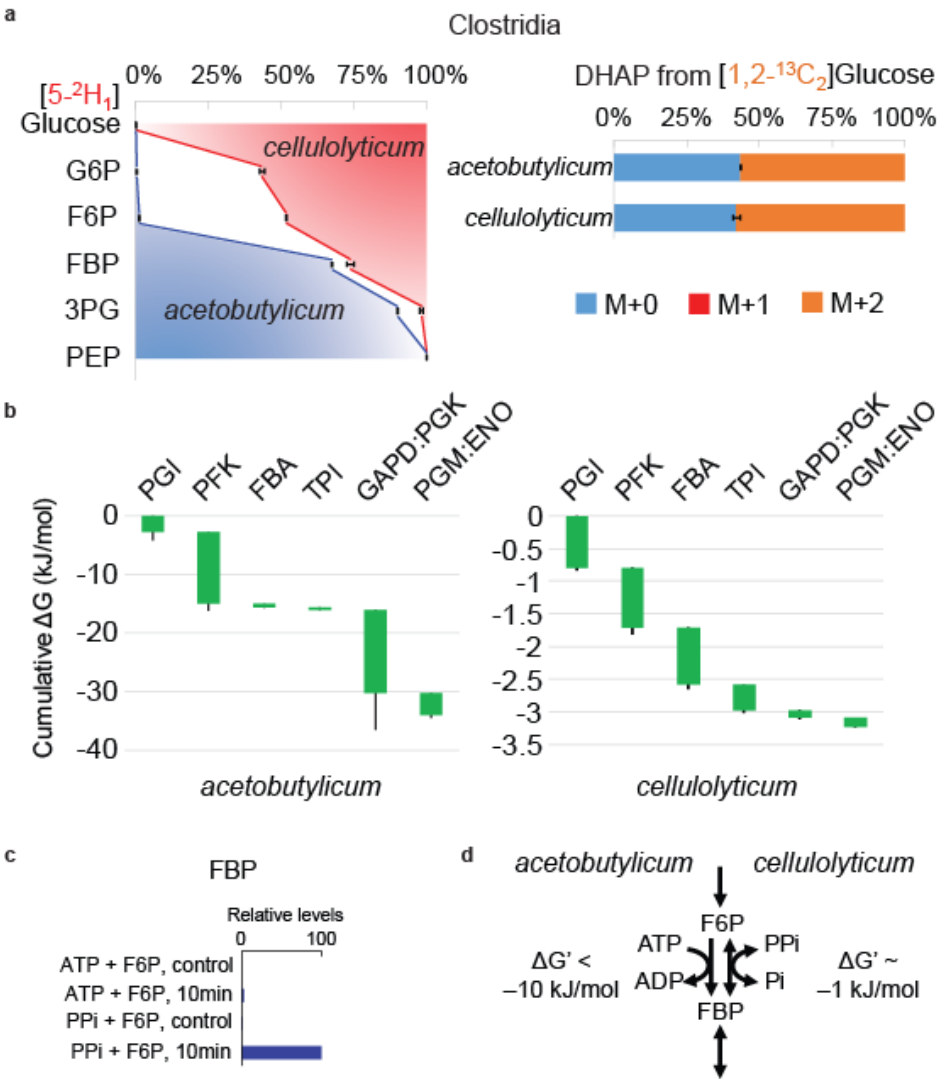


**Figure 3. Slow glycolysis of *C. cellulolyticum* operates near equilibrium using PPi-dependent PFK.** (**a**) Isotope tracing of glycolytic reversibility in the obligate anaerobes *C. cellulolyticum* and *C. acetobutylicum* cultured on [5-$^{2}H_{1}$]- or [1,2-$^{13}C_{2}$]-glucose. (**b**) All glycolytic reactions of *C. cellulolyticum* were close to equilibrium with ΔG > –1 kJ/mol. The resulting cumulative ΔG from G6P to PEP was approximately –3 kJ/mol, a tenth of that of *C. acetobutylicum*. (**c**) The greatest difference between canonical (e.g., *E. coli*, *C. acetobutylicum*, and mammlian) and *C. cellulolyticum* glycolysis was in the ΔG of phosphofructokinase (PFK). In *C. cellulolyticum* cell lysate, fructose-1,6-bisphosphate (FBP) was produced in the presence of pyrophosphate (PPi) but not ATP. Control experiments were carried out by excluding cell lysate from assay mix. (**d**) The weakly forward-driven PFK in *C. cellulolyticum* was due to the use of PPi-Pi pair instead of ATP-ADP pair.

## Part III: Other major technological and scientific advances

1. Determined that many metabolites are frequently mis-measured by LC-MS due to confusion between metabolites and in-source fragmentation products. Identified key instances of this and how to avoid them.
   Please see: Xu, Y. F., Lu, W., Rabinowitz, J. D. (2015) Avoiding misannotation of in-source fragmentation products as cellular metabolites in liquid chromatography-mass spectrometry-based metabolomics. *Anal. Chem.*, 87: 2273-81.
2. Developed 13C and 2H tracers for folate metabolism, which holds the potential for a novel means of $CO_2$ incorporation into biomass and shuttling of electrons into NADPH from abiotic sources via formate.
   Please see: Ducker, G. S., Chen, L., Morscher, R. J., Ghergurovich, J. M., Esposito, M., Teng, X., Kang, Y., Rabinowitz, J. D. (2016) Reversal of cytosolic one-carbon flux compensates for loss of the mitochondrial folate pathway. *Cell Metab.*, 23: 1140-53.
3. Identified an important deficiency of orbitrap mass spectrometry in terms of inaccurate measurement of isotope ratios (i.e. ratio of M, M+1, M+2 peaks, etc.). Show that this deficiency occurs selectively at higher mass resolution due to longer scan times, and can be mitigated by either decreasing mass resolution or selecting narrow mass windows for scanning.
   Please see: Su, X., Lu, W., Rabinowitz, J.D. (2017) Metabolite Spectral Accuracy on Orbitraps. (2017) *Anal. Chem.*, 89: 5940-5948.
4. Integrated metabolomics with phospho-histidine analysis to understand the interplay between glycolysis and serine biosynthetic pathway regulation.
   Please see: Oslund, R. C., Su, X., Haugbro, M., Kee, J. M., Esposito, M., David, Y., Wang, B., Ge, E., Perlman, D. H., Kang, Y., Muir, T. W., Rabinowitz, J. D. (2017) Bisphosphoglycerate mutase controls serine pathway flux via 3-phosphoglycerate. *Nat. Chem. Biol.*, 13: 1081-1087.
5. Solved a long-standing mystery in isotope tracing: How does $D_2O$ transmit its deuterium into fat? We showed that this occurs via Flavin-enzyme mediated H-D exchange between water and NADPH. The redox active hydride of NADPH is then incorporated into fat. This fundamental discovery enables for the first time quantitatively accurate tracing of NADH and NADPH hydride sources, and should be broadly useful in tracing efforts to bioengineer fatty acid production, as well as that of related molecules like fatty alcohols.
   Please see: Zhang, Z., Chen, L., Liu, L., Su, X., Rabinowitz, J. D. (2017) Chemical basis for deuterium labeling of fat and NADPH. *J. Am. Chem. Soc.*, 139: 14368-14371.