Title

Engineering formate assimilation in E.coli using stable isotope tracers and thermodynamic analysis

Author

Xu, Jimmy

Publication Date

2020

UNIVERSITY OF CALIFORNIA

Los Angeles

Engineering formate assimilation in *E.coli* using stable isotope tracers and thermodynamic analysis

A thesis submitted in partial satisfaction of the requirements for the Masters of Science degree in Chemical Engineering

by

Jimmy Xu

2020

ABSTRACT OF THESIS

Engineering formate assimilation in *E.coli* using stable isotope tracers and thermodynamic analysis

by

Jimmy Xu

Masters of Science in Chemical Engineering

University of California, Los Angeles, 2020

Professor Junyoung O. Park, Chair

Metabolite concentrations, fluxes, and free energies constitute the basis for understanding and controlling metabolism. Recently, using high-resolution mass spectrometry and multi-isotope tracing, improved flux quantitation led to determination of metabolic fates, pathway and nutrient contributions, and Gibbs free energy of reaction (ΔG) in central carbon metabolism using a relationship between reaction reversibility and thermodynamic driving force. We applied this quantitative analysis scheme towards the rational design of a pathway to assimilate formate in *E. coli*. Formate, which can be readily produced from CO_2 by electrocatalysis, is an attractive one-carbon feed substrate for the bioproduction of chemicals. Several synthetic pathways utilizing formate have been constructed in *E. coli*; however, these pathways either use formate as an electron donor rather than a carbon source or lose carbons as carbon dioxide. We have designed a pathway, the Gluconeogenesis-Non-Oxidative Glycolysis (GNG-NOG) cycle, capable of incorporating formate into central carbon metabolism with minimal carbon loss. Formate fed to the cell enters gluconeogenesis as pyruvate via pyruvate formate lyase (*pfl*), producing a hexose phosphate that then enters non-oxidative glycolysis (NOG). NOG produces acetate without CO_2 production while generating energy in the form of ATP. Of the six engineered strains theoretically capable of turning the GNG-NOG cycle, we determined, through media sampling and tracing of the metabolic fate of ^{13}C-formate, that the JXG2 strain incorporated formate into gluconeogenesis and NOG. This strain, which has gluconeogenesis reconstructed to include the gluconeogenic *gapB* gene from *B. subtilis* and *pfl* overexpressed, was shown to consistently consume formate in glucose-containing media. Formate was driven up gluconeogenesis, responsible for 0.4% of carbons in hexose phosphate. Then, the labeled hexose phosphates were catabolized down NOG and contributed 0.5% of carbons in acetyl phosphate, which has the potential to be converted to any biomolecule. Thus, we have engineered an *E.* coli strain that can serve as a platform for one-carbon utilization and sustainable biotechnology. At this point, further pathway optimization is needed, as flux through GNG-NOG cycle is low; this may involve conducting targeted strain evolution on formate or knockout of competing pathways, as we also identified extensive labeling in purine and CoA synthesis. The GNG-NOG cycle, when optimized, will be able to convert one-carbon compounds to acetate. In the future, a GNG-NOG strain can be coupled in tandem with an acetate-consuming microorganism to produce a wide variety of value-added chemicals.

The thesis of Jimmy Xu is approved.

Yi Tang

Harold Monbouquette

Junyoung O. Park, Committee Chair

University of California, Los Angeles

2020

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my adviser Dr. Junyoung O. Park for his continuous support throughout my degree and thesis. I would also like to thank Dr. Yi Tang and Dr. Harold Monbouquette for agreeing to be part of my thesis committee. I want to extend a special thanks to Alec Jaeger and Derrick Chuang for providing guidance and materials throughout my thesis. I want to thank Dr. Daniel Amador-Noguez, Julia Martien, Cole Gilbertson, and Junyu Ma for their contribution to Chapter 1, as well as Dr. Yinjie Tang and Dr. Ludmilla Aristilde for editing the review for publication in Current Opinion in Biotechnology. I want to thank Dr. Yu Chen for his assistance with running samples at the Molecular Instrumentation Center. I want to thank Dr. James Liao for allowing me to use facilities and materials. Finally, I would like to thank Hans Wahono, Jacob Prohroff, Junyu Ma, and Justine Reblando for their assistance in various experiments. As the first author of the article "Metabolic flux analysis and fluxomics-driven determination of reaction free energy using multiple isotopes" in the Current Opinion in Biotechnology journal, I am permitted by Elsevier to use the article in a thesis, as long as the thesis is not published commercially.

Chapter 1: Thesis Outline

Stable isotope tracers and liquid chromatography-mass spectrometry (LC-MS) have been used to study metabolism for nearly a century. Advances in these analytical capabilities have enabled comprehensive characterization and quantitation of metabolites, including determining fates of key metabolites, relative contributions of different nutrients and pathways, as well as the favorability of key reactions. This bevy of information can be used towards the rational design of novel pathways, such as the utilization of one-carbon compounds by non-methylotrophs.

In order to better understand metabolic trends, some recent studies [1-4] have employed stable isotope tracing and a relationship between reaction fluxes and Gibbs free energy of reaction to determine the thermodynamics of various metabolic reactions and pathways. This method does not require knowing standard Gibbs free energies or intracellular concentrations, and as such, is much easier to utilize. In Chapter 2, we delve into some recent studies that utilize stable isotope tracing to examine metabolism, including the determination of thermodynamic driving forces.

In Chapter 3, we sought to apply stable isotopes and thermodynamic analysis towards the design of a new metabolic pathway in *E. coli*. We wished to utilize formate as a feed substrate, as it is a one-carbon compound that can be produced from carbon dioxide. The designed pathway, named the Gluconeogenesis-Non-Oxidative Glycolysis cycle (GNG-NOG) cycle, would drive formate up gluconeogenesis to produce sugar phosphates. The gluconeogenic products would then be catabolized through non-oxidative glycolysis, a novel pathway developed by Bogorad *et al.* [5] and optimized by Lin *et al.* [6] to generate acetate from glucose while minimizing carbon loss. A functional GNG-NOG cycle would thus assimilate formate throughout central carbon metabolism, providing a sustainable route for one-carbon compound consumption.

We analyzed the driving forces of the GNG-NOG cycle using Flux Balance Analysis (FBA) and the eQuilibrator thermodynamic calculator, finding that the proposed pathway is possible and overall forward-driven. Through FBA, we also identified key reactions that either promote or inhibit the pathway, which provided targets for genetic upregulation or knockout. Starting from an NOG-capable strain evolved by Lin *et al*., NOG26, we constructed six strains theoretically capable of conducting the GNG-NOG cycle. Then, through [U-$^{13}C_1$]formate, we traced its fate throughout the pathway. We found that the JXG2 strain incorporated formate into the GNG-NOG cycle significantly more than NOG26. While we were able to demonstrate that a strain turned the GNG-NOG cycle, further optimization is needed to increase pathway flux, as we identified significant formate incorporation in orthogonal pathways, including purine, THF, and CoA synthesis.

Overall, this project aims to compile recent applications of stable isotope tracers, particularly their ability to elucidate thermodynamic properties. We then applied these concepts towards the design of a formate assimilation pathway in *E. coli*. Insofar, we believe we have demonstrated the utility of these techniques towards the rational design of metabolic pathways, while also laying the groundwork for a pathway that can convert formate to useful products.

Chapter 2: Metabolic flux analysis and fluxomics-driven determination of reaction free energy using multiple isotopes

Adapted from: Metabolic flux analysis and fluxomics-driven determination of reaction free energy using multiple isotopes

Jimmy Xu, Julia Martien, Cole Gilbertson, Junyu Ma, Daniel Amador-Noguez, and Junyoung O. Park

Current Opinion in Biotechnology

Volume 64, 2020, Pages 151-160

2.1 Introduction

Quantitative analysis of metabolism has improved our understanding of biological systems [7,8], identified disease-state metabolic markers [9,10], and directed engineering of microorganisms for bioproduct synthesis [11-14]. Advances in analytical capabilities such as LC-MS played an important role by enabling comprehensive identification and quantitation of metabolites. Isotopes augment the utility of LC-MS by providing a means to trace metabolites through complex biochemical networks.

Since their discovery a century ago, isotopes have increasingly become the cornerstone of quantitative metabolic analysis (**Figure 1a**). The first use of stable isotopes to study metabolism was in 1935, when Rudolph Schoenheimer and David Rittenberg synthesized deuterated linseed oil to analyze lipid metabolism in mice. In 1955, Lifson *et al.* devised a doubly labeled water ($^{2}H_{2}{}^{18}O$) method for measuring average metabolic rate in animals, a technique used to this day. Today, a variety of single and multiple isotope tracing methods are used to quantify metabolite concentrations, fluxes, and free energies across a wide range of cell types, environments, and diseases [3,15-18].

Despite the long history and diversity of non-carbon isotopes, ^{13}C has taken center stage in modern metabolism research with the rise of metabolic flux analysis (MFA) [19]. Metabolic fluxes have been measured by both the kinetics of isotopes traveling through metabolic pathways and the steady-state labeling patterns created by different pathways shuffling the heavy and light atoms. However, based on the reaction atom mappings in MetaCyc [20], we found that a large portion of metabolic reactions (nearly 60%) do not involve the formation or cleavage of carbon-carbon bonds (**Figure 1b**). Non-carbon tracers are needed to examine these reactions.

Recently, improved flux measurement using multiple isotopes and high-resolution LC-MS has enabled quantitation of Gibbs free energy of reaction (ΔG) in central carbon metabolism using the extent of reaction reversibility [3]. This method computes ΔG based on the ratio of forward to reverse flux through a reaction, rather than the ratio of product to substrate concentrations. Quantitation of ΔG across multiple organisms and environments has revealed fundamental design principles: efficient enzyme usage, efficient energy yield, and fluid adaptivity in changing environments [1,2].

In this review, we highlight the recent developments in multi-isotope tracing that augment the toolset for determining metabolic fates, pathway contributions, and thermodynamic driving force. Integrating this information brings us closer to robust metabolic control in engineered cells and human diseases.

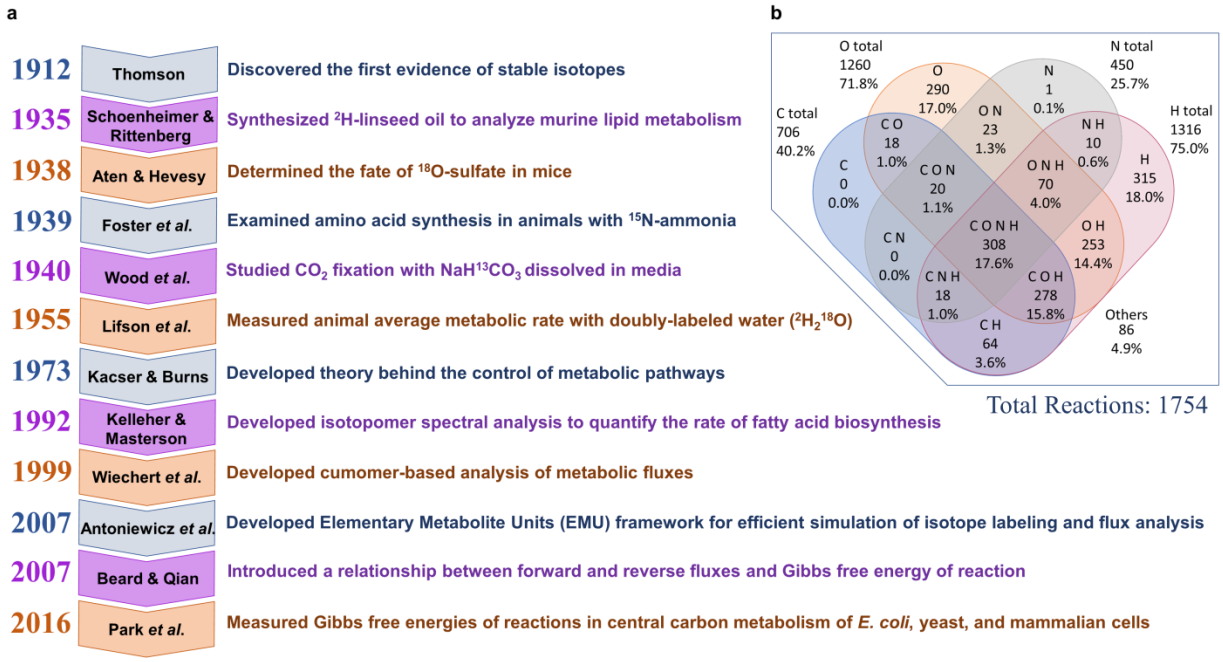


Figure 1 The past and the future of multi-isotope tracing in metabolism.

(a) Timeline of the use of stable isotope tracing in quantitative metabolic analysis. **(b)** Categorization of E. coli metabolic reactions by elements (carbon C, nitrogen N, oxygen O, and hydrogen H) that are involved in bond formation or cleavage. From the MetaCyc databases, the atom mapping of 1754 reactions in E. coli K-12 MG1655 revealed that oxygen and hydrogen are involved in more reactions than carbon. Many reactions involve transferring multiple elements. Therefore, tracing multiple isotopes would enhance quantitative metabolic analysis and deepen our understanding of metabolism.

2.2 Counting the reactions involving carbon, hydrogen, oxygen, and nitrogen transfers

To generate Figure 1B, a combination of the MetaCyc and EcoCyc databases were used in conjunction with MATLAB to determine the number of reactions involving the transfer of C, H, O, and/or N in *E. coli* K-12 substr. MG1655, totaling 1928 reactions. Of these, 174 reactions were ill-defined; they had more than one reaction arrow. The transfer of an atom was defined as the removal and subsequent addition of an atom from one molecule to a second molecule, where the second molecule has a different parent chain or ring or lacks a parent chain or ring. Transfer of the acidic hydrogen on a hydroxyl group was not counted. For example, in the phosphofructokinase reaction, oxygen would be considered as the only transferred atom (since we are not considering P). The algorithm began by iterating through each reaction known to *E.coli* and finding the corresponding reaction in SMILES syntax, an array of characters, from the MetaCyc database. For every reaction, the algorithm retrieved the reaction string and then divided the string so that each individual reactant and product was represented by its own string. The reaction was then classified as either an isomerization reaction or a cleavage/formation reaction. Separate functions were created for each case to operate on the reactant and product strings and determine whether a C, H, O, or N was transferred during the reaction. Since hydrogen atoms were not explicitly stated in reactions, each atom's connections were analyzed to determine the transfer of hydrogen in both isomerization and cleavage/formation reactions. For the latter, an additional string parsing was employed to determine the transfer of C, O, and N. To find which bonds break or form, for each numbered atom, its connectivity in the reactant was compared with its connectivity in the product. If chemical groups or moieties were transferred, all their constituent elements were counted.

2.3 Non-carbon isotope tracers

The major elements constituting the cell are carbon, hydrogen, oxygen, nitrogen, sulfur, and phosphorus [21]. Except for P, these elements have heavier stable isotopes suitable as mass-based metabolic tracers. Unlike commonly used ^{13}C, hydrogen and oxygen isotopes have a disadvantage in that many reactions dissociate these elements into the solvent water [22]. Deuterated molecules are also prone to undesirable kinetic isotope effects [23]. One downside of nitrogen and sulfur isotopes is that relatively few metabolic reactions involve these two elements [21]. However, these non-carbon isotopes offer distinctive benefits.

Deuterium is the ideal tracer for studying isomerase reactions and dehydrogenase reactions that involve redox cofactors such as NADH and NADPH [23,24]. Likewise, ^{18}O is ideal for non-radioactive tracing of phosphate [25]. While application of ^{18}O does result in kinetic isotope effects, they are substantially lower than those of deuterium and both can be adjusted for [23,26,27]. In addition, the mass addition of two neutrons by heavy ^{18}O replacing light ^{16}O (denoted as M+2) allows for easier identification of labeled metabolites compared to M+1 labeling by ^{13}C, ^{2}H, and ^{15}N. The use of ^{15}N benefits both metabolomics and proteomics, and provides a simple method to trace the flux of amino acids and nucleotides [28,29]. In combination with ^{13}C, non-carbon isotopes enhance our ability to study the kinetics, thermodynamics, and physiological function of broader metabolic networks. In the future, we envision that the utility of ^{2}H and ^{18}O may be comparable to that of ^{13}C due to their prevalence in biological reactions [23-25,30] (**Figure 1b**). Thus, the increased use of non-carbon isotopes would expand the scope of quantitative metabolic analysis.

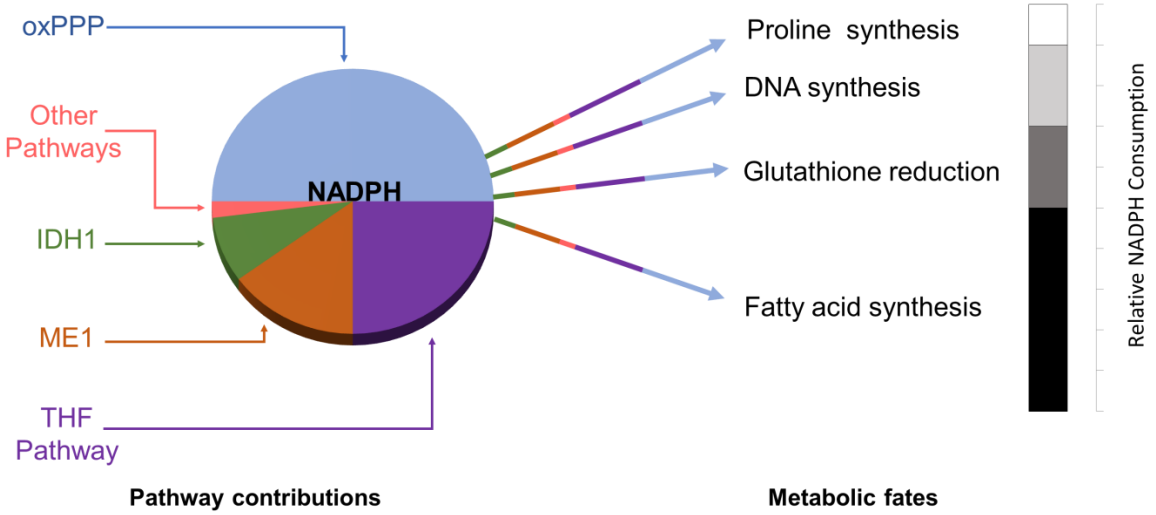


Figure 2 Applications of metabolic flux analysis in metabolism.

Isotope tracing reveals relative contributions of different metabolic pathways to the production of key metabolites, and the metabolic fates of the key metabolites along. For example, NADPH in mammalian cells is generated largely by oxidative pentose phosphate pathway (oxPPP) and THF pathway, with also contributions also from isocitrate dehydrogenase (IDH1) and malic enzyme 1 (ME1). The generated NADPH is used in various anabolic pathways, such as proline, DNA, and fatty acid synthesis, as well as combatting oxidative stress via reducing glutathione. Adapted from Refs. [24,31]. These values vary depending on cell types, organisms, and environments.

2.4 Tracing metabolic fates

Following the propagation of isotopes through metabolism reveals the connectivity between metabolites and the activity of downstream pathways (**Figure 2**). In metabolic engineering, this information is particularly helpful in identifying the steps of synthesis pathways and in maximizing the conversion of substrates into products. Lanfermann *et al.* elucidated the odorant sotolon synthesis pathway in the fungus *Laetiporus sulphureus* [32]. They found that [U-$^{13}C_6$]isoleucine, not [U-$^{13}C_6$]leucine, produced ^{13}C-labeled sotolon. Using $H_2{}^{18}O$, they were able to attribute sotolon's oxygen atoms to hydratase versus oxygenase [32].

As water is involved in many reactions, heavy water can test the activity of many parts of metabolism. In differentiating 3T3-L1 preadipocytes, Kirkwood *et al.* traced the fate of $H_2{}^{18}O$ throughout metabolism [25]. The labeled oxygen was incorporated into choline-containing metabolites (e.g., phosphocholine and CDP-choline, which likely carry ^{18}O-labeled phosphate), glutamate, and short peptides, tying elevated activity of choline metabolism, glutaminolysis, and proteasomes to preadipocyte differentiation [25]. In another study, Kloehn *et al.* employed ^2H_2O to quantify the turnover of *Leishmania* parasite's DNA, RNA, proteins, and lipids in murine lesions by measuring the kinetics of ^2H incorporation [33].

At metabolic branching points, isotopes can be used to differentiate the paths that metabolites take. One important example is the fate of glucose through glycolysis versus oxidative pentose phosphate pathway (oxPPP). The oxPPP flux corresponds to how much glucose is diverted to produce NADPH. [1,2-$^{13}C_2$]glucose facilitates quantitation of the two pathway fluxes as it loses the first carbon as $^{13}CO_2$ in the last step of oxPPP, resulting in M+1 pentose phosphate and subsequently M+1 lower glycolytic metabolites whereas glycolysis would produce M+2 and M+0 products [34].

Isotopes can be used to discover previously unknown fates; for example, Morscher *et al.* identified a key fate of one-carbon units [35]. Using [2,3,3-$^2H_3$]serine on HCT116 CRISPR knockout cell lines, they observed that the mitochondrial folate enzyme serine hydroxymethyltransferase 2 (SHMT2) was important for oxidative phosphorylation and cellular respiration. In SHMT2 knockout cells, [2,3,3-$^2H_3$]serine generated M+2 dTTP, suggesting cytosolic folate pathway flux, whereas in SHMT2 re-expression, M+1 dTTP appeared due to mitochondrial folate pathway [36]. They found that the ^{13}C of [3-$^{13}C_1$]serine generates, via

SHMT2, ^{13}C-labeled 5,10-methylene tetrahydrofolate (THF), which went on to produce ^{13}C-labeled 5-taurinomethyluridine (τm^5U), modifying mitochondrial tRNAs that translate specific guanosine-ending codons such as AAG (lysine) and UUG (leucine) in mitochondria [35].

In animals, isotopes enable tracking the flow of nutrients to different tissues and organs [28,37,38]. Neinast *et al.* carried out *in vivo* isotope tracing to study whole-body branched chain amino acid (BCAA) metabolism [28]. While elevations in BCAAs are implicated in cancer, diabetes, and heart failure, how the body maintains homeostasis of their levels is incompletely understood [28,39]. After introduction of ^{13}C-BCAAs into mice and measurement of blood and tissue metabolite labeling, they quantified the whole-body distribution of BCAA oxidation and protein synthesis fluxes in different tissues, finding that BCAA degradation products comprised 20% TCA cycle carbons in the pancreas and that insulin increased the rate of BCAA oxidation in striated muscle [28].

Tracing the fate of a drug is an important prerequisite for safely administering it. Fosmetpantotenate (RE-024) was an investigational therapy in 2019 phase 3 clinical trials for treating pantothenate kinase-associated neurodegeneration (PKAN). The drug was designed to replenish coenzyme A (CoA) by effectively supplying phosphopantothenate, a CoA-precursor that was low due to defective pantothenate kinase-2 in PKAN [40]. Upon treating cells and mice with the ^{13}C, ^{15}N, and ^{18}O-labelled fosmetpantotenate, Elbaum *et al.* observed labeled CoA [40]. They showed that approximately half of CoA originated from fosmetpantotenate in the mouse liver and brain, restoring CoA synthesis as intended [40].

Exemplifying the utility of ^{15}N-tracers, Spinelli *et al.* found an alternative fate of ammonia, which is typically cleared from mammals as a waste [29]. They used [γ-$^{15}N_1$]- and [U-$^{13}C_5{}^{15}N_2$]-

glutamine tracers to track the fate of ammonia in breast cancer cell lines T47D and MCF7, finding that glutamate dehydrogenase (GDH) efficiently recycled ammonia into glutamate, subsequently generating ^{15}N-labeled proline and aspartate [29]. Tracing the fate of ^{15}N in mice intraperitoneally injected with $^{15}NH_4Cl$, they found higher fractions of ^{15}N-labeled proline and aspartate in T47D xenograft tumors than in liver or plasma, indicating that ammonia recycling plays a role in cancer metabolism [29].

2.5 Quantifying pathway contributions

Cellular metabolism provides multiple ways to generate key metabolites, and adaptations to environmental changes often involve changing the relative contributions of alternative pathways. For example, organic cofactors such as ATP and NADPH – for cell maintenance, biomass production, and bioproduct synthesis – are produced by varying contributions of several metabolic pathways (**Figure 2**). Depending on environmental conditions such as oxygen availability, the dominant pathway contributing to cofactor production may change. The use of multiple tracers helps quantify fluxes through these pathways, improving our understanding of cellular physiology and adaptation.

Electron carriers NADH and NADPH are produced by various reactions in central carbon and one-carbon metabolism, fueling respiratory ATP generation, anabolism, and antioxidant defense [31]. As reducing these cofactors involves hydride transfers, deuterium is ideal for tracing NADH and NADPH production. Chen *et al.*, using [2,3,3,4,4-$^2H_5$]glutamine and [3-$^2H_1$]glucose, studied three major cytosolic NADPH-generating routes in mammalian cells: oxPPP, malic enzyme 1 (ME1), and isocitrate dehydrogenase I (IDH1) [24]. In HCT116 colon cancer cells, they observed that ~20% of NADPH was labeled from [3-$^2H_1$]glucose, indicating that oxPPP

contributes to generating ~40% of NADPH. [2,3,3,4,4-$^{2}H_5$]glutamine, which can label NADPH via ME1 or IDH1, resulted in an increased fraction of labeled NADPH in cells with glucose-6-phosphate dehydrogenase (G6PD) knockout, which blocks oxPPP, compared to wild-type cells. While these results indicated that oxPPP was the dominant NADPH production pathway and ME1 and IDH1 recover production upon its impairment, G6PD knockout also had drastic metabolic consequences such as accumulation of $NADP^+$ and dihydrofolate. These findings revealed an $NADP^+$-dependent inhibition of dihydrofolate reductase (DHFR), tying folate metabolism to oxPPP [24].

Deuterium tracing can also be used to determine how NADPH production routes depend on cell types and environments. For example, Liu *et al.* traced NADPH production in differentiating 3T3-L1 mouse adipocytes. They identified ME1 as the main source of NADPH under normoxia whereas oxPPP was the main NADPH source under hypoxia [23]. They used [2,2,3,3-$^{2}H_4$]dimethyl-succinate, a membrane-permeable succinate analog, and [1-$^{2}H_1$]-, [3-$^{2}H_1$]-, and [4-$^{2}H_1$]glucose to distinguish between NADPH generated from malate via ME1 and NADH generated in the oxPPP via G6PD and 6-phosphogluconate dehydrogenase.

Cellular energy in the form of ATP is produced in glycolysis, the TCA cycle, and oxidative phosphorylation using various nutrient sources [41]. At the hub of metabolism, many nutrients can turn the TCA cycle to produce ATP via respiration. To quantify major nutrient contributions to the TCA cycle in SFxL human glioma cells, Yang *et al.* measured labeled citrate originating from [U-$^{13}C_6$]glucose or [U-$^{13}C_5$]glutamine [42]. Upon inhibition of mitochondrial pyruvate carrier by UK5099 treatment, citrate labeling from glucose decreased, but increased from glutamine. To enter the TCA cycle, glutamine goes through glutamate and α-ketoglutarate (αKG) via either glutamate dehydrogenase (GDH) or transaminase. To differentiate these paths, they fed

cells [α-$^{15}N_1$]glutamine and measured $^{15}NH_4^+$ versus [$^{15}N_1$]alanine. UK5099 treatment resulted in increased $^{15}NH_4^+$ production and thus increased glutamine contribution to the TCA cycle via activating GDH, which sustained cell growth [42].

Seim *et al.* observed dynamic reprogramming in the TCA cycle of macrophages stimulated with lipopolysaccharide (LPS) and interferon-γ (IFN- γ) [43]. They used [U-$^{13}C_6$]glucose, [U-$^{13}C_5{}^{15}N_2$]glutamine, [U-$^{13}C_2$]acetate, and [U-$^{13}C_{16}$]palmitate to quantify TCA cycle contributions and fluxes during the metabolic transition [43]. In the early stage (6-24 hours), IDH1 flux was abolished, causing *cis*-aconitate to accumulate and divert to itaconate. Itaconate blocks succinate dehydrogenase and prolyl hydroxylases (PHDs), which resulted in succinyl-CoA, acetyl-CoA, and S-adenosyl methionine accumulation, potentially affecting posttranslational modification. PHD inhibition increases expression of hypoxia-inducible factor-1α (HIF-1α) [43]. HIF-1α and decreased lipoic acid production downregulate pyruvate dehydrogenase (PDH) and αKG dehydrogenase (αKGDH) complexes. Subsequently in the "late" (48-72 hours) phase, TCA metabolites trend towards pre-stimulation levels [43].

In addition to cofactors and TCA intermediates, key chemical moieties are produced via different pathways. Spanning multiple organelles, one-carbon metabolism supports methylation and nucleotide biosynthesis. Folate intermediates, such as 5,10-methenyl THF, 5-methyl THF, and 5,10-methylene THF, differ by oxidation states; as such, deuterium facilitates step-by-step tracing. Ducker *et al.* identified that HCT116 colon cancer cells produced one-carbon units primarily through the mitochondrial pathway, as evidenced by largely M+1 dTTP from [2,3,3-$^2H_3$]serine [36]. Serine-derived formate in mitochondria can be incorporated into cytosolic one-carbon metabolism for nucleotide synthesis. Knockout of the mitochondrial pathway resulted in

M+2 dTTP from [2,3,3-$^{2}H_3$]serine, demonstrating that cytosolic production of one-carbon units can compensate for the defective mitochondrial pathway. [36].

2.6 Quantifying nutrient contributions

Mixed substrate utilization can be observed by cofeeding labeled and unlabeled substrates. Kukurugya *et al.* found that *Pseudomonas putida* grew faster when fed glucose and benzoate (lignin-derived aromatic) simultaneously than glucose alone [44]. Using [U-$^{13}C_6$]glucose and unlabeled benzoate, they found that glucose is catabolized via the non-canonical Entner-Doudoroff (ED) glycolytic pathway (89%) [44]. They observed the formation of M+4 fructose-1,6-bisphosphate (FBP) from [1,5,6-$^{13}C_3$]glucose, which implied the flux through ED glycolysis is directed towards upper Embden-Meyerhof-Parnas (EMP) glycolysis via reverse FBP aldolase reaction. The labeling patterns of the rest of central carbon metabolism suggested that glucose is confined to upper glycolysis and the PPP while benzoate supported the TCA cycle and lower glycolysis [44]. In another study, to understand the contributions of nutrient supplements to engineered *Yarrowia lipolytica*, Czajka *et al.* fed [U-$^{13}C_6$]glucose and unlabeled secondary carbon sources to cells and measured the dilution of isotopically labeled metabolites in different pathways, determining the relative importance of each nutrient [45]. By feeding a labeled substrate in a nutrient mixture, each substrate's contribution can be quantified.

Mammalian cells are surrounded by many nutrient sources due to their metabolic requirement. Different disease states may alter nutrient contributions. Elia *et al.* used individual nutrient labeling to discover the role of pyruvate in cancer-cell-dependent collagen hydroxylation that supports metastatic breast cell growth [46]. Multiple nutrients including lactate and glucogenic amino acids can initiate gluconeogenesis, which helps maintain systemic glucose homeostasis in

humans. In type 2 diabetes, dysregulated glucagon stimulation causes hepatic glucose overproduction by increased gluconeogenesis and glycogenolysis, causing hyperglycemia [47]. While lactate remained the main gluconeogenic substrate, using [U-$^{13}C_3$]lactate and [U-$^{13}C_5$]glutamine, Miller *et al.* found that glucagon stimulation in primary hepatocytes increased glutamine's contribution to gluconeogenesis to nearly the same extent as lactate's [47].

In animals, diet and macromolecule degradation supply a slew of metabolites in cells; thus, differentiating between diet, degradation, and *de novo* synthesis is an important task. In adipocytes, Wallace *et al.* observed *de novo* synthesis of monomethyl branched-chain fatty acids (mmBCFA), which were thought to be diet-derived [48]. Using $^{13}C^{15}N$-labeled branched chain amino acids (BCAAs) and $^{2}H_2O$, they quantified the contribution of newly synthesized versus diet-derived mmBCFAs *in vivo*.

In fasting nutrient-limited conditions, metabolites like acetate and ketone bodies can serve as nutrients. To identify the source of acetate, Liu *et al.* used various ^{13}C, ^{2}H, and ^{18}O tracers. Using [U-$^{13}C_3$]- and [3,3,3-$^{2}H_3$]-pyruvate, they showed that thiamine-dependent keto acid dehydrogenases (i.e. PDH and αKGDH) convert pyruvate into acetate [30]. Using $^{18}O_2$ and $H_2{}^{18}O_2$, they showed that reactive oxygen species also contribute to decarboxylation of pyruvate to acetate. The pyruvate-derived acetate contributed to generating acetyl-CoA, fatty acids, and other acetylated molecules and rescued ATP citrate lyase knockout cells [30].

In amino acid deficient environments, oncogenic Ras signaling stimulates extracellular protein uptake by inducing macropinocytosis, allowing pancreatic cancers to grow. In *K-Ras*-mutant pancreatic ductal adenocarcinoma cells, Nofal *et al.* quantified free amino acid uptake and intact protein scavenging fluxes by feeding $^{13}C^{15}N$-labeled amino acids and unlabeled serum proteins

[49]. While a prior study posited that mTORC1 activity inhibited cancer growth by suppressing degradation of scavenged proteins [50], Nofal *et al.* showed that the growth inhibition was instead due to imbalance between limited amino acid supply and high amino acid demand for translation [49]. In amino acid deprivation, partial mTOR inhibition slows protein synthesis rate and brings it on par with amino acid supply from extracellular protein scavenging and catabolism.

2.7 Quantifying thermodynamic driving force using metabolic fluxes

Thermodynamic quantitation is a critical component in analysis of any chemical or biochemical system including metabolism. Gibbs free energy of reaction indicates not only reaction feasibility but also enzyme efficiency based on reaction reversibility (**Figure 3a**). Furthermore, ΔG values are closely tied to metabolic control: the rate determining step of a pathway is largely thermodynamically forward driven ($\Delta G \ll 0$). A pathway step with a strong thermodynamic driving force will achieve a higher net flux given a fixed amount enzyme than one closer to equilibrium. A step closer to equilibrium (e.g., $\Delta G \approx -1$ kJ/mol) will be less enzyme efficient, and overexpressing its enzyme would bring it even closer to equilibrium ($\Delta G \approx 0$) without affecting the net flux (**Figure 3a**). Thermodynamic analysis can therefore provide unique insights into pathway design principles by identifying bottlenecks, pinpointing the enzymes for which changes in activity will have the largest effect on flux, and predicting the most efficient route for product synthesis. However, obtaining physiological ΔG values has been challenging because, traditionally, ΔG is calculated by:

$$\Delta G = \Delta G^{\circ\prime} + RT \, ln \, Q \qquad (1)$$

where $\Delta G^{\circ\prime}$ is ΔG at standard biochemical conditions, R is the gas constant, T is temperature in kelvin, and Q is the reaction quotient (the ratio of substrate to product activities, which are effective concentrations in a non-ideal solution). **Equation 1** is difficult to use in practice because accurate $\Delta G^{\circ\prime}$ values are difficult to obtain and it is technically challenging to measure the intracellular concentrations of all substrates and products of a reaction, especially considering the spatial localization or compartmentalization of metabolites and enzymes within a cell. Missing the concentration of any one metabolite involved in a reaction precludes ΔG calculation.

Recently, an alternative approach was developed for more direct measurement of physiological ΔG using reaction reversibility, here defined as the backward-to-forward flux ratio (J^-/J^+):

$$\Delta G = RT\ ln\ (J^-/J^+) \qquad (2)$$

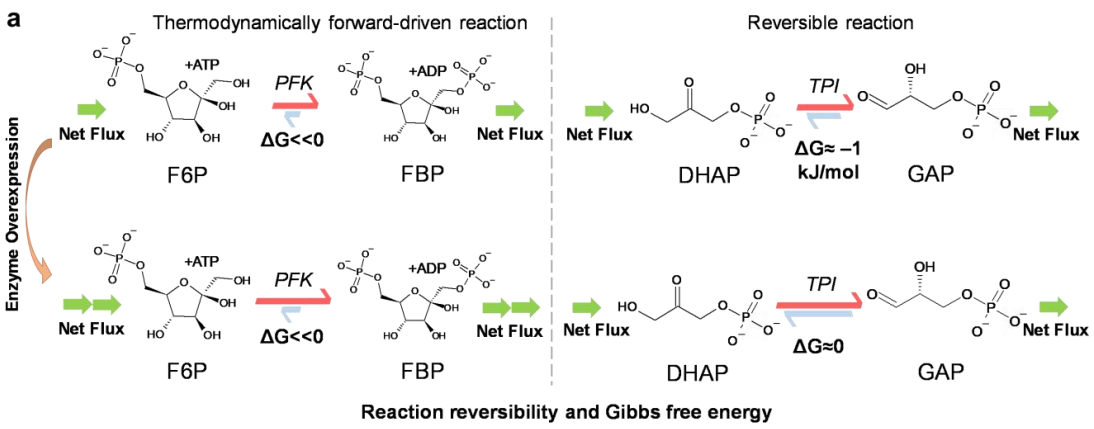
Equation 2 does not require $\Delta G^{\circ\prime}$ or intracellular metabolite concentrations, but instead requires measuring the backward to forward flux ratio of a given reaction, which can be done with isotope tracers.

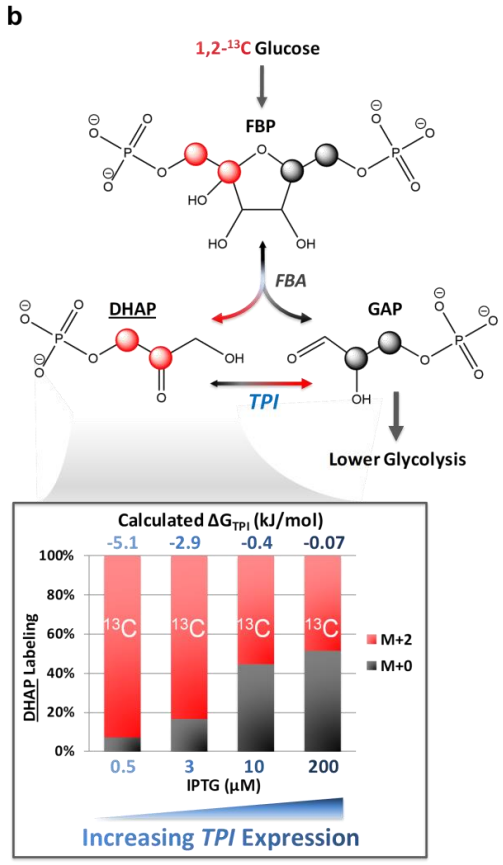
Substrates harboring isotopes at certain positions generate metabolites with distinct labeling patterns from the forward reaction versus the reverse reaction. For example, [1,2-$^{13}C_2$]glucose generates [1,2-$^{13}C_2$]FBP, which subsequently produces M+2 dihydroxyacetone phosphate (DHAP) and unlabeled (M+0) glyceraldehyde-3-phosphate (GAP) via FBP aldolase. The forward triosephosphate isomerase (TPI) reaction converts the M+2 DHAP to M+2 GAP, and DHAP remains labeled. However, the reverse TPI reaction converts M+0 GAP to M+0 DHAP (**Figure 3b**). The ratio of M+0 DHAP to M+2 DHAP can therefore be used to determine the backward-to-forward flux ratio of the TPI reaction. At low levels of TPI expression, little M+0 DHAP is observed, indicating a small reverse flux and a large forward driving force ($\Delta G < -5$

kJ/mol). On the other hand, increasing TPI expression increases M+0 DHAP since the increased enzyme capacity results in increased reverse flux, bringing the reaction closer to equilibrium [3].

An additional benefit of the flux-ratio-based ΔG quantitation is that, by combining **equations 1 and 2**, absolute metabolite concentrations that are previously unmeasured or imprecisely measured can be computed. Using this integrative approach, Park *et al.* measured physiological ΔG in *E. coli*, yeast, and mammalian central carbon metabolism while also obtaining comprehensive absolute metabolite concentrations [3]. The ΔG values showed that, across widely divergent organisms, available Gibbs free energy was relatively evenly distributed across reactions in glycolysis [3]. Unlike the textbook model, which considered all non-kinase steps to be near equilibrium ($\Delta G \approx 0$), the observed partitioning of thermodynamic driving force was proposed to provide a fitness advantage by more efficient enzyme usage [3].

In 2019, another Park *et al.* study measured ΔG of glycolytic reactions in different organisms and changing conditions using [5-$^{2}H_{1}$]glucose, which dissipates the deuterium faster from metabolites with increasing glycolytic reversibility (**Figure 3c**) [1]. They found that slow glycolysis possessed multiple steps that are near equilibrium. In *E. coli*, they found that an increase in nutrient availability (i.e., nitrogen or phosphorus upshift) increased glycolytic flux and thermodynamic driving force within 5 minutes [1]. Similarly, in mammalian cells, oligomycin treatment, which inhibits ATP synthase, increased both glycolytic flux and thermodynamic driving force within 30 minutes [1].


a
Thermodynamically forward-driven reaction
Reversible reaction
Enzyme Overexpression
Net Flux
+ATP
PFK
+ADP
F6P
FBP
ΔG<<0
TPI
DHAP
GAP
ΔG≈ –1 kJ/mol
ΔG≈0
Reaction reversibility and Gibbs free energy


b
1,2-13C Glucose
FBP
FBA
DHAP
GAP
TPI
Lower Glycolysis
Calculated ΔGTPI (kJ/mol)
-5.1
-2.9
-0.4
-0.07
DHAP Labeling
M+2
M+0
IPTG (μM)
0.5
3
10
200
Increasing TPI Expression

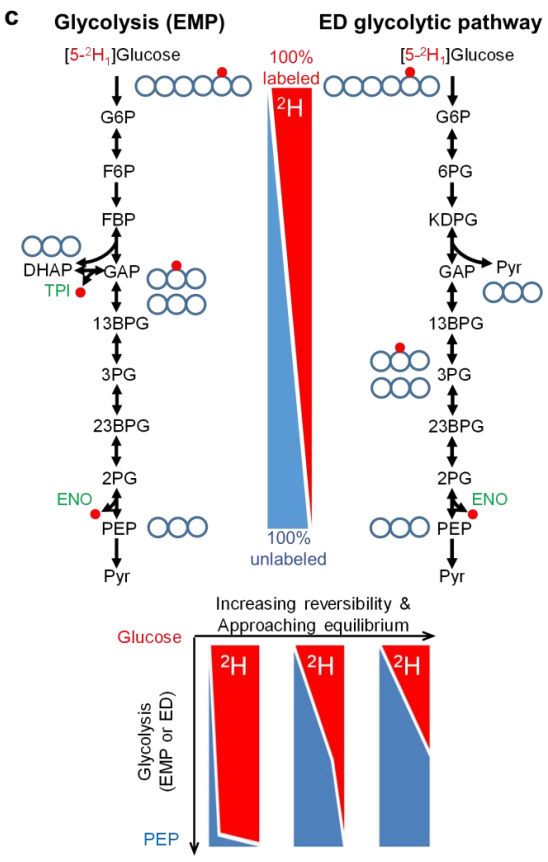

c
Glycolysis (EMP)
ED glycolytic pathway
[5-2H1]Glucose
100% labeled
100% unlabeled
G6P
F6P
FBP
DHAP
GAP
TPI
13BPG
3PG
23BPG
2PG
ENO
PEP
Pyr
6PG
KDPG
Increasing reversibility & Approaching equilibrium
Glucose
Glycolysis (EMP or ED)
PEP

Figure 3 Reaction reversibility reveals Gibbs free energy of reaction (ΔG)

(a) Quantifying the forward and reverse fluxes of individual reactions reveals Gibbs free energy of reaction (ΔG). Triosephosphate isomerase (TPI) is close to equilibrium while phosphofructokinase (PFK) is strongly forward driven. For TPI, enzyme overexpression brings the reaction closer to equilibrium, but results in little change to the net pathway flux. For PFK, enzyme overexpression results in an increase in net pathway flux. DHAP denotes dihydroxyacetone phosphate; GAP, glyceraldehyde 3-phosphate; F6P, fructose-6-phosphate; and FBP, fructose 1,6-bisphosphate. **(b)** In glycolysis, [1,2-$^{13}C_2$]glucose generates [1,2-$^{13}C_2$]FBP, which is split into M+2 DHAP and unlabeled GAP via aldolase. The reverse TPI reaction introduces unlabeled carbons to DHAP. Increasing TPI expression results in greater TPI reversibility since increasing the enzyme levels at a non-rate-determining step puts the spare enzymes to work in the unproductive reverse direction. Adapted from [3]. **(c)** In glycolysis, the deuterium of [5-$^{2}H_1$]glucose is lost to water via TPI and enolase (ENO). The reversibility of glycolytic reactions generates a descending gradient of deuterated metabolites. The greater the reversibility of glycolysis (equivalent to smaller thermodynamic driving force), the steeper the gradient of deuterium loss from glycolytic metabolites. Adapted from [1]. G6P denotes glucose-6-phosphate; 13BPG, 1,3-bisphosphoglycerate; 3PG, 3-phosphoglycerate; 23BPG, 2,3-bisphosphoglycerate; 2PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate; Pyr, pyruvate; 6PG, 6-phosphogluconate; KDPG, 2-keto-3-deoxy-6-phosphogluconate; EMP, Embden-Meyerhof-Parnas; ED, Entner-Doudoroff.

Together, the ΔG values calculated across conditions in both bacterial and mammalian cells indicate that operating glycolysis close to equilibrium may confer its own fitness advantage [1]. The underutilized pool of glycolytic enzymes may poise the cell to rapidly increase glucose consumption in response to environmental perturbations on a much faster timescale than transcriptional regulation [1].

Quantitation of ΔG provides new insights into the metabolism of non-model organisms. In the same 2019 Park *et al.* study, they measured ΔG of glycolytic reactions in *Clostridium cellulolyticum*, a bacterium adapted to growth on cellulose [1]. They found that the entire *C. cellulolyticum* glycolysis was near equilibrium at least in part due to the use of pyrophosphate-dependent PFK instead of canonical ATP-dependent PFK [1]. This finding revealed another potential fitness benefit associated with near equilibrium glycolysis: a greater ATP yield per glucose.

Using various ^{2}H and ^{13}C tracers, Jacobson *et al.* quantified ΔG in central carbon metabolism of *Zymomonas mobilis*, a bacterium that exclusively use the non-canonical ED glycolytic pathway (**Figure 3c**) [2]. Consistent with a prior computational study, they found that *Z. mobilis* ED pathway was substantially more thermodynamically favorable than canonical glycolysis [2,51]. This coincided with a higher rate of glucose consumption in *Z. mobilis* compared to organisms using canonical glycolysis [2]. The interdependency between the thermodynamic favorability of glycolysis and glycolytic flux highlights the importance of favorable thermodynamic driving force when attempting to maximize flux through engineered pathways.

In addition to determining ΔG values for entire pathways, differential metabolite labeling can be used in a focused context. For example, Enjalbert *et al.* used ^{13}C-labeled glucose and acetate to investigate the reversibility of acetate uptake in *E. coli*, where they determined that acetate was simultaneously produced and consumed at high rates during growth on a mixture of glucose and acetate [52]. Using a computational model based on labeling dynamics, they showed that this fast exchange of acetate enabled rapid changes in the direction of net acetate flux as determined by thermodynamic favorability [52]. They go on to describe the evolutionary advantage of a thermodynamically controlled carbon assimilation pathway, showing that *E. coli* is able to respond to increased extracellular concentrations of acetate within two minutes [52].

2.8 Summary of articles reviewed

In this chapter, we reviewed recent studies utilizing various isotope tracers. For studying metabolism, we have divided studies between those examining the fate of key metabolites, those examining the relative contributions of different fed nutrients to cells, those examining the relative contributions of different pathways towards different key metabolites such as NADPH

and ATP, and finally those examining the thermodynamics of metabolic reactions. While these categories do not encompass all uses for isotope tracers in metabolomics, we believe them to be a good introduction for those relatively unfamiliar with the field. We have summarized recent studies, including the application, tracer used, traced metabolites, and experimental rationales of the studies, in **Table 1**.

Table 1: Application of isotope tracers towards measuring key metabolites to obtain fluxes, free energies, and other useful information

All abbreviations for Table 1 can be found in Appendix A.

Application	Tracer	Key Traced Metabolites	Rationale
Studies on metabolic fate			
Elucidating production pathway of the odorant sotolon in the fungus *Laetiporus sulphureus* *Pathway(s)*: Sotolon production	[U-$^{13}C_6$]isoleucine	^{13}C-sotolon(M+6); ^{13}C-3-methyl-dihydro-2(3H) furanone(M+5)	Leucine and isoleucine were considered possible precursors to sotolon, with this study confirming isoleucine leads to sotolon. Labeled water identified hydratase versus oxygenase as an oxygen source for the furanone ring [32].
	[2-$^{13}C_1$]leucine	^{13}C-3-hydroxy-3-methyl-butanoic acid(M+1); ^{13}C-4-methyl-dihydro-2(3H) furanone(M+1)	
	[U-$^{18}O_1$]water	^{18}O-sotolon(M+2)	
Identifying changes in cellular metabolism of differentiating 3T3-L1 preadipocytes *Pathway(s)*: TCA, FA, glu, CDP-choline cycle	[U-$^{18}O_1$]water	^{18}O-peptides(M+2); ^{18}O-choline(M+2); ^{18}O-CDP-choline(M+2); ^{18}O-phosphocholine(M+2); ^{18}O-glutamate(M+2)	^{18}O-labeled water was used for untargeted metabolic screening in preadipocytes. Significant labeling was found in short-chain peptides, CDP-choline cycle, and glutaminolysis [25].
Growth rate, fatty acid, amino acid and RNA/DNA synthesis rates in *Leishmania mexicana* *Pathway(s)*: FA, RNA/DNA, AA	[U-$^{13}C_6$]glucose	^{13}C-deoxyribose(M+5)	Glucose examined uniformly-labeled deoxyribose incorporation into DNA. Heavy water measured turnover rates of fatty acids, RNA/DNA, and all amino acids [33].
	5% ^2H_2O	^{2}H-linoleic acid; ^{2}H-deoxyribose; ^{2}H-ribose; ^{2}H-alanine	
Folate-dependent tRNA methylation and dTTP synthesis upon single gene knockouts in HCT116 and HEK293T cell lines *Pathway(s)*: TCA, oxidative phosphorylation, one-carbon, tRNA modification	[U-$^{13}C_6$]glucose	^{13}C-citrate(M+2)	Labeled glucose and glutamine were used to examine respiration, represented by citrate labeling. ^{13}C-labeled serine and methionine confirmed incorporation of one carbon from serine into the τm^5U group and from methionine into the f^5C group of tRNAs. Deuterated serine tracked dTTP production through mitochondrial (M+1) and cytosolic (M+2) routes [35].
	[U-$^{13}C_5$]glutamine	^{13}C-citrate(M+4, M+5)	
	[3-$^{13}C_1$]serine	^{13}C-p-τm^5U(M+1); ^{13}C-f^5C(M+1)	
	[U-$^{13}C_5$]methionine	^{13}C-p-τm^5U(M+1); ^{13}C-f^5C(M+1)	
	[2,3,3-$^2H_3$]serine	^{2}H-dTTP(M+1, M+2)	
Breakdown of BCAAs in mice *Pathway(s)*: TCA, BCAA	[U-$^{13}C_5$]valine	^{13}C-α-ketoisovaleric acid (M+5); ^{13}C-3-HIB(M+4); ^{13}C-succinate(M+3); ^{13}C-malate(M+3)	Labeled BCAAs were tracked to BCKAs and TCA metabolites. Succinate and malate were labeled M+2 or M+3 depending on their

catabolism	[U-$^{13}C_6$]leucine	^{13}C-α-ketoisocaproic acid (M+6); ^{13}C-succinate(M+2); ^{13}C-malate(M+2)	synthesis pathway, with M+2 arising from BCAA-derived acetyl-CoA and M+3 from BCAA-derived succinyl-CoA [28].
	[U-$^{13}C_6$]isoleucine	^{13}C-α-keto-beta-methylvaleric acid (M+6); ^{13}C-succinate(M+2, M+3); ^{13}C-malate(M+2, M+3)	
Breakdown of drug fosmetpantotenate, or RE-024, in mice, monkeys, rats and in human neuroblastoma cells for treatment of CoA-deficiency *Pathway(s)*: RE-024 breakdown	[$^{13}C_3$, $^{15}N_1$, $^{18}O_1$]RE-024	^{13}C, ^{15}N, ^{18}O-CoA(M+4, M+6); ^{13}C, ^{15}N, ^{18}O-PPA(M+4, M+6); ^{13}C, ^{15}N-PA(M+4)	Carbon and nitrogen labels were applied to track the drug; the oxygen label differentiated direct incorporation and re-phosphorylation of the drug towards forming CoA. Labels were tracked to two RE-024 intermediates and CoA [40].
Nitrogen assimilation and recycling of ammonia through glutamate in T47D and MCF7 breast cell lines and in mice *Pathway(s)*: TCA, nucleotide synthesis, AA	[U-$^{13}C_5$,$^{15}N_2$]glutamine	^{13}C, ^{15}N-glutamate(M+1, M+6); ^{13}C-αKG(M+5)	Uniformly labeled glutamine was used to determine relative rate of ammonia recycling, with M+6 glutamate coming directly from glutamine while significant M+1 glutamate formed from ammonia recycling. ^{15}N-labeled glutamine was used to track recycled ammonia into various amino acids and nucleotides. Ammonium chloride was used to determine if free ammonia could be assimilated [29].
	[γ-$^{15}N_1$]glutamine	^{15}N-glutamate(M+1); ^{15}N-UMP(M+1); ^{15}N-proline(M+1); ^{15}N-aspartate(M+1)	
	[U-$^{15}N_1$]ammonium chloride	^{15}N-glutamate(M+1); ^{15}N-glutamine(M+1, M+2); ^{15}N-proline(M+1); ^{15}N-aspartate(M+1)	
Studies on pathway contribution			
Determining pathway contributions toward NADPH production after single and double gene CRISPR knockouts in various mammalian cell lines *Pathway(s)*: Glycolysis, oxPPP, TCA, ME1, one-carbon, FA	[U-$^{13}C_6$]glucose	^{13}C-palmitate(M+N, where N is even and 2≤N≤16)	[U-$^{13}C_6$]glucose was used to track fatty acid synthesis. [U-$^{13}C_5$]glutamine was used to examine prevalence of ME1 (indicated by M+4 malate and M+3 pyruvate) versus reductive carboxylation coupled to ATP citrate lyase (indicated by M+3 malate and M+2 pyruvate), while M+3 lactate was measured as the upper limit of ME1 flux. Labeled palmitate was fed to measure fatty acid oxidation. Deuterated glucose, glutamine, and water accounted for oxPPP, TCA, and flavin enzyme contribution to cytosolic
	[U-$^{13}C_5$]glutamine	^{13}C-malate(M+3, M+4); ^{13}C-pyruvate(M+2, M+3); ^{13}C-lactate(M+3)	
	[U-$^{13}C_{16}$]palmitate	^{13}C-palmitoleic acid(M+16); ^{13}C-citrate(M+2)	
	[2,3,3,4,4-$^{2}H_5$]glutamine	^{2}H-NADPH(M+1); ^{2}H-malate(M+2); ^{2}H-isocitrate(M+5)	
	[3-$^{2}H_1$]glucose	^{2}H-NADPH(M+1)	
	[U-$^{2}H_2$]water	^{2}H-NADPH(M+1)	
	[2,3,3-$^{2}H_3$]serine	^{2}H-NADPH(M+1);	

			NADPH production respectively. Deuterated serine accounted for one-carbon contribution to NADPH [24].
Determining metabolite contributions toward NADPH production under normoxic and hypoxic conditions in differentiating 3T3-L1 adipocytes *Pathway(s)*: Glycolysis, oxPPP, TCA, one-carbon, FA, nucleotide synthesis	[U-$^{13}C_6$]glucose	^{13}C-pyruvate(M+3); ^{13}C-malate(M+2, M+3); ^{13}C-acetyl-CoA(M+2); ^{13}C-palmitate(M+N, where N is even and 2≤N≤16)	Uniformly-labeled glucose and glutamine were used to measure fluxes throughout glycolysis and TCA cycle. Fatty acid formation was also traced. Using [1,2-$^{13}C_2$]glucose, M+1 metabolites diverted through oxPPP was differentiated from M+2 metabolites that travel down glycolysis. Labeled serine was used to find contribution of mitochondrial and cytosolic one-carbon pathways towards dTTP. Labeled DMS and [4-^{2}H]glucose were shown to be novel tracers for ME1 flux and its contribution to NADPH. [1-$^{2}H_1$]glucose and [3-$^{2}H_1$]glucose labeled NADPH produced in the first and third step of oxPPP respectively, and their incorporation into fatty acids was measured [23].
	[U-$^{13}C_5$]glutamine	^{13}C-3PG(M+2, M+3); ^{13}C-pyruvate(M+2, M+3); ^{13}C-malate(M+4); ^{13}C-acetyl-CoA(M+2); ^{13}C-palmitate(M+N, where N is even and 2≤N≤16)	
	[1,2-$^{13}C_2$]glucose	^{13}C-F6P(M+1, M+2); ^{13}C-FBP(M+1, M+2); ^{13}C-R5P(M+1)	
	[2,3,3-$^{2}H_3$]serine	^{2}H-NADPH(M+1); ^{2}H-dTTP(M+1, M+2)	
	[2,2,3,3-$^{2}H_4$]DMS	^{2}H-NADPH(M+1); ^{2}H-succinate(M+2, M+4); ^{2}H-malate(M+1, M+2); ^{2}H-palmitate(M+1, M+2)	
	[4-$^{2}H_1$]glucose	^{2}H-NADPH(M+1); ^{2}H-NADH(M+1); ^{2}H-malate(M+1); ^{2}H-palmitate(M+1, M+2)	
	[1-$^{2}H_1$]glucose; [3-$^{2}H_1$]glucose	^{2}H-NADPH(M+1); ^{2}H-G6P(M+1); ^{2}H-6PG(M+1); ^{2}H-palmitate(M+1 to M+4)	
Quantifying contribution of glutamine versus glucose to TCA cycle, fatty acid synthesis, and acetyl-CoA after blockade of pyruvate transporter in SFxL glioma cells *Pathway(s)*: Glycolysis, TCA, FA, AA	[U-$^{13}C_5$]glutamine; [3-$^{13}C_1$]pyruvate	^{13}C-malate(M+4); ^{13}C-citrate(M+4, M+5, M+6); ^{13}C-acetyl-CoA(M+2); ^{13}C-palmitate(M+N, where N is even and 2≤N≤16)	Uniformly-labeled glutamine was used to trace TCA cycle and fatty acid synthesis with and without single-labeled pyruvate. Addition of labeled pyruvate increased the M+5 fraction of citrate. Uniformly-labeled glucose was used to compare glucose and glutamine-derived anaplerosis and measure fatty acid synthesis. [1-$^{13}C_1$]pyruvate was fed to determine pyruvate consumption upon treatment of UK5099. [3-$^{13}C_1$]glutamine and [5-$^{13}C_1$]glutamine accounted for acetyl-CoA production via PDH and ACL respectively. [α-
	[U-$^{13}C_6$]glucose	^{13}C-pyruvate(M+3); ^{13}C-citrate(M+2); ^{13}C-acetyl-CoA(M+2); ^{13}C-palmitate(M+N, where n is even and 2≤N≤16)	
	[1-$^{13}C_1$]pyruvate	^{13}C-lactate(M+1); ^{13}C-alanine(M+1); ^{13}C-bicarbonate(M+1)	
	[3-$^{13}C_1$]glutamine	^{13}C-pyruvate(M+1); ^{13}C-acetyl-CoA(M+1); ^{13}C-palmitate(M+1, M+2)	
	[5-$^{13}C_1$]glutamine	^{13}C-pyruvate(M+1); ^{13}C-acetyl-CoA(M+1); ^{13}C-palmitate(M+1)	
	[α-$^{15}N_1$]glutamine	^{15}N-glutamate(M+1); ^{15}N-ammonium(M+1); ^{15}N-alanine(M+1)	

			$^{15}N_1$]glutamine was used to determine contribution of GDH versus transaminase [42].
Reprogramming of TCA cycle in developing mouse macrophages after stimulation by lipopolysaccharide and interferon-γ *Pathway(s)*: Glycolysis, TCA, FA, AA	[U-$^{13}C_6$]glucose	^{13}C-acetyl-CoA(M+2) ^{13}C-citrate(M+2, M+5); ^{13}C-aconitate(M+2, M+5); ^{13}C-itaconate(M+1, M+4); ^{13}C-αKG(M+2, M+4); ^{13}C-malate(M+2, M+3)	Labeled glucose and glutamine were used to probe fluxes within TCA cycle. For glucose, the lower weight metabolites derived from acetyl-CoA, while the higher weight metabolites arose from pyruvate's conversion to malate. Certain labelling patterns, such as the decrease in M+4 malate after labeled glutamine introduction, indicated reprogramming of TCA cycle. Labeled palmitate and acetate accounted for alternative sources of acetyl-CoA [43].
	[U-$^{13}C_5$, $^{15}N_2$]glutamine	^{13}C-acetyl-CoA(M+2); ^{13}C-citrate(M+4, M+5); ^{13}C-αKG(M+5); ^{13}C-malate(M+4)	
	[U-$^{13}C_{16}$]palmitate	^{13}C-acetyl-CoA(M+2)	
	[U-$^{13}C_2$]acetate	^{13}C-acetyl-CoA(M+2); ^{13}C-citrate(M+2)	
Determining effect of CRISPR gene knockouts on one-carbon pathway in HCT116 colon cells *Pathway(s)*: Glycolysis, oxPPP, TCA, one-carbon, nucleotide synthesis, AA	[U-$^{13}C_1$]formate	^{13}C-10-formyl THF(M+1); ^{13}C-dTTP(M+1)	Labeled formate accounted for ATP formation via 10-formyl THF, as well as dTTP formation. ^{13}C-serine tracked fate of serine to formate and dTTP. ^{2}H-serine helped differentiate mitochondrial and cytosolic dTTP, and measured fatty acid synthesis, which was labeled via NADP^{2}H. Labeled malate was used as a reporter metabolite for NADH contribution from malate dehydrogenase [36].
	[3-$^{13}C_1$]serine	^{13}C-formate(M+1); ^{13}C-dTTP(M+1)	
	[2,3,3-$^{2}H_3$]serine	^{2}H-NADPH(M+1); ^{2}H-dTTP(M+1, M+2); ^{2}H-formate(M+1); ^{2}H-malate(M+1); ^{2}H-palmitate(M+1 to M+3)	
Studying impact of BCAT1/BCAT2 and IDH1/IDH2 inhibition on nitrogen metabolism, particularly glutamate synthesis, in various mammalian cell lines and in mice *Pathways(s)*: TCA, AA, BCAA catabolism	[U-$^{13}C_5$]glutamine	^{13}C-glutamate(M+5); ^{13}C-αKG (M+5); ^{13}C-2HG(M+5); ^{13}C-citrate(M+4, M+5)	Uniformly-labeled glutamine tracked glutaminase activity in IDH mutants via its incorporation into glutamate, αKG, and 2HG. Uniformly-labeled leucine and isoleucine were compared against glutamine to measure BCAA consumption, represented by labeled citrate enrichment. ^{15}N-labeled leucine was fed to measure BCAT activity, along with downstream products of glutamate. Contribution of other
	[U-$^{13}C_6$]leucine	^{13}C-citrate(M+2)	
	[U-$^{13}C_6$]isoleucine	^{13}C-citrate(M+2, M+3)	
	[1-$^{13}C_1$]glucose	^{13}C-glutamate(M+1)	
	[U-$^{15}N_1$]leucine	^{15}N-α-ketoisocaproic acid (M+1); ^{15}N-glutamate(M+1); ^{15}N-aspartate(M+1); ^{15}N-GSH(M+1)	
	[$^{15}N_1$]isoleucine; [$^{15}N_1$]valine; [α-$^{15}N_1$]glutamine; [$^{15}N_1$]alanine	^{15}N-glutamate(M+1)	

	[γ-$^{15}N_1$]glutamine	^{15}N-ammonia(M+1)	glutamate sources was also studied, including labeled glucose, isoleucine, valine, glutamine, and alanine. [γ-$^{15}N_1$]Glutamine was used to determine contribution of GDH compared to transamination of GLU and PYR to αKG and alanine [9].
PP2A-mediated redirection of glucose to pentose phosphate pathway in B cells *Pathway(s)*: Glycolysis, oxPPP, nucleotide synthesis	[1,2-$^{13}C_2$]glucose	^{13}C-3PG(M+1, M+2); ^{13}C-lactate(M+1, M+2); ^{13}C-R5P(M+1, M+2); ^{13}C-IMP(M+1, M+2)	[1,2-$^{13}C_2$]glucose introduction resulted in M+1 metabolites derived from oxPPP due to loss of the labeled first carbon, while M+2 metabolites derived from glycolysis [34].
Collagen hydroxylation by αKG-dependent P4HA in MCF10A and 4T1 breast cells *Pathway(s)*: TCA, pyruvate to alanine transamination	[U-$^{13}C_6$]glucose; [U-$^{13}C_3$]pyruvate	^{13}C-αKG(M+2, M+3); ^{13}C-alanine(M+3)	Transamination of glutamate and pyruvate to αKG and alanine was examined as a major source of αKG. αKG originated from acetyl-CoA(M+2), malic enzyme (M+3), and glutamine (M+5). Alanine labeling, which was only present from labeled glucose and pyruvate, was indicative of transamination [46].
	[U-$^{13}C_5$]glutamine	^{13}C-αKG(M+5)	
Studies on nutrient contribution			
Mapping simultaneous glucose and benzoate usage in *Pseudomonas putida* *Pathway(s)*: Glycolysis, ED pathway, TCA, nucleotide synthesis, AA, PPP, benzoate catabolism	[U-$^{13}C_6$]glucose	^{13}C-serine(M+3); ^{13}C-IMP(M+5, M+7, M+10); ^{13}C-glycerol-phosphate(M+3)	Uniformly labeled glucose was fed either independently or with unlabeled benzoate, to measure nutrient fate into biomass biosynthesis including amino acids, membrane precursors, and nucleotides. [1,5,6-$^{13}C_3$]glucose was fed with unlabeled benzoate, finding M+3 FBP from glucose and M+4 FBP from reversal of ED to EMP glycolysis. Through labeling of central carbon pathways, glucose-derived carbons were shown to be largely confined to upper glycolysis and oxPPP while benzoate-derived carbons were largely confined to TCA cycle [44].
	[1,5,6-$^{13}C_3$]glucose	^{13}C-6PG(M+3); ^{13}C-FBP(M+3, M+4); ^{13}C-R5P(M+2, M+4); ^{13}C-OAA(M+1, M+2); ^{13}C-acetyl-CoA(M+2); ^{13}C-pyruvate(M+1, M+2)	
Engineering heterologous production of β-	[U-$^{13}C_6$]glucose	^{13}C-3PG(M+3); ^{13}C-citrate(M+2); ^{13}C-acetyl-CoA(M+2);	Cells were initially cultured with labeled glucose, and unlabeled ethanol, glycerol,

ionone in *Yarrowia lipolytica* *Pathway(s)*: TCA, glycolysis, oxPPP, FA, AA, mevalonate pathway		^{13}C-mevalonate(M+2, M+4, M+6)	isoleucine, and valine were pulsed in later. Central carbon pathways were measured, with dilution of labeling and β-ionone titers indicating the importance of the secondary nutrients [45].
Increased usage of glutamine versus lactate toward gluconeogenesis after glucagon stimulation in C57bl/6-isolated liver cells and in mice *Pathway(s)*: Gluconeogenesis, TCA	[U-$^{13}C_5$]glutamine	^{13}C-glucose(M+3); ^{13}C-αKG(M+5); ^{13}C-glutamate(M+5); ^{13}C-malate(M+4); ^{13}C-aspartate(M+4)	Either labeled glutamine or lactate was fed alongside their unlabeled counterparts to track their contribution to gluconeogenesis. Labeled glutamine entered TCA cycle from GDH and was exported as malate, which proceeded to M+3 glucose via PEPCK. Labeled lactate was primarily introduced into TCA cycle by pyruvate carboxylation to OAA, which also exited as malate. This route would generate M+3 glucose. Via PDH, M+2 acetyl-CoA is generated and leads to M+2 glucose [47].
	[U-$^{13}C_3$]lactate	^{13}C-glucose(M+2, M+3); ^{13}C-αKG(M+2); ^{13}C-malate(M+3); ^{13}C-aspartate(M+3)	
Confirming the existence of monomethyl BCFA *de novo* synthesis from branched chain amino acids in 3T3-L1 preadipocytes and in mice. *Pathway(s)*: Glycolysis, oxPPP, TCA, FA, AA	[U-$^{13}C_6$]glucose	^{13}C-*iso*-C16:0(M+N, where N is even and 2≤N≤12); ^{13}C-palmitate(M+N, where N is even and 2≤N≤16); ^{13}C-acetyl-CoA(M+2)	Uniformly labeled glucose was used to detect monomethyl and saturated fatty acid synthesis. Labeled glutamine was fed to quantify its acetyl-CoA contribution. Labeled BCAAs were used to confirm their incorporation into BCFAs; only one molecule of each BCAA was incorporated, and variations in mass were due to addition of BCAA-derived acetyl-CoA. Heavy water was used to measure *de novo* synthesis in germ-free, pathogen-free mice, controlling for diet and microbiota. Deuterated glucose accounted for NADPH production and its labeling of fatty acids. [48]
	[U-$^{13}C_5$]glutamine	^{13}C-acetyl-CoA(M+2)	
	[U-$^{13}C_6$,$^{15}N_1$]leucine	^{13}C-*iso*-C17:0(M+5, M+7); ^{13}C-acetyl-CoA(M+2); ^{13}C-isovaleryl-CoA(M+5);	
	[U-$^{13}C_6$,$^{15}N_1$]isoleucine	^{13}C-*anteiso*-C17:0(M+2, M+5, M+7); ^{13}C-acetyl-CoA(M+2); ^{13}C-2-methylbutyryl-CoA(M+5);	
	[U-$^{13}C_5$,$^{15}N_1$]valine	^{13}C-*iso*-C16:0(M+4); ^{13}C-acetyl-CoA(M+2); ^{13}C-isobutyryl-CoA(M+4)	
	[1-$^{2}H_1$]water	^{2}H-*iso*-C16:0 ^{2}H-palmitate	
	[3-$^{2}H_1$]glucose	^{2}H-*iso*-C16:0(M+1 to M+4); ^{2}H-palmitate(M+1 to M+5); ^{2}H-NADPH(M+1)	
Direct conversion pathways of pyruvate to acetate in various mammalian cell lines *Pathway(s)*: Glycolysis, TCA, FA, AA, oxidation of	[U-$^{13}C_6$]glucose	^{13}C-pyruvate(M+3); ^{13}C-acetate(M+2); ^{13}C-lactate(M+3); ^{13}C-acetyl-CoA(M+2)	Uniformly labeled glucose was used to quantify its fate to pyruvate, acetate, or lactate. Glutamine was fed as a comparative carbon source to glucose. After considering other potential
	[U-$^{13}C_5$]glutamine	^{13}C-pyruvate(M+3); ^{13}C-acetate(M+2); ^{13}C-lactate(M+3);	

pyruvate to acetate		^{13}C-acetyl-CoA(M+2)	sources of acetate, labeled pyruvate was used to determine contribution of pyruvate dehydrogenase, while H_2O_2 and O_2 accounted for reactive oxygen species conversion of pyruvate, along with other by-products [30].
	[U-$^{13}C_3$]pyruvate	^{13}C-acetate(M+2); ^{13}C-acetyl-GSH(M+2)	
	[3,3,3-$^{2}H_3$]pyruvate	^{2}H-acetate(M+3)	
	$H_2{}^{18}O_2$ $^{18}O_2$	^{18}O-acetate(M+2); ^{18}O-kynurenine(M+2); ^{18}O-MetO(M+2)	
Extracellular uptake of protein and mTORC1's regulatory role in mouse pancreatic cancer cells, fibroblasts harboring oncogenic *K-Ras.* *Pathway(s)*: Macropinocytosis, amino acid uptake	[$^{13}C_6{}^{15}N_3$]histidine; [$^{13}C_6{}^{15}N_2$]lysine; [$^{13}C_9{}^{15}N_1$]phenylalanine; [$^{13}C_4{}^{15}N_1$]threonine; [$^{13}C_5{}^{15}N_1$]valine	^{13}C, ^{15}N-histidine(M+9); ^{13}C, ^{15}N-lysine(M+8); ^{13}C, ^{15}N-phenylalanine(M+10); ^{13}C, ^{15}N-threonine(M+5); ^{13}C, ^{15}N-valine(M+6)	Labeled amino acids were used to differentiate macropinocytosis and amino acid uptake to measure protein scavenging flux. Release of unlabeled amino acids was considered to be breakdown of unlabeled albumin fed to cells [49].
Contribution of circulating lactate to TCA cycle in mice. *Pathways(s)*: TCA	[U-$^{13}C_6$]glucose	^{13}C-succinate(M+2); ^{13}C-malate(M+2); ^{13}C-glucose(M+3, M+6); ^{13}C-glutamine(M+2); ^{13}C-lactate(M+3)	Contributions of circulating glucose, glutamine and lactate to the TCA cycle were determined. Succinate and malate were measured due to their high abundance in tissues. Labeling enrichment in serum glucose, glutamine, and lactate for each labelled nutrient was determined to classify it as "circulating" [7].
	[U-$^{13}C_5$]glutamine	^{13}C-succinate (M+4); ^{13}C-malate (M+4); ^{13}C-glucose(M+2); ^{13}C-glutamine(M+2, M+4, M+5); ^{13}C-lactate(M+2, M+3)	
	[U-$^{13}C_3$]lactate	^{13}C-succinate (M+2); ^{13}C-malate (M+2); ^{13}C-glucose(M+2, M+3); ^{13}C-glutamine(M+2, M+3); ^{13}C-lactate(M+2, M+3)	
Studies on thermodynamics of metabolism			
Thermodynamics of reactions in glycolysis, amino acids synthesis, and fatty acid synthesis through metabolic fluxes and concentrations in *E. coli*, *S. cerevisiae*, and mammalian cells. *Pathway(s)*: Glycolysis, PPP, TCA, one-carbon, FA, AA	[U-$^{13}C_6$]glucose in mammalian cells	^{13}C-6PG(M+5, M+6); ^{13}C-malate(M+2, M+3); ^{13}C-serine(M+1, M+2, M+3); ^{13}C-pyruvate(M+2, M+3)	*E. coli* was fed either 50% [U-$^{13}C_6$]glucose or [3-$^{13}C_1$]glucose, yeast was fed 50% [U-$^{13}C_6$]glucose, and iBMK cells were fed entirely [U-$^{13}C_6$]glucose or [U-$^{13}C_5$]glutamine in order to measure fluxes and subsequently, free energies of reactions, within major pathways. Labeled pyruvate was used to investigate the reversibility of pyruvate kinase [3]. All three organisms were fed [1,2-$^{13}C_2$]glucose to measure
	50% [U-$^{13}C_6$]glucose	^{13}C-6PG(M+3, M+5, M+6); ^{13}C-FBP(M+3, M+6); ^{13}C-malate(M+2, M+3); ^{13}C-serine(M+1, M+2, M+3); ^{13}C-pyruvate(M+2, M+3)	
	[U-$^{13}C_5$]glutamine	^{13}C-citrate(M+3, M+5); ^{13}C-αKG(M+2, M+4); ^{13}C-malate(M+0, M+1, M+4)	
	[U-$^{13}C_3$]pyruvate	^{13}C-3PG(M+3); ^{13}C-PEP(M+3); ^{13}C-malate(M+2, M+3); ^{13}C-aspartate(M+2, M+3)	

	[3-$^{13}C_1$]glucose	^{13}C-6PG(M+1, M+2); ^{13}C-FBP(M+1, M+2); ^{13}C-X5P(M+1, M+2); ^{13}C-αKG(M+1); ^{13}C-malate(M+1); ^{13}C-serine(M+0); ^{13}C-pyruvate(M+0, M+1)	oxPPP flux (M+1 lower glycolytic intermediates come from oxPPP) and resolve reversibility of TPI. Supplementary table 2 in this paper shows this information in detail.
	[1,2-$^{13}C_2$]glucose	^{13}C-FBP(M+2, M+4); ^{13}C-DHAP(M+2); ^{13}C-GAP(M+2); ^{13}C-R5P(M+1, M+2); ^{13}C-X5P(M+1, M+2, M+3, M+4);	
Thermodynamics of glycolysis in *E. coli*, mouse kidney epithelial cells, and *C. cellulolyticum*, through metabolic fluxes *Pathway(s)*: Glycolysis	[1,2-$^{13}C_2$]glucose	^{13}C-FBP(M+2, M+4); ^{13}C-DHAP(M+0, M+2); ^{13}C-3PG(M+0, M+2)	[1,2-$^{13}C_2$]glucose was fed to resolve reversibility of aldolase and TPI. [5-$^{2}H_1$]glucose was used to trace glycolysis, excluding TPI. The loss of ^{2}H via TPI and ENO and the reversibility of glycolysis generate a gradient of deuterium labeling [1].
	[5-$^{2}H_1$]glucose	^{2}H-G6P(M+0, M+1); ^{2}H-F6P(M+0, M+1); ^{2}H-FBP(M+0, M+1); ^{2}H-3PG(M+0, M+1); ^{2}H-23BPG(M+0, M+1); ^{2}H-PEP(M+0, M+1);	
EMP and ED glycolysis preference in *Z. mobilis* and other carbon cycling phenomenon *Pathway(s)*: Glycolysis, ED pathway, oxPPP, TCA	[1-$^{13}C_1$]glucose	^{13}C-G6P(M+1); ^{13}C-pyruvate(M+1); ^{13}C-6PG(M+1); ^{13}C-S7P(M+1, M+2); ^{13}C-malate(M+1)	^{2}H and ^{13}C glucose were used to determine flux and thermodynamics of ED pathway and EMP glycolysis. [1-$^{13}C_1$]glucose was used to trace glycolysis, oxPPP, and TCA cycle. [3-$^{13}C_1$]- and [6-$^{13}C_1$]-glucose were used to prove the exclusive use of ED pathway. [4-$^{13}C_1$]glucose was used to differentiate between the fates of pyruvate: anaplerosis versus pyruvate dehydrogenase. Deuterated glucose variants were used to differentiate between ED pathway and EMP glycolysis and measure reversibility of ED pathway steps [2].
	[3-$^{13}C_1$]glucose	^{13}C-KDPG(M+1); ^{13}C-pyruvate(M+1)	
	[4-$^{13}C_1$]glucose	^{13}C-pyruvate(M+1); ^{13}C-citrate(M+1, M+2); ^{13}C-CO_2(M+1); ^{13}C-succinate(M+1, M+2)	
	[6-$^{13}C_1$]glucose	^{13}C-3PG(M+1); ^{13}C-pyruvate(M+1); ^{13}C-KDPG(M+1, M+2); ^{13}C-citrate(M+1, M+2)	
	[1,2-$^{13}C_2$]glucose,	^{13}C-FBP(M+2, M+4); ^{13}C-DHAP(M+2); ^{13}C-GAP(M+2)	
	[2-$^{2}H_1$]glucose; [4-$^{2}H_1$]glucose; [5-$^{2}H_1$]glucose	^{2}H-G6P(M+1); ^{2}H-GAP(M+1); ^{2}H-KDPG(M+1)	
Reversibility of Pta-AckA pathway in *E. coli* for acetate uptake and impact of knockout of various enzymes involved in acetate metabolism *Pathway(s)*: glycolysis, Pta-AckA pathway	[U-$^{13}C_6$]glucose	^{13}C-acetate(M+2); ^{13}C-acetyl-CoA(M+2); ^{13}C-acetyl-phosphate(M+2);	Acetate turnover rates via Pta-AckA pathway and acetyl-CoA synthetase was examined with uniformly labeled glucose. Pta-AckA intermediates acetyl-CoA and acetyl-phosphate. Labeled acetate was co-fed with glucose to examine its co-utilization with glucose. [52]
	[U-$^{13}C_2$]acetate	^{13}C-acetate(M+2); ^{13}C-acetyl-CoA(M+2); ^{13}C-acetyl-phosphate(M+2); ^{13}C-PEP(M+2); ^{13}C-malate(M+2)	

2.9 Conclusions

One of the key objectives in biotechnology and medicine is robust control of metabolism as metabolic fluxes directly influence cellular functionality and phenotypes [53]. The use of ^{13}C, ^{2}H, ^{18}O, and ^{15}N tracers in cells and animals has improved flux quantitation, providing insights into the fate of nutrients, drugs, and chemical moieties as well as the contribution of individual pathways to metabolic functions. While fluxes are now routinely measured, metabolic control strategies are seldom successful.

The missing puzzle piece in the present control schemes is thermodynamics, which has been challenging to comprehend due to the need for complete knowledge of involved metabolite concentrations and standard free energies. We reviewed here recent progresses in thermodynamic quantitation. Comprehensive and precise quantitation of Gibbs free energy of reaction has emerged as the consequence of improved flux quantitation, which scrutinized individual reactions' forward and backward fluxes. Reaction ΔG reveals enzyme efficiency, pathway control capability, and adaptability in changing environments. The upshot of ΔG quantitation is that we have an improved understanding of the metabolic pathway design principles.

To successfully control pathway fluxes, we must first go back to the basics of how fluxes are determined: the fundamental thermodynamic and kinetic rate laws that incorporate the levels of metabolites (substrates, products, regulators) and proteins (metabolic enzymes). Therefore, comprehensive quantitation of multi-omics datasets, kinetic parameters, and thermodynamic driving forces would lead to robust metabolic control strategies to accelerate metabolic engineering and therapeutic development endeavors.

Chapter 3: Engineering *E. coli* gluconeogenesis and non-oxidative glycolysis to assimilate formate into acetate

3.1 Introduction

Carbon dioxide (CO_2) emissions have been the largest contributor to global greenhouse warming, comprising nearly 2/3 of the global heating contribution in 2018 after reaching a historic 33.1 Gt CO_2 [54]. CO_2 capture and storage (CCS) technologies have been increasingly implemented by various industries, with long-term storage typically involving injection into deep rock formations [54]. However, while effective at reducing emissions, CCS technologies are mostly unable to convert CO_2 into useful chemicals. There have been numerous electrochemical [55,56] and catalytic [57,58] techniques developed that can convert CO_2 to more valuable chemicals. Chemical-based production is typically faster and more specific than biological production, and avoids accumulation of cell waste and biomass; however biological production can produce more structurally-complex products via a lengthy pathway, typically has lower energy demands, and does not require highly purified starting materials [59,60]. Fixation by engineered biological organisms is a versatile option to reduce CO_2 emissions while recycling CO_2 into industrially-relevant products.

There are four major naturally-occurring autotrophic pathways that are typically recognized. Traditional CO_2 fixation in plants and cyanobacteria occurs through the Calvin-Benson-Bassham (CBB) cycle, which produces triose phosphate and other downstream sugars [61]. Methanogens and acetogens commonly utilize the Wood-Ljungdahl pathway to produce acetyl-CoA from CO_2 [62]. These pathways, along with the Arnon-Buchanan and 3-hydroxypropionate cycles, can convert CO_2 to various useful substrates, but require extensive pathway engineering and strain evolution to implement in heterotrophs [61]. Glycolysis, on the other hand, is highly conserved

across nature and produces both energy and biomass, both essential to cellular growth [63]. Consequently, gluconeogenesis shares many of the same enzymes as glycolysis besides the largely irreversible hexokinase, phosphofructokinase, and pyruvate kinase steps, and is an overall anabolic pathway. Thus, if gluconeogenesis could be modified to produce sugar phosphates from CO_2, this pathway would be possible in a wide range of heterotrophs and assimilate one-carbon compounds throughout central carbon metabolism.

E. coli is well-characterized and cannot naturally grow on only one-carbon compounds, and thus is an ideal organism to implement a one-carbon assimilation pathway [64]. There have been multiple pathways engineered in *E. coli* that consume one-carbon compounds as carbon sources. Gleizer *et al*. introduced the CBB cycle into *E. coli*, and achieved entirely autotrophic growth after around 340 days of evolution [64]. Bang and Lee reconstructed the tetrahydrofolate (THF) cycle in *E. coli*, which was able to produce serine and pyruvate while fed only formate and CO_2 [65]. Kim *et al*. also introduced the THF cycle into *E. coli*, although they produced ethanol and fed glycine instead of CO_2 alongside formate [66]. Kim *et al*. achieved *E. coli* growth on methanol and formate in sugar-free media, using the reductive glycine pathway to assimilate formate and methanol dehydrogenase overexpression to convert methanol to formate [67]. The aforementioned pathways each offer unique advantages towards the consumption of one-carbon compounds, but typically required the extensive introduction of heterologous genes, oxygen to function, and lose carbons as CO_2. Thus, we sought to engineer a carbon-neutral pathway that relies largely on native genes and functions with and without oxygen. The latter is relevant to industrial production, as processes will often switch between aerobic and anaerobic conditions to prioritize either growth or production [68,69]. The pathway would utilize the one-carbon compound formate to synthesize a value-added product.

Formate is one of the major products of mixed-acid fermentation and can be produced from CO_2 via electrocatalysis [70]. In *E. coli*, formate is converted to CO_2 via one of three formate dehydrogenases (FDHs), which have been shown to be reversible *in vivo* [71]. Formate and acetyl-CoA can form pyruvate via the reversible anaerobic enzyme pyruvate formate lyase (*pfl*) [72]. With pyruvate being the primary starting substrate in gluconeogenesis, the collective reverse reactions of FDH and *pfl* could conceivably recycle CO_2 into central carbon metabolism (**Figure 4**).

Upon producing hexose phosphates from gluconeogenesis, it is necessary to redirect these metabolites through another pathway to produce energy and biomass. Allowing these metabolites to proceed back down native glycolysis inefficiently utilizes enzymes (as gluconeogenesis and glycolysis share some machinery) and carbon (since each produced pyruvate loses a carbon to CO_2 upon oxidation to acetyl-CoA). Bogorad *et al*. developed a non-oxidative glycolytic pathway in *E. coli*, capable of producing acetate from glucose with minimal CO_2 production [5]. This pathway (see **Figure 4**) utilizes enzymes native to *E. coli* aside from phosphoketolase (*xpk*), which converts the pentose phosphate intermediate xylulose-5-phosphate (Xu5P) to glyceraldehyde-3-phosphate (GAP) and acetyl-phosphate (AcP). *Xpk* is essential to non-oxidative glycolysis (NOG), as the produced AcPs can become acetate or acetyl-CoA via the *pta-acka* pathway. NOG was further improved by Lin *et al*. to support growth on glucose-only minimal media [6]. An optimized strain, NOG26, was able to grow and prioritize acetate production, with a doubling time of 3.6 hours [6]. The strain was first grown under aerobic conditions, and upon reaching the desired cellular density, was switched to anaerobic conditions, where acetate production was maximized. NOG26 served as the starting point for this study.

The proposed pathway employs NOG to convert six-carbon sugars produced via gluconeogenesis to acetate, effectively converting formate to acetate (**Figure 4**). Formate enters gluconeogenesis as pyruvate via *pfl*. Glucose is co-fed with formate to enable growth and produce ATP necessary for gluconeogenesis. This pathway, the Gluconeogenesis-Non-Oxidative Glycolysis (or GNG-NOG) cycle, can end as acetate, or can recycle acetyl-CoA to further contribute to *pfl* flux and gluconeogenesis.

There were several major challenges to engineering the GNG-NOG cycle. First, both gluconeogenesis and the *pfl* reaction are known to be thermodynamically unfavorable under the majority of conditions [3,72]. To address concerns about pathway viability, we conducted Flux Balance Analysis (FBA) to confirm that growth while having non-zero gluconeogenesis and reverse *pfl* flux was possible. We also conducted an overall metabolite balance to confirm that the cycle was overall forward-driven, and determined an optimal range for the glucose:formate feeding ratio. Second, buildup of formate is known to be toxic, causing oxidative stress and intracellular acidification [66,73]. To test formate toxicity, we examined growth at different formate concentrations, and observed negligible growth inhibition at 10 mM of formate. Finally, we were concerned that other reactions would consume formate, detracting from the GNG-NOG cycle. One major reaction is catalyzed by the aforementioned FDHs; we considered knocking out the major fermentative FDH *fdhF*, but we identified a genomic truncation in NOG26, the precursor to our engineered strains, which had already removed *fdhF*. Other potentially interfering pathways would be purine and THF biosynthesis, which incorporates formate via the phosphoribosylglycinamide formyltransferase 1 and 2 (*purT* and *purU*) reactions. However, significant incorporation of fed formate into purines has not yet been reported in *E. coli*, although Sah *et al*. did identify that purine auxotrophy in an *E. coli* strain with a re-constructed

THF cycle could be remediated by formate addition [74]. Surprisingly, we observed extensive formate incorporation throughout purines and one-carbon metabolites in our strains.

Starting from NOG26, we engineered six strains with restored gluconeogenesis and other genetic modifications to increase formate uptake and gluconeogenesis flux. We then conducted time course experiments to measure glucose and formate uptake along with acetate production. In addition, we traced the fate of fed formate using [U-^{13}C]formate. Finally, we estimated the intracellular metabolite concentrations of a suite of relevant metabolites, using the prior study developed by Park *et al* [3]. We have identified a strain, JXG2, which incorporated formate into both gluconeogenesis and NOG while producing primarily acetate. While we believe JXG2 was able to turn the GNG-NOG cycle, less than 1% of carbons in the pathway originated from formate, with purine, THF, and CoA synthesis acting as large formate sinks. As such, further optimization is needed to improve GNG-NOG cycle viability.

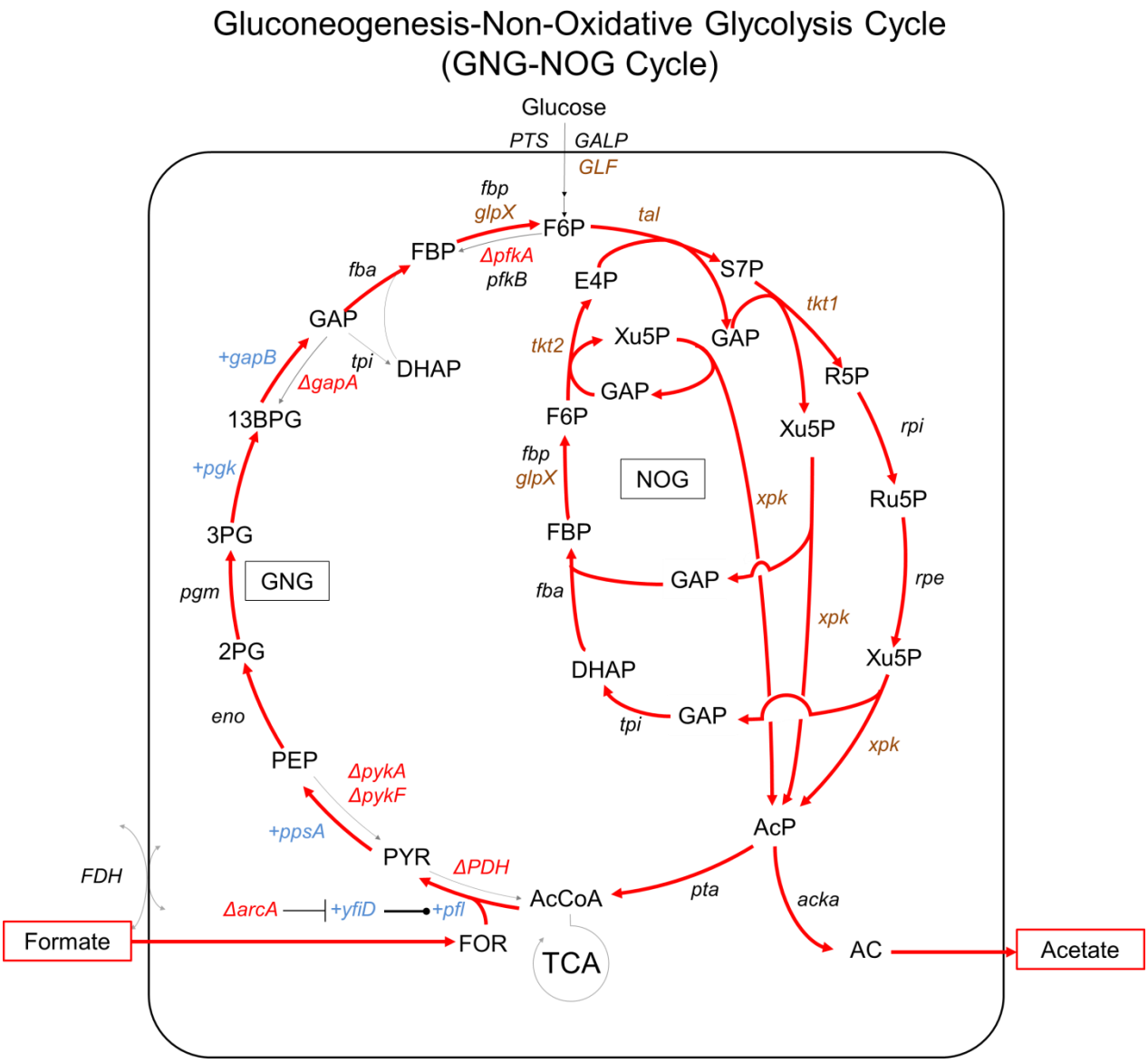


Figure 4 The Gluconeogenesis-Non-Oxidative Glycolysis (GNG-NOG) cycle

A diagram of the Gluconeogenesis-Non-Oxidative Glycolysis (GNG-NOG) cycle. Starting from formate, the pathway follows red arrows, ending either as exported acetate or recycled into acetyl-CoA. The major pathways are gluconeogenesis (GNG), non-oxidative glycolysis (NOG), with the tricarboxylic acid cycle (TCA) branching off acetyl-CoA. Metabolites are in larger black font between arrows, enzymes are in small, italicized fonts. Black enzymes are native to *E. coli*, red enzymes are enzymes knocked out during this study or the original Lin *et al*. study, brown enzymes are native and overexpressed by the original Lin *et al*. study, and blue enzymes are overexpressed via plasmids in this study.

3.2 Results

Feasibility of formate to acetate conversion confirmed by stoichiometric and thermodynamic analysis

We first wished to identify the net glucose requirements in order to facilitate formate uptake up gluconeogenesis, an energetically-expensive pathway. A net metabolite balance (**Table 2**) showed a cost of 32 ATP to convert six formate into one net glucose (**B5**). This ATP cost includes other cofactors (*i.e.* NADH, NADPH, and $FADH_2$) converted to ATP via transhydrogenase and oxidative phosphorylation. ATP can be generated via the *pta-acka* pathway consuming acetyl phosphate (AcP) from NOG (**B3**) or via the TCA cycle (**B4**). If only *pta-acka* is utilized, a maximum ratio of 2.5:1 of glucose:formate generates the necessary ATP while producing acetate (**B6**). This represents the maximum ratio of glucose to be fed alongside formate, as it assumes all ATP generation via NOG, the less-efficient pathway. If all ATP is produced by the TCA cycle, ATP generation is significantly more efficient, generating the required ATP with a glucose:formate ratio of 0.017:1, but produces carbon dioxide instead, effectively wasting formate (**B7**). Thus, a mixture of these two pathways should be utilized, with the TCA cycle supporting growth while some fed and generated glucose can be diverted towards acetate. We chose to examine the condition in which exactly 1 mole of formate results in 1 mole of CO_2, as this represents the minimum glucose:formate ratio that would generate the necessary ATP while losing less than a full carbon to CO_2. This ratio was found to be 0.25:1 glucose:formate (**B8**). Thus, a molar glucose:formate feeding ratio between 0.25:1 and 2.5:1 should be utilized. Using the eQuilibrator thermodynamic calculator and known intracellular metabolite concentrations, we also estimated the Gibbs free energies of reaction for all relevant balances and reactions, with overall balances being heavily forward-driven (**Table 2**). Thus,

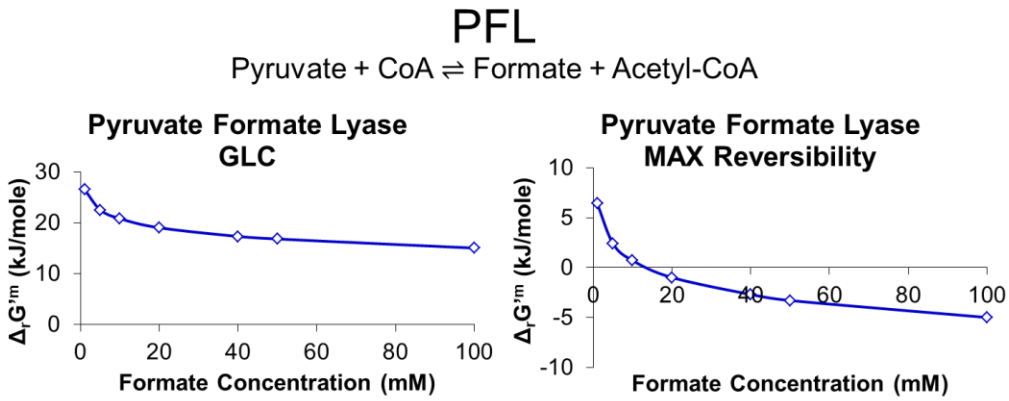
while gluconeogenesis and production of ATP are energetically-expensive, the overall pathway balances indicate the pathway is favorable.

Table 2 Net metabolite balances of the GNG-NOG cycle

Name	Net Input	Net Output	$\Delta_r G'^{o}$ (kJ/mol)
Pathways			
NOG	Glc + ATP + 2Pi	3AcP + ADP	-217.9
GNG	2AcP + 2For + 6ATP + 2NADH	F6P + 7Pi + 6ADP + $2NAD^+$	9.7
TCA	AcP + $NADP^+$ + $2NAD^+$ + ADP + FAD	$2CO_2$ + NADPH + 2NADH + ATP + $FADH_2$	17.3
Reactions			
Hexokinase+*pgi* **[R1]**	Glc + ATP	F6P + ADP	-14.7
ackA **[R2]**	AcP + ADP	Ac + ATP	-13.4
Transhydrogenase **[R3]**	NADPH + NAD^+	NADH + $NADP^+$	-1.0
Oxidative Phosphorylation **[R4]**	1NADH + 2.5ADP + 2.5Pi; $1FADH_2$ + 1.5ADP + 1.5Pi	2.5ATP + $1NAD^+$; 1.5ATP + FAD	44.6; 18.1
Cofactor Substitution Balances			
TCA+R3+R4 **[B1]**	AcP + 10ADP + 9Pi	$2CO_2$ + 10ATP	151.2
GNG-R1-R4 **[B2]**	2AcP + 2For + 10ATP	Glc + 10ADP + 12Pi	-133.2
ATP Generation Balances			
NOG+3*R2 **[B3]**	Glc + 2Pi + 2ADP	3Ac + 2ATP	-258.1
NOG+3*B1 **[B4]**	Glc + 29ADP + 29Pi	$6CO_2$ + 29ATP	235.6
Overall Balances			
NOG+1.5*B2 **[B5]**	3For + 16ATP	0.5Glc + 16ADP + 16Pi	-417.8
B5+8*B3 **[B6]**	7.5Glc + 3For	24Ac	-2482.7
B5+0.55*B4 **[B7]**	0.052Glc + 3For	$3.31CO_2$	-287.7
B5+0.75*B3+0.5*B4 **[B8]**	0.75Glc + 3For	2.25Ac + $3CO_2$	-424.5

We were also interested in estimating the reversibility of *pfl* though eQuilibrator's custom concentrations feature. We found the reverse reaction to be unfavorable across the majority of concentrations, estimated using the range of intracellular concentrations determined by Bennett *et al*. and pyruvate concentrations in *pykF* mutant *E. coli* by Siddiquee *et al*. (**Figure 5)** [75,76]. However, with intracellular concentrations set to the extremes that favored the reverse reaction while still within the 95% confidence range, the *pfl* reaction becomes reversible at 10 mM of formate. This reversibility is supported by estimates generated by Zelcbuch *et al*., who instead assumed a broad range of intracellular concentrations and estimated a *pfl* lower $\Delta_r G$ bound of -70

kJ/mol. From these estimates and the study by Zelcbuch *et al*., we believe that *pfl* is capable of significant flux in the reverse direction [72].



Metabolite	Intracellular Concentration (mM)	Intracellular Concentration (mM)
Pyruvate	3.66	0.0883
CoA	1.37	0.05
Acetyl-CoA	0.606	1.74

Figure 5 Estimates of Gibbs Free Energy of Reaction ($\Delta_r G'^m$) of pyruvate formate lyase

Using the eQuilibrator online thermodynamic calculator and absolute intracellular metabolite concentrations by Bennett *et al.* and Siddiquee *et al*., the Gibbs energy of reaction of pyruvate formate lyase (*pfl*) was calculated [75,76]. These were done for exact published values when fed glucose (Left) and at the maximum value of the acetyl-CoA concentration range estimate and the minimum values of the CoA and pyruvate concentration range estimate (Right).

After confirming that glucose was required to fund the GNG-NOG cycle and that *pfl* can be reversible when fed formate, we utilized the COBRA toolbox to conduct FBA and confirm that the pathway was possible. From FBA, we saw that NOG26, in order to maximize growth, prioritizes NOG and TCA cycle under aerobic conditions (**Figure 6a**), while under anaerobic conditions (**Figure 6b**), NOG and acetate production are maximized. This is consistent with the findings of Lin *et al*., where NOG26 was observed to switch from growth to production mode upon loss of oxygen [6]. Upon introduction of gluconeogenic *gapB* from *B. subtilis* and *pgk*, at microaerobic conditions, we saw a similar flux profile to NOG26's at anaerobic conditions, but no gluconeogenic flux.

We then examined knockout of *pykA* and having *pfl* be functional under microaerobic conditions, which has been shown to be possible in engineered *E. coli* [77,78]. This strain was named JXG4, and was capable of the GNG-NOG cycle. The suggested changes showed the restoration of gluconeogenesis when the objective function was acetate export (**Figure 6c**). We did see an increase in acetate production under microaerobic conditions when fed formate compared to anaerobic NOG26 (**Table 3**). Strikingly, we also saw large fluxes in *ppsA* and PDH, two reactions that consume pyruvate. While some pyruvate was produced by *pfl*, a large portion was produced by reactions not shown in the figure (via phosphoenolpyruvate carboxykinase or the phosphoenolpyruvate-carbohydrate phosphotransferase system). However, in our model, the *pfl* flux was limited by the assumed formate uptake rate (10 mmol gDW^{-1} hr^{-1}), which would likely be exceeded by engineered or evolved strains; as such, this demand for pyruvate may encourage reverse *pfl* flux.

While gluconeogenesis was activated in JXG4, the flux was still low relative to the *pfl* flux (**Table 3**). We examined a strain with PDH knockout, JXG6, as removal of PDH should eliminate a reaction competing with gluconeogenesis. However, we found that gluconeogenic flux decreased from JXG4, likely due to the increased growth stress from the loss of a relevant NADH source (**Table 3, Figure 6d**). Indeed, we saw a lower formate uptake rates, *pfl* fluxes, and maximum growth rates compared to JXG4 (**Table 3, Figure 6d**). While these simulations showed low flux through gluconeogenesis, they indicated that the GNG-NOG cycle flux was possible in both JXG4 and JXG6, which could be improved with targeted strain engineering.

One weakness of the COBRA toolbox is its unrealistic view on relative contributions of alternative pathways, with less-efficient pathways being entirely ignored. For example, the malic enzymes could be a potential source of pyruvate, but had zero flux in every GNG-NOG

simulation. While it was promising that the toolbox determined that the desired pathway was theoretically possible, quantitative estimates of reaction fluxes must be considered skeptically. Thus, different strains were engineered to examine their respective abilities to turn the GNG-NOG cycle.

Table 3 Metabolic flux estimates for GNG-NOG cycle via Flux Balance Analysis

Strain/Conditions; Obj. Function	Max Growth Rate (hr^{-1})	Acetate Export ($mmol*gCDW^{-1}* hr^{-1}$)	PFL Flux ($mmol*gCDW^{-1}* hr^{-1}$)	Gluconeogenic GAPDH Flux* ($mmol*gCDW^{-1}*hr^{-1}$)
NOG26 Aerobic G+F; Biomass	1.387	0	10	0
NOG26 Aerobic G; Biomass	1.334	0	0	0
NOG26 Microaerobic G+F; Biomass	0.730	17.036	10	0
NOG26 Microaerobic G; Biomass	0.730	17.036	0	0
NOG26 Anaerobic G+F; Biomass	0.162	53.843	2.136	0
NOG26 Anaerobic G; Biomass	0.162	41.965	0	0
JXG4 Aerobic G+F; Biomass	1.387	0	10	0
JXG4 Aerobic G; Biomass	1.320	0	0	0
JXG4 Microaerobic G+F; Biomass	0.688	45.412	10	0
JXG4 Microaerobic G; Biomass	0.673	43.136	0	0
JXG4 Anaerobic G+F; Biomass	0.100	53.843	2.136	0
JXG4 Anaerobic G; Biomass	0.100	53.314	0.533	0
JXG6 Aerobic G+F; Biomass	1.387	0	10	0
JXG6 Aerobic G; Biomass	1.320	0	0	0
JXG6 Microaerobic G+F; Biomass	0.347	29.299	5.303	0
JXG6 Microaerobic G; Biomass	0.336	28.282	0	0
JXG6 Anaerobic G+F; Biomass**	0.102	56.912	1.487	0
JXG6 Anaerobic G; Biomass**	0	0	0	0
NOG26 Microaerobic G+F; Max Ac	0.100	53.847	2.131	0
NOG26 Microaerobic G; Max Ac	0.100	53.314	0.533	0
JXG4 Microaerobic G+F; Max Ac	0.100	54.862	10	2.029
JXG4 Microaerobic G; Max Ac	0.100	53.076	0	0
JXG6 Microaerobic G+F;Max Ac	0.100	53.453	1.124	0.352
JXG6 Microaerobic G; Max Ac	0.100	53.076	0	0

*Gluconeogenic GAPDH flux relative to formate uptake rate of -10 mmol gDW^{-1} hr^{-1} when maximizing acetate export

**Rate of glucose uptake was raised to -20 mmol gDW^{-1} hr^{-1}

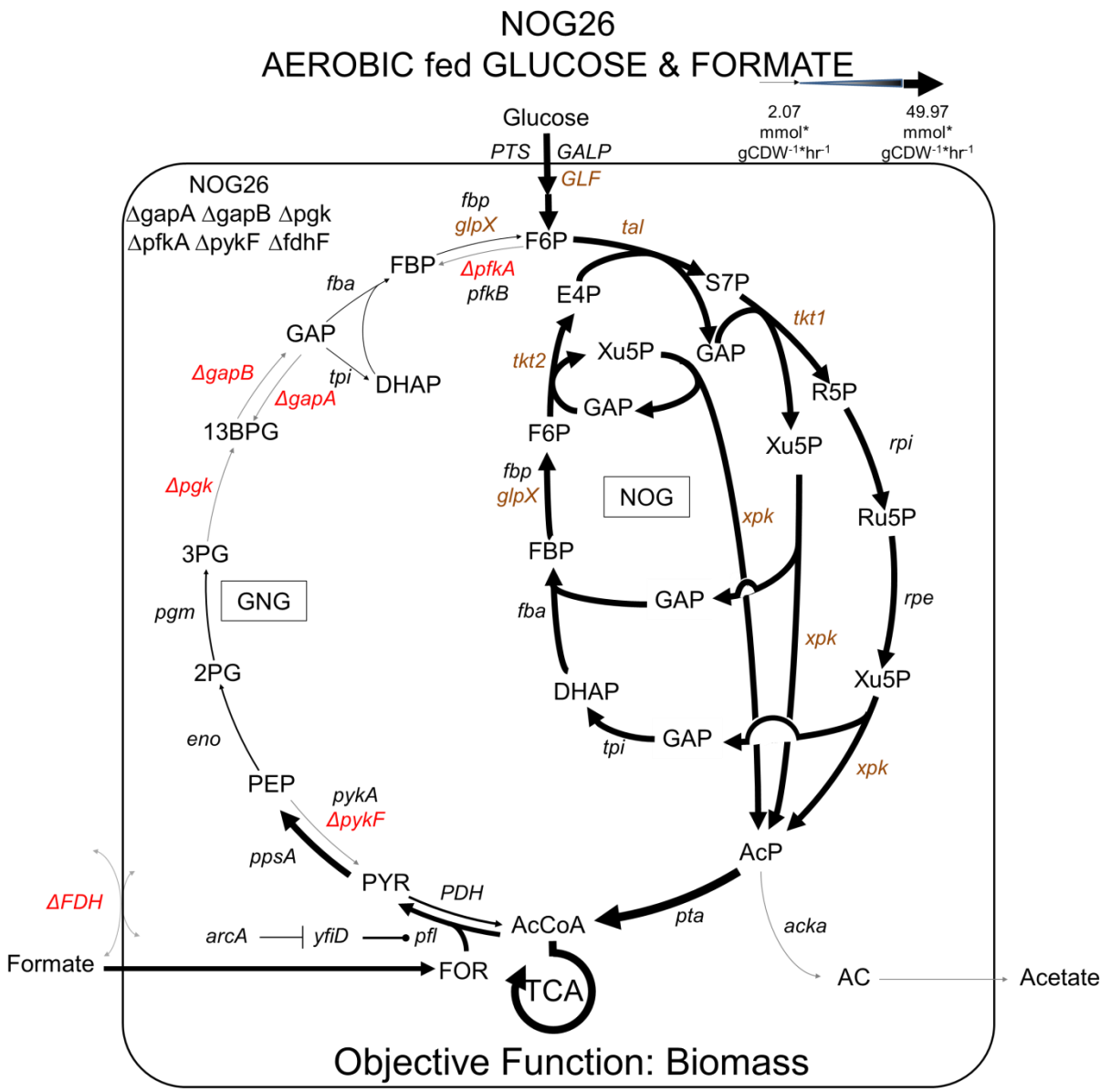


Figure 6a Flux balance analysis on NOG26 under aerobic conditions

Using the COBRA toolbox, NOG26 fluxes were simulated and quantitative flux estimates were calculated. The objective function was Biomass, indicating fluxes while maximizing growth. All arrow widths are relative to the maximum reaction flux value, which was assigned an arrow thickness of 10 pts; fluxes corresponding to maximum and minimum arrow thicknesses are displayed. Each simulated flux value was adjusted on a base 2 logarithmic scale and then the resulting value was compared to the maximum flux to generate an arrow thickness. Under aerobic conditions, NOG26 prioritizes non-oxidative glycolysis (NOG) and TCA cycle, with little flux through acetate export. All abbreviations in Figures 6a are shared with Figure 3.

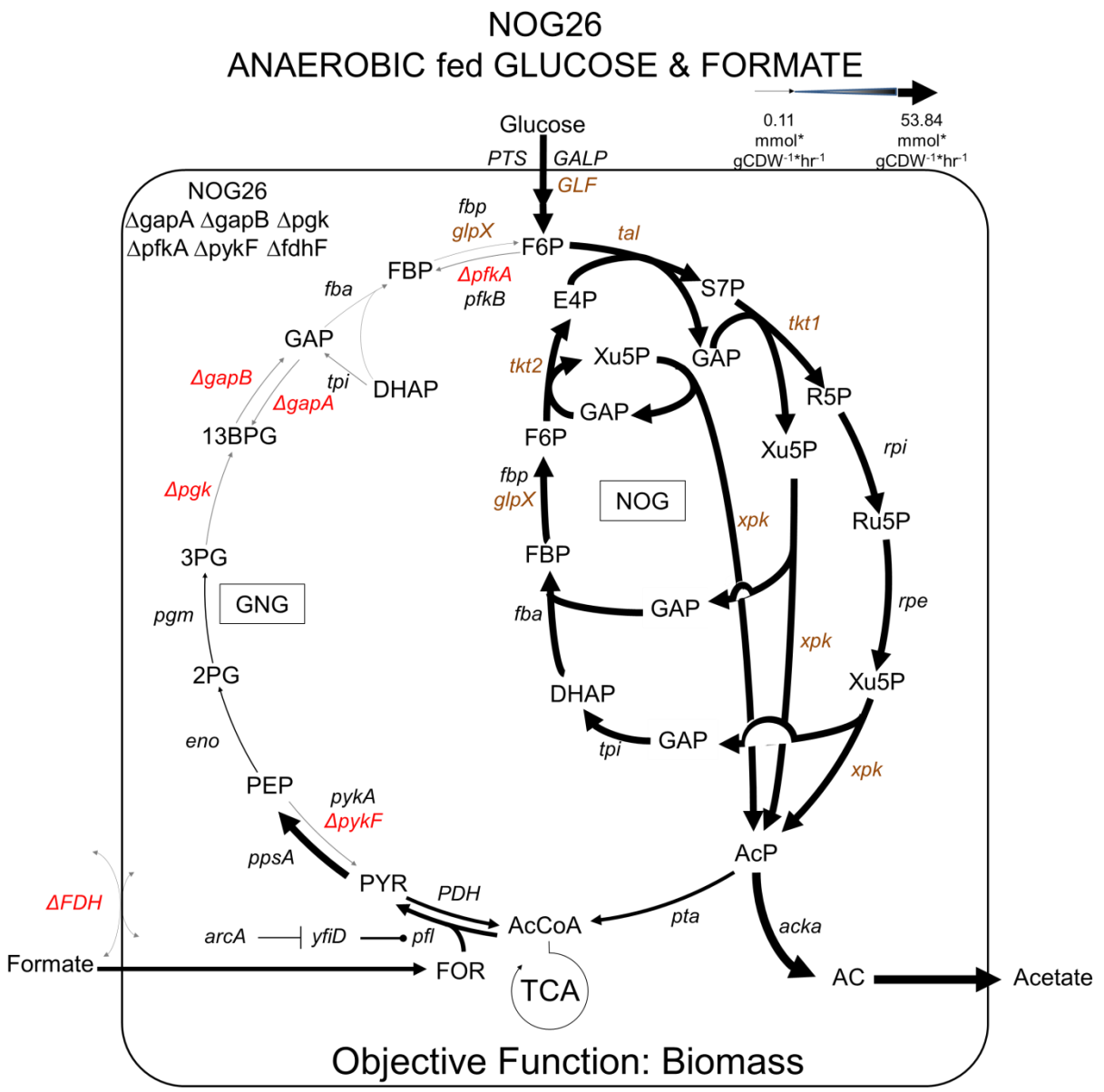


Figure 6b Flux balance analysis on NOG26 under anaerobic conditions

Using the COBRA toolbox, NOG26 fluxes were simulated and quantitative flux estimates were calculated. The objective function was Biomass, indicating fluxes while maximizing growth. All arrow widths are relative to the maximum reaction flux value, which was assigned an arrow thickness of 10 pts; fluxes corresponding to maximum and minimum arrow thicknesses are displayed. Each simulated flux value was adjusted on a base 2 logarithmic scale and then the resulting value was compared to the maximum flux to generate an arrow thickness. Under anaerobic conditions (lower), NOG26 relies on NOG and subsequently acetate export, as the *acka* reaction produces ATP. All abbreviations in Figures 6b are shared with Figure 3.

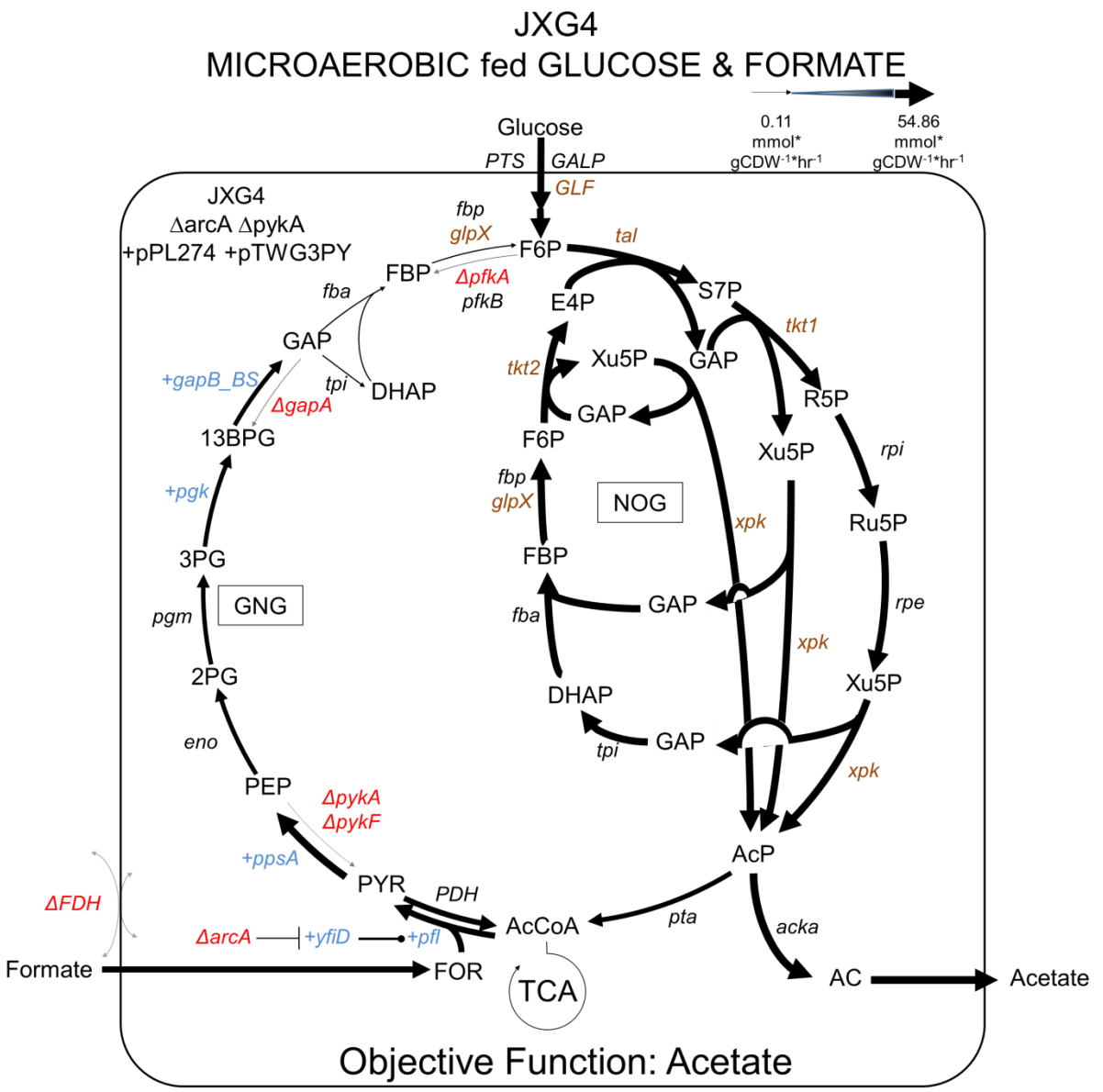


Figure 6c Flux balance analysis on JXG4 under microaerobic conditions

Flux estimates on *E. coli* with *gapB*_BS, *pgk*, *ppsA*, *pfl*, and *yfiD* upregulated, along with *pykA* knockout (JXG4). The objective function in this case was acetate, which is similar to the flux diagram under Biomass maximization, with the exception of gluconeogenesis flux. Under microaerobic conditions, with pyruvate kinase knockout, we see acetate production being maximized with significant NOG flux and non-zero GNG flux. All arrow widths are relative to the maximum reaction flux value, which was assigned an arrow thickness of 10 pts; fluxes corresponding to maximum and minimum arrow thicknesses are displayed. Each simulated flux value was adjusted on a base 2 logarithmic scale and then the resulting value was compared to the maximum flux to generate an arrow thickness. All abbreviations are shared with Figure 3.

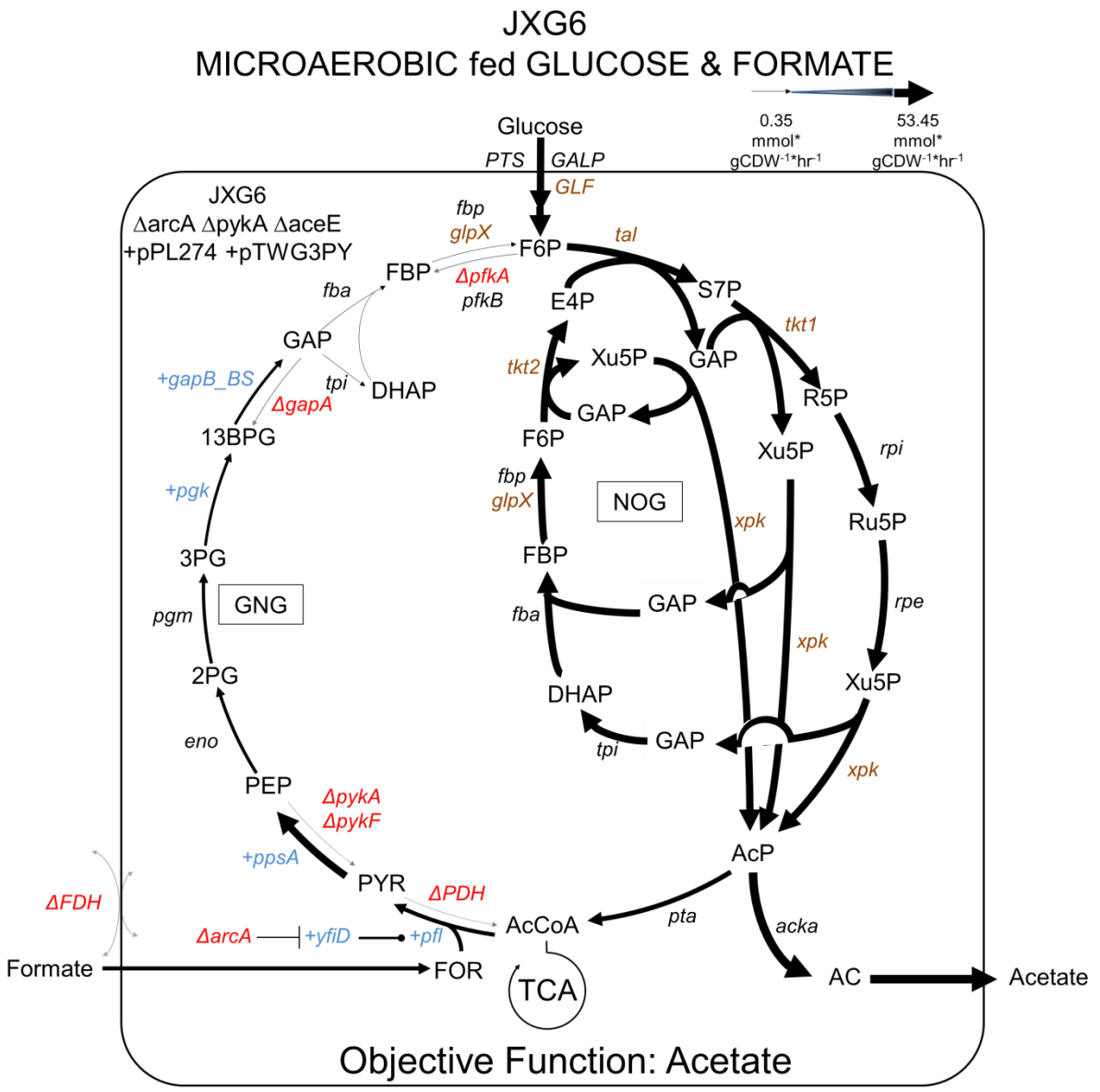


Figure 6d Flux balance analysis on JXG6 under microaerobic conditions

Flux estimates on *E. coli* with *gapB*_BS, *pgk*, *ppsA*, *pfl*, and *yfiD* upregulated, with *pykA* and PDH knockout (JXG6). The objective function was acetate, which is similar to the flux diagram under Biomass maximization, with the exception of gluconeogenesis flux. Under microaerobic conditions, with pyruvate kinase and PDH knockout, we see acetate production being maximized with significant NOG flux and non-zero GNG flux. There is a decrease in GNG flux compared to JXG4, still with significant NOG flux. All arrow widths are relative to the maximum reaction flux value, which was assigned an arrow thickness of 10 pts; fluxes corresponding to maximum and minimum arrow thicknesses are displayed. Each simulated flux value was adjusted on a base 2 logarithmic scale and then the resulting value was compared to the maximum flux to generate an arrow thickness. All abbreviations in Figures 6d are shared with Figure 3.

Design and engineering of six strains capable of GNG-NOG cycle

To engineer strains capable of the GNG-NOG cycle, we first needed to add genes essential to the pathway via an inducible genetic cassette on an existing plasmid in NOG26. To restore gluconeogenesis in NOG26, GAPDH and *pgk* needed to be re-added. However, if the native *E. coli* enzymes were used, EMP glycolysis could be restored, eliminating the need for NOG. In other bacterial strains like *Bacillus subtilis* and *Haloferax volcanii*, distinct GAPDH isoenzymes have been reported, with one variant functional in glycolysis and the other in gluconeogenesis [79,80]. *GapB* from *B. subtilis subsp. 168* (*gapB* _BS) was chosen for this study, as it has been regarded as irreversible in the gluconeogenic direction [81]. It is also strongly expressed under gluconeogenic conditions, and is proteolyzed upon a shift to glucose media [79]. There are no similarly irreversible *pgk* isoenzymes, so the native version was used. The plasmid generated after addition of *gapB*_BS and *pgk* to the multiple cloning site (MCS) of pTW371 was named pTWGP. NOG26 with pTW371 replaced by pTWGP was named JXG1.

Plasmid pTWGP contains the minimum genes necessary for the desired pathway. We then built upon pTWGP and added other genes that may increase GNG-NOG flux. Since *pfl* is the key enzyme facilitating formate's entry into gluconeogenesis, its reverse flux is likely be a major bottleneck [82]. Overexpression of a largely reversible enzyme will bring it closer to equilibrium [83], and *pfl*'s reversibility has been demonstrated [72,84,85]; thus, *pfl* was the first of the non-essential genes considered for overexpression.

Via FBA, we identified phosphoenolpyruvate synthase (*ppsA*) as a key enzyme for the pathway, as all successful scenarios showed significant flux through this enzyme. *PpsA* is one of several gluconeogenic routes pyruvate can take [86]. Other major pathways, such as phosphoenolpyruvate carboxykinase (PEPCK) or malate dehydrogenase, lose a carbon as CO_2

and thus would not be minimizing carbon loss; FBA concurred that *ppsA* was the more efficient path for formate utilization [87]. Thus, *ppsA* became the second of the non-essential genes targeted for inducible overexpression.

It is also desirable that *pfl* be functional under microaerobic conditions, as acetate production by NOG can be maintained while having faster growth compared to purely anaerobic conditions [88,89]. *Pfl* is cleaved under aerobic conditions, as it contains an oxygen-sensitive glycyl radical; the autonomous radical glycyl cofactor, *yfiD*, has been shown to possess considerable homology to the cleaved *pfl* peptide, and studies have found that it can reactivate *pfl* even after oxygen cleavage under microaerobic conditions [77,78]. Thus, *yfiD* was the last of the non-essential genes added to pTWGP. The final constructed plasmid with *pfl*, *ppsA*, and *yfiD* added was named pTWG3PY. NOG26 with pTW371 replaced by pTWG3PY was named JXG2. Confirmation of plasmid construction can be found in Appendix C1.

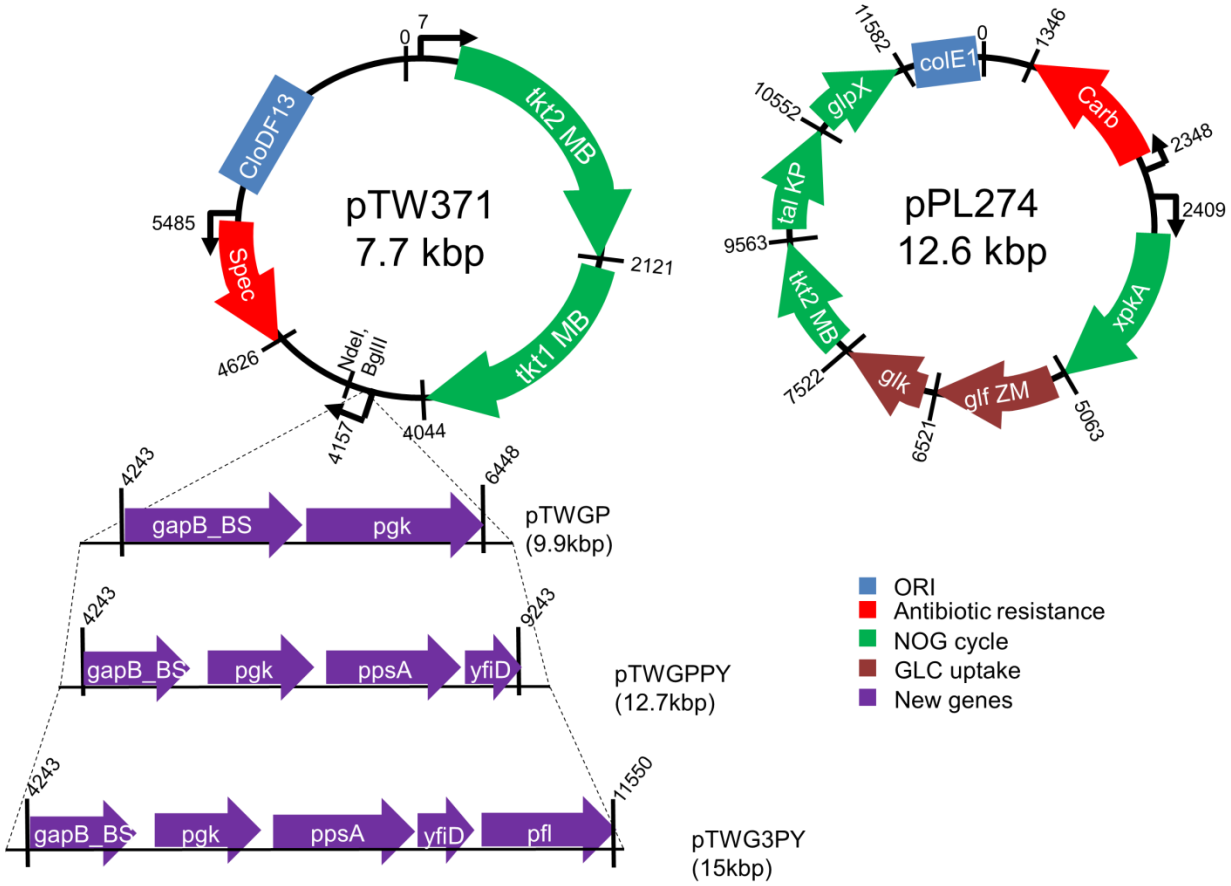


Figure 7 Plasmid maps of pPL274, pTW371, and newly engineered plasmids

The two plasmids that were transformed into NOG26 by Lin *et al*., pPL274 and pTW371, are shown above. pPL274 (right) contains genes found in NOG (green), along with genes that help facilitate glucose uptake (brown), and a carbenicillin (Carb) resistance selection marker (red). The pTW371 plasmid (left) contained two transketolase genes for NOG (green), a spectinomycin (Spec) resistance selection marker (red), as well as a multiple cloning site (MCS) for future cloning. At this site, we first added glyceraldehyde-3-phosphate dehydrogenase B (*gapB*) from *B. subtilis* and native phosphoglycerate kinase (*pgk*) (named pTWGP). We then created an intermediary vector, named pTWGPPY, adding in phosphoenolpyruvate synthetase (*ppsA*) and the autonomous radical glycyl cofactor (*yfiD*). Finally, we added an additional copy of pyruvate formate lyase (*pfl*) to create the pTWG3PY vector. NdeI and BglII are two restriction cut sites located within the MCS used during assembly. Numbering at the ends of genes indicate the position of the last base for each gene. Numbering at promoters indicate the position of the first base of the promoter. ORI denotes origin of replication; NOG, non-oxidative glycolysis; GLC, glucose; *tkt1/2 MB*, transketolase 1/2 from *Methylomicrobium buryatense* 5GB1; *xpkA*, phosphoketolase from *Bifidobacterium adolescents*; *glf ZM*; glucose facilitated diffusion protein from *Zymomonas mobilis*; *glk*, glucose kinase; *tal KP*, transaldolase from *Klebsiella pneumoniae*; and *glpX*, fructose-1,6-bisphosphatase 1.

The second stage of genetic manipulations was to perform CRISPR genetic knockouts of potentially interfering reactions. The first knockout target was *fdhF*, the primary formate dehydrogenase in *E. coli* during anaerobic fermentation [90]. *FdhF* converts formate to carbon dioxide, which would divert a portion of the fed formate away from the desired pathway. While attempting to knockout this gene, we were unable to amplify it from the NOG26 genome, while being able to easily PCR the gene in its wild-type precursor JCL16 (**Figure 8**). Upon further examination, we identified a known genomic truncation in a previous version of NOG26 about 1 kbp away from the *fdhF* gene [6]; this recorded mutation likely extended into *fdhF*, leaving NOG26 without a functional copy of the gene and no need to perform the knockout. There are two other FDH variants, (*fdhN* and *fdhO*), and we were able to confirm *fdhN* was intact in NOG26 (**Figure 8**) [90]. However, *fdhO* is expressed at relatively low levels (around 10% of FDH activity), while *fdhN* is not induced by formate, but rather nitrate, which was not present in the microaerobic media used in this study [90]. Thus, FDH should be a small contributor to formate consumption.

The second knockout target was *arcA*, a transcriptional regulator of many genes involved in aerobic respiration [91]. Prior studies on the *pfl-yfiD* interaction in *E.coli* found that expression of *yfiD* increased *pfl* flux in an *arcA* mutant [92]. While *arcA* is traditionally thought of as a positive regulator of anaerobic genes, its inhibition has been shown to significantly increase *pfl* flux and enhance *yfiD* expression [78,93]. This seems to be due to the fact that the *pfl* operon is regulated by two systems: the Arc and Fnr systems [93]. The Fnr system is known to be the stronger activator of *pfl*, and upon *arcA* mutation under microaerobic conditions, elevated levels of Fnr protein were observed, indirectly increasing *pfl* flux [93]. With *pfl* being the critical entry point for formate into the pathway, an *arcA* knockout was conducted in several strains. Several

studies have noted a mild decrease in *pfl* expression levels in Δ*arcA* mutants [94,95], so *arcA* knockout in combination with pTWG3PY (which overexpresses an additional copy of *pfl*) would be interesting to examine. It should also be noted that for microaerobic growth, *arcA* deletion resulted in drastic flux changes in *E. coli*, which may impact multiple genes within the GNG-NOG cycle, not just *pfl* [96].

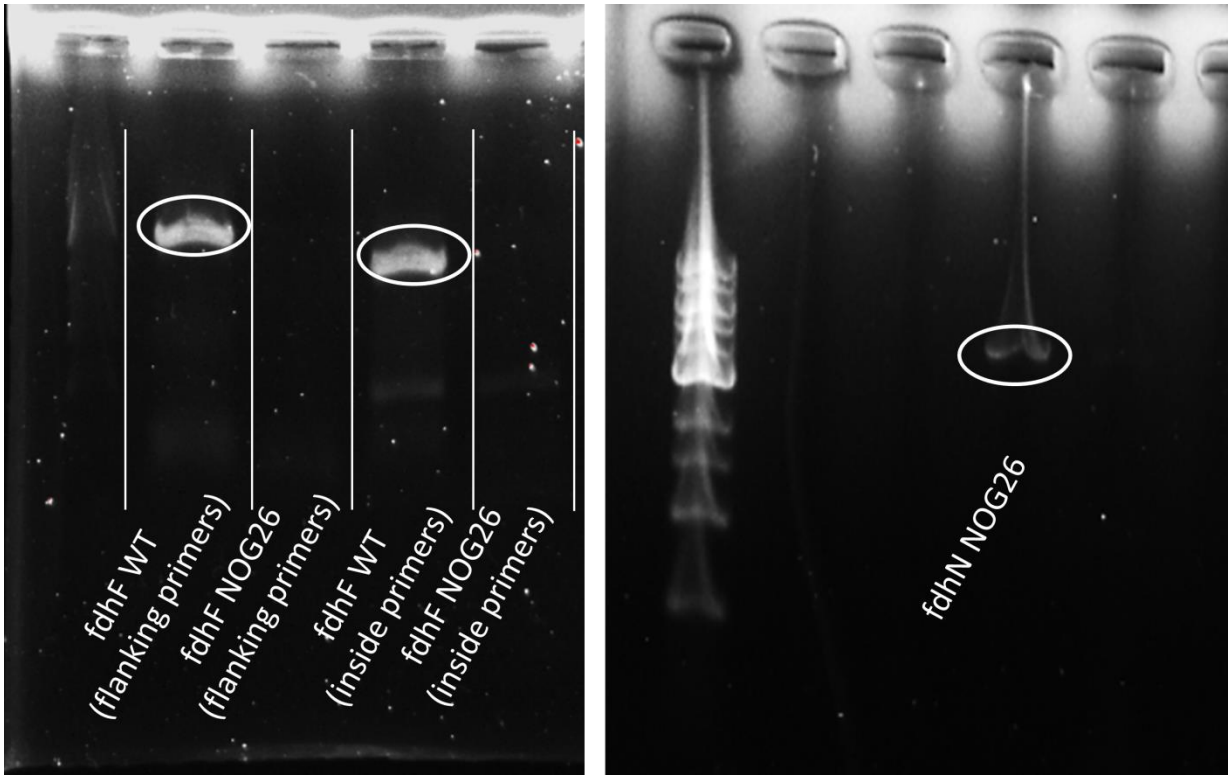


Figure 8 The primary formate-inducible formate dehydrogenase is not functional in NOG26

(Left) The primary formate dehydrogenase (FDH) involved in anaerobic fermentation has been removed from NOG26, likely due to genomic truncation during strain evolution. The wild-type precursor clearly contains a copy of *fdhF*, as evidenced by primers binding flanking the gene, as well as within the gene. NOG26 is missing these bands. (Right) The other major FDH, *fdhN*, is present in NOG26, while we were unable to confirm the low-level, constitutively-expressed *fdhO*. These variants are not likely to significantly impact the proposed pathway.

The third knockout target was *pykA*, the less prevalent pyruvate kinase [97]. While *pykF* had been knocked out, *pykA* was left in NOG26. As pyruvate kinase is the second most forward-

driven glycolytic reaction in *E. coli* (behind only phosphofructokinase) and runs opposite of *ppsA*, it may interfere with driving formate up gluconeogenesis [1]. Two strains were engineered with Δ*arcA* and Δ*pykA* knocked out: JXG3 (containing pTWGP) and JXG4 (containing pTWG3PY).

The final knockout target we considered was the pyruvate dehydrogenase complex (PDH). Retaining PDH could create a redundant loop with reverse *pfl* flux and detract from gluconeogenesis. Alexeeva *et al*. found that maximal *pfl* expression coincided with complete inactivation of PDH [95]. However, impaired PDH function is known to significantly retard growth, which we observed ultimately resulted in lower gluconeogenesis flux from FBA [98]. *AceE*, the E1 component of PDH, was the chosen target, as it is known to impair PDH function but is not involved in other reactions like alpha-ketoglutarate dehydrogenase [99,100]. Two strains were engineered with Δ*arcA*, Δ*pykA*, and Δ*aceE*: JXG5 (containing pTWGP) and JXG6 (containing pTWG3PY). Further confirmation of strain knockouts can be found in Appendix C2.

Another glycolytic isoenzyme that could inhibit the pathway was *pfkB*, which accounts for approximately 10% of *pfk* activity in *E. coli* [101]. This had been acknowledged by Lin *et al*. in their NOG optimization study; however, using [3,4-$^{13}C_2$]-glucose, they confirmed that NOG21, a precursor to NOG26, did not utilize EMP glycolysis [6]. Thus, while knockout of *pfkB* remains a good candidate for future pathway optimization, we did not feel it necessary and focused on removing other genes.

To summarize, starting from the NOG26 strain from the Lin *et al*. pathway engineering study, we constructed six strains that are theoretically capable of turning the GNG-NOG cycle (**Table 4**) [6]. Of the six, three (JXG1, JXG3, JXG5) contain pTWGP, which reintroduced *gapB* from *B.*

subtilis and *pgk* from *E. coli*. The other three strains (JXG2, JXG4, JXG6) contain pTWG3PY, which in addition to the two genes added in pTWGP also added an extra copy of *pfl*, *yfiD*, and *ppsA*. All genes are regulated by an IPTG-inducible promoter. Two strains (JXG1 and JXG2), did not possess any genomic modifications beyond the ones already done in NOG26. Two strains (JXG3 and JXG4) had gene knockouts of *arcA* and *pykA*. The final two strains (JXG5 and JXG6) had gene knockouts of *arcA*, *pykA* and *aceE*.

During growth, production, and labeling studies, a wild-type precursor to NOG26 (JCL16), or NOG26 itself, was used as a negative control [6]. Two strains (JXG4 and JXG5) experienced significantly longer doubling times, and were not considered for several experiments.

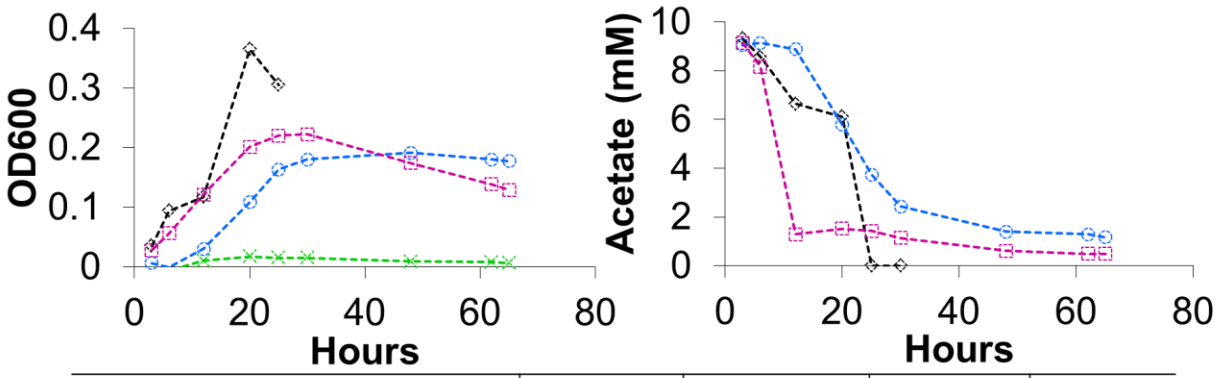
Table 4: Genetic profile of strains involved in experiments

Genetic Modification(s)	JCL16 (WT)	NOG26 (Ref.)	JXG1	JXG2	JXG3	JXG4	JXG5	JXG6
Δ*gapA*, Δ*gapB*, Δ*pgk*, Δ*pfkA*, Δ*pykF*		✓	✓	✓	✓	✓	✓	✓
Δ*arcA*, Δ*pykA*					✓	✓	✓	✓
Δ*aceE*							✓	✓
+*gapB*_BS, +*pgk*			✓	✓	✓	✓	✓	✓
+*ppsA*, +*yfiD*, +*pfl*				✓		✓		✓
Δ*fdhF*		✓	✓	✓	✓	✓	✓	✓

Aerobic growth on different carbon sources showcases engineered strains' gluconeogenic capabilities and mild formate toxicity

We were interested in the gluconeogenic capabilities of engineered strains compared to NOG26. Under aerobic conditions, NOG26 is reliant on glucose to grow, while growth on acetate-only media would require gluconeogenesis to supply larger metabolites. The NOG pathway is theoretically capable of manufacturing larger metabolites by reversing direction, but we were

unsure if reversed NOG was capable of sustaining growth. We attempted to grow cells with acetate as the only carbon source. We chose JXG1 and JXG3 to compare against the wild-type JCL16 and NOG26, as they contained the minimal genes necessary for gluconeogenesis. JCL16, JXG1, and JXG3 were all able to grow, reaching decline phase approximately when the acetate had been depleted (**Figure 9**). NOG26, on the other hand, did not grow, indicating reverse NOG flux was not capable of sustaining growth. Only the strains with the reintroduced gluconeogenic enzymes via pTWGP were able to do so. We have thus restored a functional gluconeogenesis pathway in our engineered strains.



Genetic Modification	JCL16	NOG26	JXG1	JXG3
ΔgapA, ΔgapB, Δpgk		✓	✓	✓
ΔarcA, ΔpykA				✓
ΔaceE				
+gapB_BS, +pgk			✓	✓
+pfl, +yfiD, +ppsA				

Figure 9 Aerobic growth on acetate is possible by engineered strains but not NOG26

OD600 and acetate concentration in media over time of JCL16 (WT), NOG26, JXG1, and JXG3 when fed 10 mM acetate in M9 minimal media. Only NOG26 did not grow. Glucose and formate production was not detected, while stationary/decline phase appears to be reach around when acetate nearly all consumed. Tabulated genetic modifications here are also reflected in Table 4.

Having confirmed growth on acetate in two of our engineered strains, we were more confident that formate would be able to take the place of acetate and contribute to gluconeogenesis. We

examined aerobic growth in 10 g/L glucose-only and 10 g/L glucose with 20 mM formate minimal media (**Figure C3**).

Across all conditions, the wild-type strain JCL16 grew the fastest and reached stationary phase between 24-36 hours. Acetate and formate were produced across all strains, with acetate production beginning around OD of 0.3 and formate production beginning between OD of 0.3 and 1.2 (**Figure C3**). This rise in products corresponding to higher ODs suggested that the higher cell density resulted in a more microaerobic environment that induced fermentation. When 20 mM of formate was added with the same glucose concentration, we saw significant growth inhibition in all NOG-utilizing strains, reaching stationary phase at ODs of around 0.3-0.4, suggesting toxic conditions (**Figure C3**). Acetate production was also significantly hindered. This lack of production coincided with lower ODs, so the more aerobic growth conditions likely inhibited fermentation. Of the fed formate, around 25% (5 mM) was consumed over 72 hours, which is in stark contrast to the wild-type, which produced another 10-15 mM of formate instead of consuming the formate. While formate was being consumed by the engineered strains, there was little difference from the parent strain NOG26, indicating a need for media optimization and/or continued strain engineering.

Decreasing glucose feeding increases formate uptake at the expense of growth and acetate production in engineered strains

From our overall metabolite balances, we identified a range of glucose:formate feeding ratios that would prevent excess production of CO_2 while producing sufficient ATP to fund gluconeogenesis. The aforementioned aerobic experiments utilized a glucose:formate ratio of 2.775:1, resulting in overfeeding of glucose, causing excess usage of TCA cycle and lessening

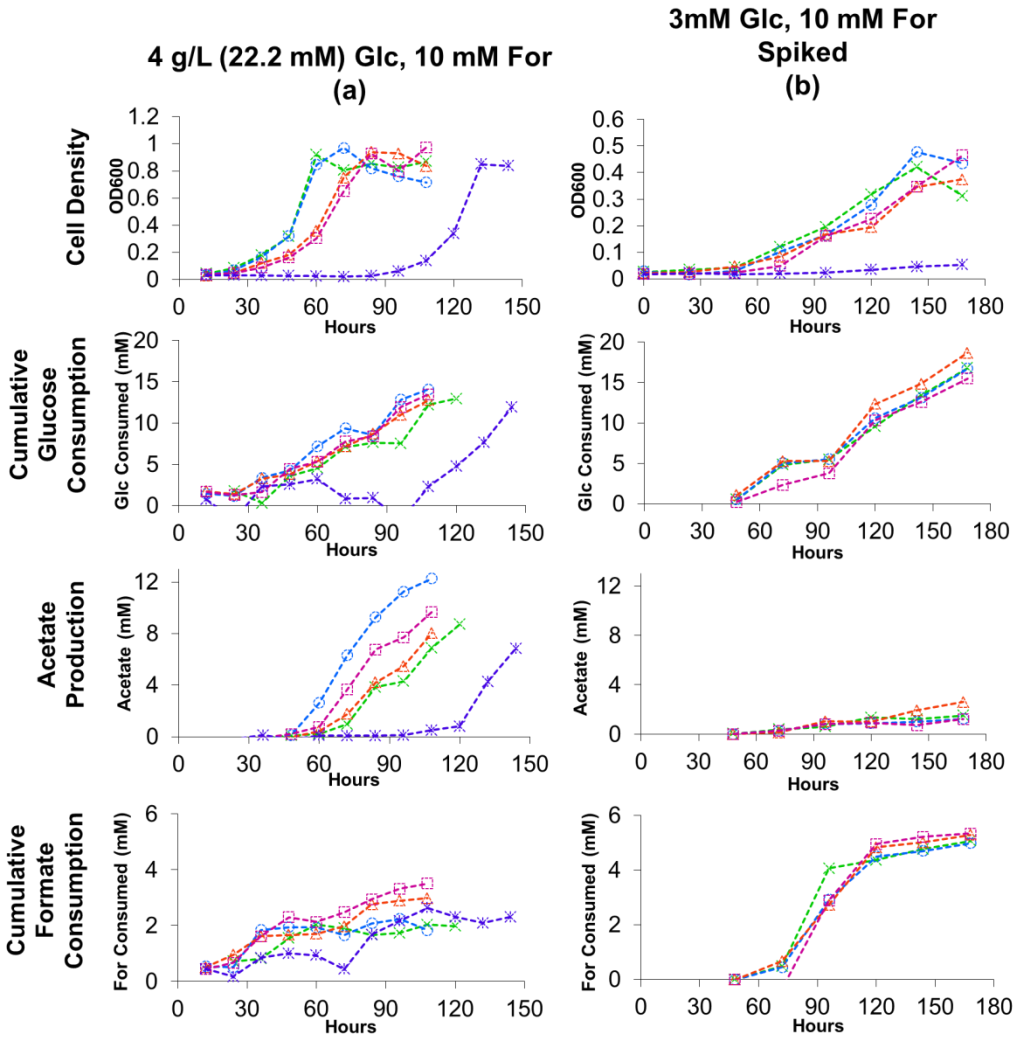
the need for the GNG-NOG cycle. We also observed growth inhibition due to formate toxicity. As a result, we lowered our media concentrations to 4 g/L glucose and 10 mM of formate. This constituted a molar feeding ratio of around 2.22:1 glucose:formate, which falls within the target range. We conducted these experiments in semi-sealed centrifuge tubes, in order limit oxygen flow to the cells and achieve a more microaerobic environment.

We observed a significant increase in the OD reached by cells at stationary phase compared to the aerobic formate feeding experiment, with similar cellular densities to aerobic growth on glucose only (**Figure 10a**). Thus, we believe there to be insignificant toxic effects at the new formate concentration. NOG26 and JXG1 had the highest growth rates of 0.064 and 0.067 hr^{-1}, with JXG2 and JXG3 having growth rates of 0.049 and 0.041 hr^{-1}. JXG4 had a significantly longer lag phase (around 96 hours) at a rate of 0.027 hr^{-1}, but all strains reached a maximum OD of 0.8-1.0. Aside from JXG4, all strains had very similar glucose consumption rates, averaging around 3.13 mM/day. JXG1 produced acetate at the fastest overall rate (around 2.73 mM/day), while NOG26, JXG2, and JXG3 all exported acetate at a similar rate (between 1.8-2.2 mM/day). Formate was consumed consistently by two strains (JXG2 and JXG3), taking up 30-35% of the starting formate, an improvement from the previous experiment. The other three strains, upon reaching exponential phase, produced and consumed formate at similar rates.

Next, we maintained formate at 10 mM but reduced the glucose concentration to 3 mM, and spiked in 3 mM glucose, 0.75 mM formate daily to sustain growth (**Figure 10b**). This ratio was above the minimum 0.25 glucose:formate ratio calculated from the overall metabolite balances, and was used to represent glucose-starved conditions, as this feeding rate was below the daily glucose consumption rate measured at 4 g/L glucose. Aside from JXG4, which did not reach exponential phase and was not included in the metabolite curves, strains had an average growth

rate of 0.019 hr^{-1}, significantly lower than the prior experiment (**Figure 10b**). Growth rates did not vary significantly between strains, suggesting the limited glucose access resulted in similar growth capabilities. Average overall glucose consumption rate was lower than the feeding rate, ranging between 2.22 and 2.66 mM/day. Acetate production decreased relative to the previous experiment, with JXG2 being the highest at 0.37 mM/day. Formate consumption, on the other hand, rose, with JXG2 and JXG3 consuming around 39%, or 5.3 mM of total fed formate. We did not see significant variation in growth rates or formate consumption between NOG26 and the engineered strains, indicating the need for further strain engineering and evolution to improve formate uptake.

By decreasing glucose media concentrations, we were able to achieve increased formate consumption. At 4 g/L glucose, 10 mM formate, we observed consistent formate consumption and higher acetate production than NOG26 by JXG2 and JXG3 (**Figure 10a**). Upon decreasing glucose concentration to 3 mM, and spiking in glucose and formate daily, we observed lower growth and acetate production, but increased formate consumption (**Figure 10b**). In the glucose-starvation experiment, we observed glucose consumption rates lower than the feeding rate, indicating the glucose concentration could be lowered even further to increase formate consumption. However, this would likely further inhibit acetate production and growth. Overall, due to their higher formate consumption rates, we identified JXG2 and JXG3 as the most promising candidates to turn the GNG-NOG cycle.



Genetic Modification	NOG26	JXG1	JXG2	JXG3	JXG4
ΔgapA, ΔgapB, Δpgk	✓	✓	✓	✓	✓
ΔarcA, ΔpykA				✓	✓
ΔaceE					
+gapB_BS, +pgk		✓	✓	✓	✓
+pfl, +yfiD, +ppsA			✓		✓

Figure 10 Decreased glucose feeding increases formate consumption, but limits acetate production

(a) OD600, glucose consumed (mM), acetate produced (mM), and formate consumed (mM) in media over time of NOG26 and JXG1-4, when fed 4 g/L glucose, 10 mM formate in M9 minimal media. OD600 values were normalized by blank and pathway correction. Consumption was normalized by a correction factor to adjust for run-to-run variation, with total consumption summed. **(b)** OD600, glucose consumed (mM), acetate produced (mM), and formate consumed (mM) in media over time of NOG26 and JXG1-4, when fed 3 mM glucose, 10 mM formate in M9 minimal media, with 3mM glucose and 0.75 mM formate spiked in daily.

In addition to acetate, and formate, we also measured other common fermentation products, including pyruvate, succinate, ethanol, and lactate, from media sampling [102]. **Figure 11** displays the distribution of fermentation products in various strains at three different points throughout stationary phase. Standard curves were constructed prior to running samples to convert peak areas to molar concentrations. In addition to measuring each fermentation product, the molar concentration of carbon dedicated to biomass production was estimated from OD600, as well as the total molar concentration of carbon consumed. There was a wide variation in the proportion of carbons that were unaccounted. For example, at 4 g/L glucose, an average of 40% of carbons were unaccounted for, while at 3 mM glucose, an average of 92% of carbons were unaccounted for. These inaccuracies likely arose from inaccurate estimates of intracellular biomass, measurement errors of media samples, and unidentified metabolites.

At 4 g/L glucose, 10 mM formate, the acetate contribution by the engineered strains (approaching or exceeding 40% of total production) was higher than NOG26 (around 20% of total production) (**Figure 11a**). The other significant metabolites were succinate, pyruvate and ethanol, from highest to lowest. Succinate contribution was nearly 60% of total products in NOG26, indicating an increased reliance on TCA cycle or glyoxylate shunt. Succinate production by JXG1 also exceeded acetate production, while in JXG2 and JXG3, acetate production was higher. Surprisingly, we did not observe lactate production in any sample. Biomass levels were relatively constant, as data points were at stationary phase.

When glucose concentration was lowered to 3 mM, the succinate contribution increased in most strains and remained high in NOG26. Only JXG2 had its acetate contribution increase and exceed that of succinate (**Figure 11b**). Pyruvate, ethanol, and overall production levels decreased, likely due to lower growth and access to sugar. We determined from FBA and overall

metabolite balances that the TCA cycle would be a significant contributor under stressed aerobic and microaerobic conditions. An increase in succinate production relative to acetate production after decreasing glucose access supports increased usage of TCA cycle.

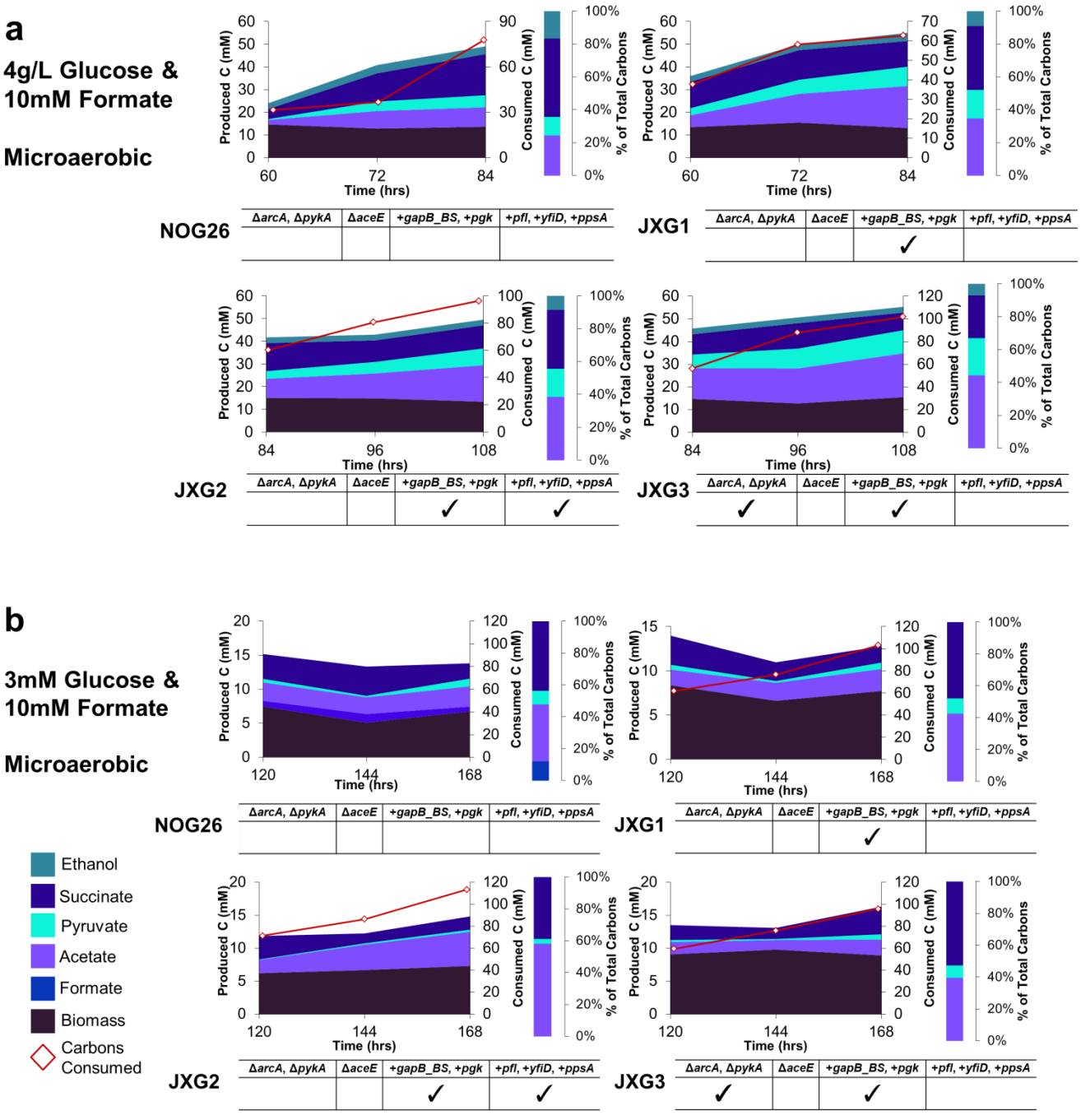


Figure 11 Relative contribution of various products as a percentage of total production shows that only JXG2 favors acetate production under both high and low glucose feeding

(a) Fermentation product distribution of NOG26, JXG1, JXG2, and JXG3 at three time points, after sampling media starting at 4 g/L glucose, 10 mM formate under microaerobic conditions. The area plots sum the total measured products in mM including biomass estimates, with the right y-axis corresponding to the red line, which denotes the total mM of carbons consumed. The percentages in the bar graphs are out of total products averaged across three runs, not including estimated biomass. **(b)** Fermentation product distribution of NOG26, JXG1, JXG2, and JXG3 at three time points, after sampling media starting at 3 mM glucose, 10 mM formate, and spiking in 3 mM glucose, 0.75 mM formate daily, under microaerobic conditions. JXG4 did not reach exponential phase, and as such, was not included in this figure.

Finally, we wished to examine growth and production under anaerobic conditions, including our triple knockout strains. We observed slow growth in all strains (between 0.006-0.02 hr^{-1}) aside from JXG4 (**Figure 12b**), including NOG26, which was reported to not grow anaerobically on glucose-only media by Lin *et al*. [6]. From **Figure 12a**, we see acetate production being the dominant fermentation product, with minor contribution from succinate. Acetate production composed about 70% in NOG26 with succinate being the remaining 30%, similar to the proportions measured by Lin *et al* [6]. JXG3, JXG5, and JXG6 all had lower acetate and higher succinate production than NOG26, while JXG1 and JXG2 had higher acetate production. One possible reason for higher succinate production in JXG3, JXG5, and JXG6 may be the knockout of aerobic regulator *arcA*; these three strains all have *arcA* knocked out, which may have resulted in increased succinate production, as *arcA* deletion is known to upregulate the TCA cycle and glyoxylate shunt [103].

We also compared anaerobic acetate and succinate hourly production rates normalized by cell density to microaerobic production rates at 4 g/L and 3 mM glucose, 10 mM formate (**Figure 12b**). Across all strains, we saw the highest production rates when fed 4 g/L of glucose. Production rates sharply decreased at 3 mM of glucose, 10 mM formate, and decreased again

under anaerobic growth. Under microaerobic growth, JXG1, JXG2, and JXG3 all exceeded NOG26 in acetate production; however, NOG26 was the highest producer of acetate anaerobically. Succinate was the other major metabolite produced by strains under all conditions; for most strains, there was at least one growth condition under which succinate production exceeded acetate production. The one exception was JXG2, which had higher acetate production rates across all conditions. While production was normalized to cell density, the anaerobic production was from a growth experiment, and was likely not at maximum cell density when sampling; as such, anaerobic production rates would likely increase if cells were allowed to grow to higher cell density and then switched to anaerobic conditions. This is likely why anaerobic production was lowest across the board, which contrasted with the findings by Lin *et al*. that NOG26 production was maximized under anaerobic conditions [6]. Overall, we have confirmed that our JXG1, JXG2, and JXG3 strains are capable of growth and acetate production in aerobic, microaerobic, and anaerobic conditions.

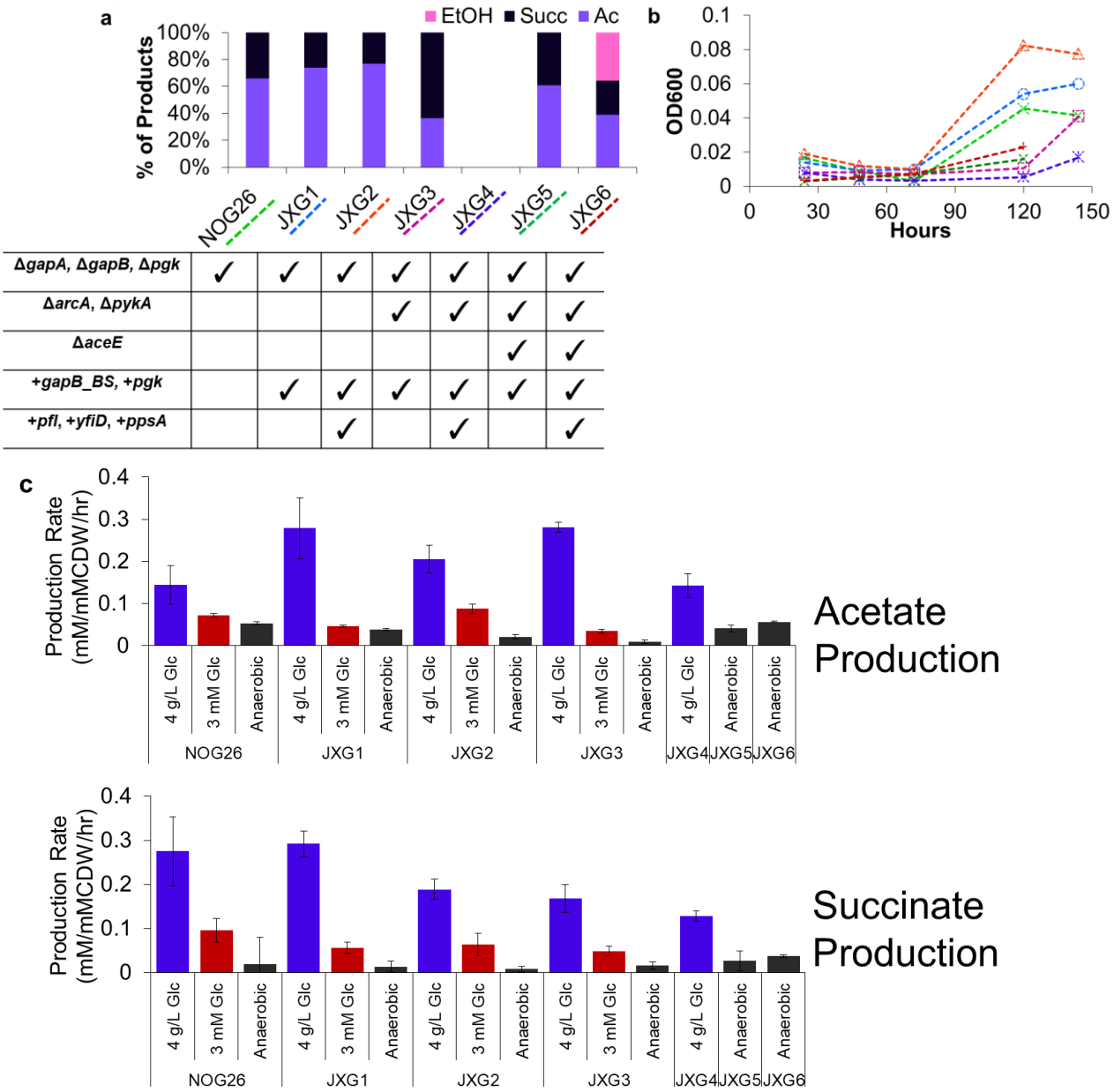


Figure 12 Anaerobic growth in strains shows that acetate and succinate dominate production, while overall production rate estimates decrease from microaerobic to anaerobic growth

(a) Proportion of fermentation products as a percentage of total products under anaerobic conditions, in 4 g/L glucose with 10 mM formate. JXG4 was not able to grow significantly, and as such, was not included. **(b)** Normalized OD600 over time of NOG26 and JXG1-6 under anaerobic conditions. Line colors are reflected in the strain-gene table. **(c)** Acetate and succinate production rates comparing microaerobic growth at 4 g/L glucose, 10 mM formate and 3 mM glucose, 10 mM formate to anaerobic growth. Microaerobic rates were averaged across three time points at stationary phase and anaerobic rates were averaged across the final two time points. Error bars are standard errors of hourly rates of production.

^{13}C-Formate tracing reveals formate uptake by the GNG-NOG cycle by JXG2 and extensive formate incorporation into purines and THF-cycle

We fed cells 10 mM [U-$^{13}C_1$]formate supplemented with 4 g/L of glucose to track the fate of formate. We chose this glucose concentration to facilitate faster growth and higher stationary phase cell density for extraction. After extraction and measuring ^{13}C-labeling, we determined the labeling proportions for essential metabolites, including central carbon metabolites, amino acids, nucleotides, and other cofactors. We also attempted tracing in JXG4 and JXG5, but were unable to achieve significant growth, so these two strains were omitted. All samples were run in triplicate, with error bars denoting the standard error of the mean. When trying to correct for naturally-abundant isotopes, we found that theoretical fractions consistently overestimated labeling. For central carbon metabolites and amino acids, we instead used NOG26 as a negative control, as it is incapable of gluconeogenesis. A two-sided Student's t-test was conducted to determine if the difference in labeling fractions between each engineered strain and NOG26 was significant; a p-value cutoff of 0.05 was used. To estimate the percent of carbons derived from formate, we adjusted the labeling fractions by the lowest NOG26 fraction for each metabolite.

We first wished to confirm that formate was incorporated into gluconeogenesis. We saw a rise in M+2 labeling in fructose-6-phosphate (F6P), glucose-6-phosphate (G6P), and fructose-1,6-bisphosphate (FBP) in JXG1 and JXG2 compared to NOG26 (**Figure 13**). M+2 labeling in hexose phosphates should only occur from two labeled triose phosphates. We would expect to see an increase in M+1 DHAP/GAP in engineered strains compared to NOG26; we were unable to confirm a significant increase, but JXG2 had the highest labeling increase amongst all strains by far, with a p-value around 0.1 (**Figure C5**). M+1 hexose phosphates were also significantly higher in the majority of engineered strains compared to NOG26, with the exception of FBP in

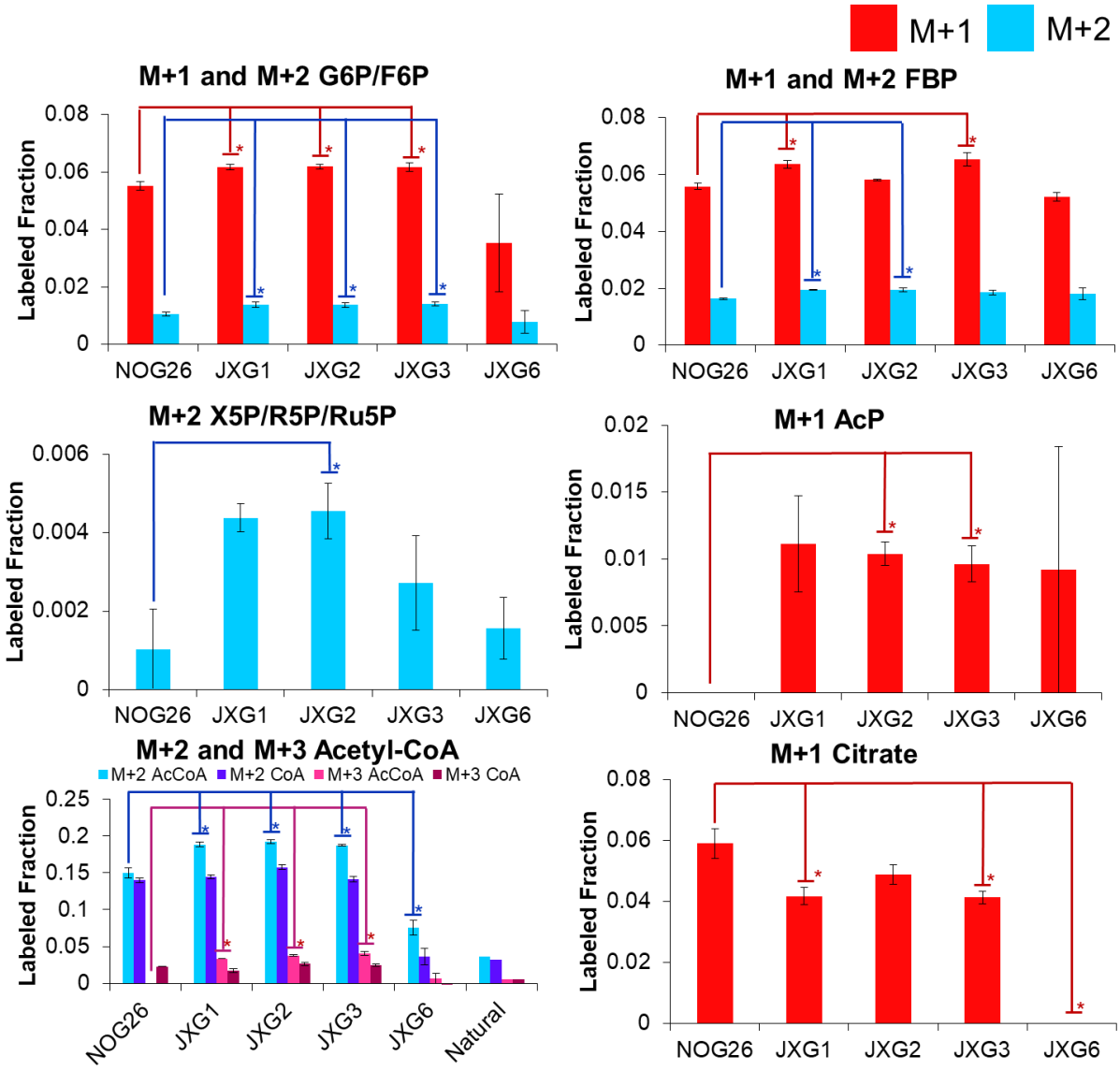
JXG2. For JXG1, JXG2, and JXG3, we estimated 0.3-0.4% of measured carbons in hexose phosphates to be derived from formate. For lower gluconeogenesis, we were unable to confirm significant differences between engineered strains and NOG26; however NOG26 is capable of partial gluconeogenesis through 3-phosphoglycerate (3PG), which Lin *et al*. hypothesized would be active in their NOG strains. Thus, we believed that formate had been incorporated up gluconeogenesis, with JXG2 being the strongest candidate due to its higher DHAP/GAP labeling and M+2 hexose phosphate fractions.

We then sought to determine incorporation of the formate-derived gluconeogenic metabolites down NOG. For only JXG2, we observed increased M+2 labeling compared to NOG26 in NOG intermediates Xu5P, R5P, Ru5P ($p<0.05$) (**Figure 13**). For another NOG intermediate, S7P, we were unable to identify a significant increase in labeling, but JXG2 again had the highest labeling fraction, with a p-value around 0.1 (**Figure C5**). We were unable to identify increased M+1 NOG metabolites. In addition, we looked at AcP, the NOG intermediate before acetate, and observed no labeling in NOG26 and increased labeling in JXG2 and JXG3. While the reverse *pta* reaction is a possible source of labeled AcP [52], the lack of M+1 AcP in the *pta*-capable NOG26 and the M+2 labeling of larger NOG metabolites suggest JXG2 had formate incorporated into NOG. In JXG2, we estimated 0.5% of AcP carbons and 0.3% of Xu5P/R5P/Ru5P carbons to be derived from formate, comparable to the incorporation observed in hexose phosphates. We were unable to directly measure labeling in acetate, as its mass-to-charge ratio was too low for our LC-MS method to confidently detect.

M+1 labeling composed over half of acetyl-CoA (AcCoA), which was significantly higher than the expected natural abundance. Examining free CoA, we found the M+1 CoA fraction exceeded M+1 AcCoA, indicating that the majority of M+1 labeling in AcCoA was on the CoA group. On

the other hand, when we looked at M+2 and M+3 labeling, we saw higher fractions in AcCoA than CoA (**Figure 13**). This difference in labeling of a second or third carbon is believed to be on the acetyl group, derived from formate via AcP and the *pta* reaction. While production of acetyl-CoA from pyruvate readily occurs via PDH, the carbon lost during this reaction would be the same carbon added via the reverse *pfl* reaction, which should result in no labeling on the acetyl group. Thus, any increase in labeling should be produced via NOG. Compared to NOG26, JXG1, JXG2, and JXG3 all had significantly higher labeling fractions M+2 and M+3 AcCoA ($p<0.05$), suggesting that all the engineered strains had higher formate uptake via *pfl*.

Using [U-^{13}C]formate, we confirmed that the JXG2 strain had significantly higher labeling in upper gluconeogenesis and NOG. We were unable to comprehensively identify increased labeling of all involved metabolites, but the pathway likely had low flux, due to lack of optimization, resulting in low peak intensities of labeled metabolites and potentially underestimating labeling. Compared to other engineered strains, only JXG2 showed increased M+2 labeling in hexose phosphates, NOG intermediates Xu5P, R5P, and Ru5P, as well as M+1 labeling in AcP and AcCoA. In addition, JXG2 had the highest labeling fraction in other key metabolites, such as S7P, DHAP and GAP. Thus, we believe that JXG2 turned the GNG-NOG cycle, albeit inefficiently, with less than 1% of involved carbons generated from formate.

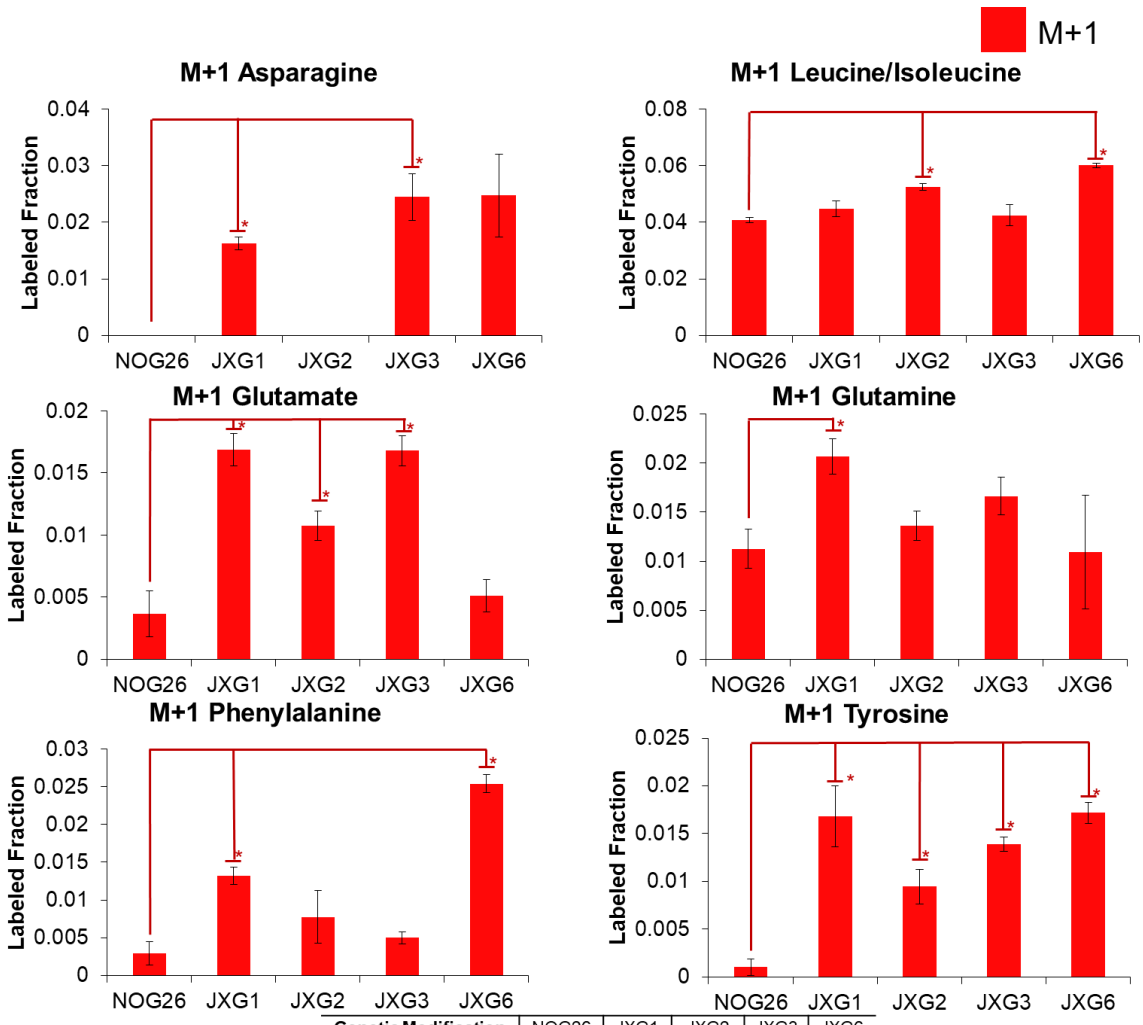


Genetic Modification	NOG26	JXG1	JXG2	JXG3	JXG6
ΔgapA, ΔgapB, Δpgk	✓	✓	✓	✓	✓
ΔarcA, ΔpykA				✓	✓
ΔaceE					✓
+gapB_BS, +pgk		✓	✓	✓	✓
+pfl, +yfiD, +ppsA			✓	✓	✓

Figure 13 Increased M+2 labeling in six carbon metabolites suggests incorporation of formate into gluconeogenesis and non-oxidative glycolysis by JXG2

Labeling fractions of metabolites involved in gluconeogenesis. M+1 and M+2 labeling fractions of glucose-6-phosphate/fructose-6-phosphate (G6P/F6P), fructose-1,6-bisphosphate (FBP), xylulose-5-phosphate (Xu5P), ribose-5-phosphate (R5P), ribulose-5-phosphate (Ru5P), acetylphosphate (AcP), acetyl-CoA, CoA, and citrate for NOG26, JXG1, JXG2, JXG3, and JXG6, with engineered strains compared against NOG26. Cells were cultured in 4 g/L mM glucose, 10 mM [U-$^{13}C_1$]formate, and extracted from OD600 0.13 to 0.62. All samples were run in triplicate, and error bars are standard errors of the mean. Two-sided student's t-tests were run to confirm significant variation between NOG26 and engineered strain, with solid lines and * denoting $p<0.05$.

As we saw significant labeling in acetyl-CoA, we were also interested in the extent of labeling in TCA cycle. In the majority of TCA metabolites, including α-ketoglutarate (αKG), fumarate, and malate, we were unable to determine a significant difference between NOG26 and engineered strains. Only citrate showed significant decrease in labeling by engineered strains from NOG26 (**Figure 13**), with lower formate incorporation in JXG1, JXG3, and JXG6. This indicated that reverse *pfl* flux was able to consume fed formate in NOG26, but directed more of the produced pyruvate towards TCA cycle, as gluconeogenesis is not possible in NOG26.



Genetic Modification	NOG26	JXG1	JXG2	JXG3	JXG6
Δ*gapA*, Δ*gapB*, Δ*pgk*	✓	✓	✓	✓	✓
Δ*arcA*, Δ*pykA*				✓	✓
Δ*aceE*					✓
+*gapB_BS*, +*pgk*		✓	✓	✓	✓
+*pfl*, +*yfiD*, +*ppsA*			✓	✓	✓

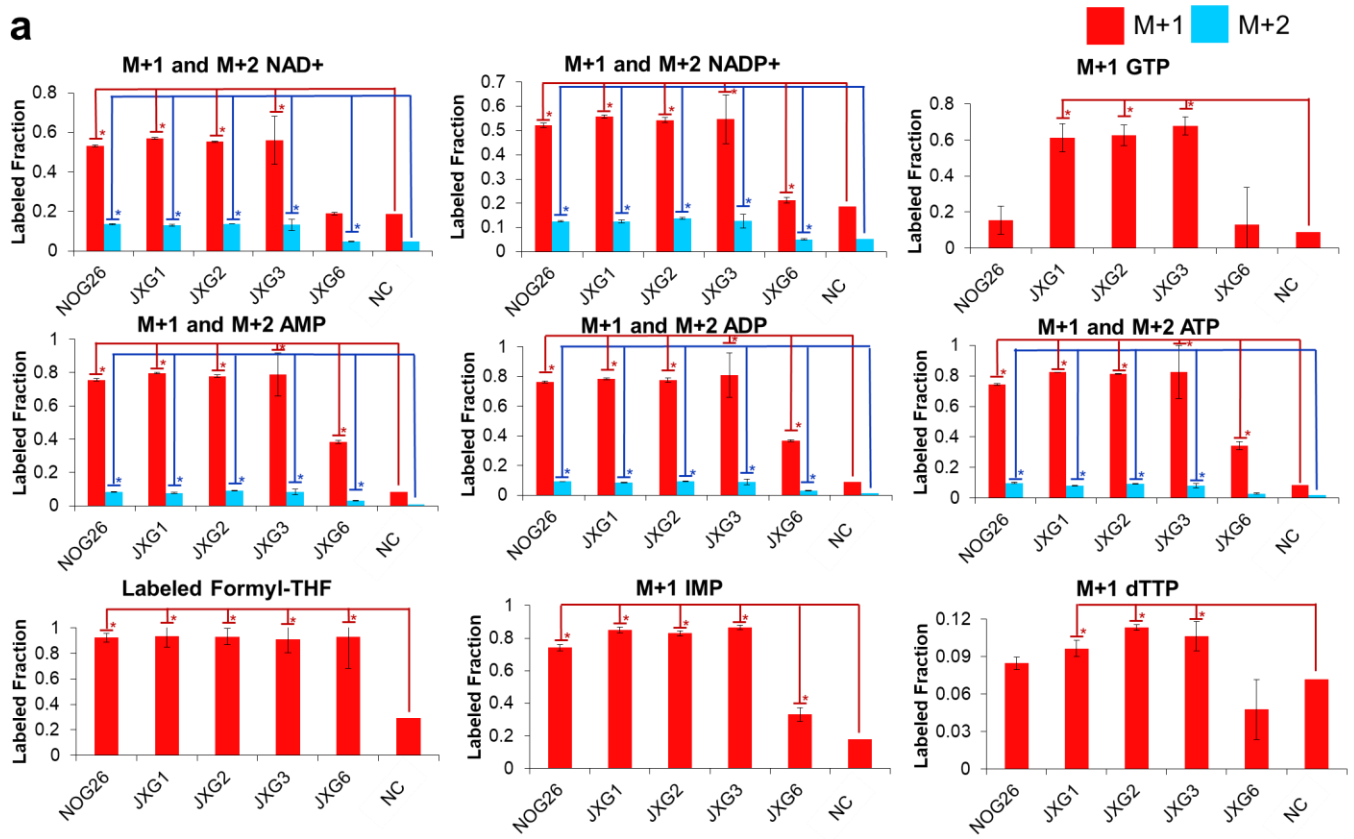
Figure 14 Higher labeling fractions in amino acids indicate higher formate consumption by engineered strains

M+1 labeling fractions of asparagine, leucine/isoleucine, glutamate, glutamine, tyrosine, and phenylalanine for NOG26, JXG1, JXG2, JXG3, and JXG6, with the engineered strain compared against NOG26. Cells were cultured in 4 g/L mM glucose, 10 mM [U-$^{13}C_1$]formate, and extracted from OD600 0.13 to 0.62. All samples were run in triplicate, and error bars are standard errors of the mean. Two-sided student's t-tests were run to confirm significant variation between NOG26 and engineered strain, with solid lines and * denoting $p<0.05$.

Amino acid synthesis is tied to various points of central carbon metabolism, so we measured formate's incorporation into different amino acids. Generally, we observed higher labeling fractions in engineered strains compared to NOG26 (**Appendix C6**). Glutamine and glutamate, two amino acids generated from αKG, had significantly increased labeling fractions in the majority of engineered strains (**Figure 14**). We also observed increased labeling in asparagine, which derives from oxaloacetate, and in leucine/isoleucine, which derive from pyruvate. We also saw higher labeling of phenylalanine and tyrosine in all engineered strains compared to NOG26. Aromatic amino acid synthesis arises from a common precursor, chorismate, which derives from PEP and erythrose-4-phosphate, metabolites which are involved in the GNG-NOG cycle [104]. The higher labeling of amino acids suggests higher formate consumption by engineered strains compared to NOG26, which is supported by the formate uptake rates measured from media sampling of cultures fed 4 g/L glucose (**Figure 10**).

We also examined labeling in nucleotides, and found significant labeling in important cofactors, including NAD^+, $NADP^+$, GTP, AMP, ADP, and ATP (**Figure 15a**). Compared to central carbon and amino acid metabolites, purine labeling fractions were much higher, typically having >70% of detectable purines labeled, while other metabolites were typically <10% labeled (**Figure 15a**). For the metabolites with significantly raised labeling, we found that negative control samples fed unlabeled glucose were relatively close to the theoretical naturally-occurring labeling abundance; thus, for these metabolites specifically, we compared all strains against the negative control using two-sided t-tests ($p<0.05$) to establish a significance of the labeling. We also measured increased labeling in purine synthesis intermediates IMP and formyl-THF. We also looked at pyrimidines, represented by dTTP, and observed a significant increase compared to the negative control, but significantly lower total labeling compared to purines (**Figure 15a**). Formate can be

incorporated into purines and THF pathway via several potential routes (**Figure 15b**), which has been noted in other studies, but the relative contributions of the different paths has not been determined [64,66]. CoA and dTTP can also be synthesized *de novo* by *E. coli* via 5,10-methylene THF, and we observed significant labeling in CoA, acetyl-CoA, and dTTP (**Figure 13**). The labeling of dTTP is especially striking, as transfer of labeling from 5,10-methylene THF to dTTP would likely require the labeling of the formyl group generated via the unfavorable reverse phosphoribosylglycinamide formyltransferase 2 (*purU*) reaction. The other potential path for pyrimidine labeling would be via CO_2 formed from formate via FDH, and then incorporated into carbamoyl phosphate and carbamoyl aspartate to form pyrimidines; we did not observe any labeling in these intermediates, which supports formate incorporation through *purU*. With labeling fractions in CoA, THF, and purine metabolism being significantly higher than the unlabeled control, we believe that the majority of consumed formate was directed towards these fates. This finding represents a significant formate sink that may need to be removed in order to facilitate the GNG-NOG cycle.



Genetic Modification	NOG26	JXG1	JXG2	JXG3	JXG6
ΔgapA, ΔgapB, Δpgk	✓	✓	✓	✓	✓
ΔarcA, ΔpykA				✓	✓
ΔaceE					✓
+gapB_BS, +pgk		✓	✓	✓	✓
+pfl, +yfiD, +ppsA			✓	✓	✓

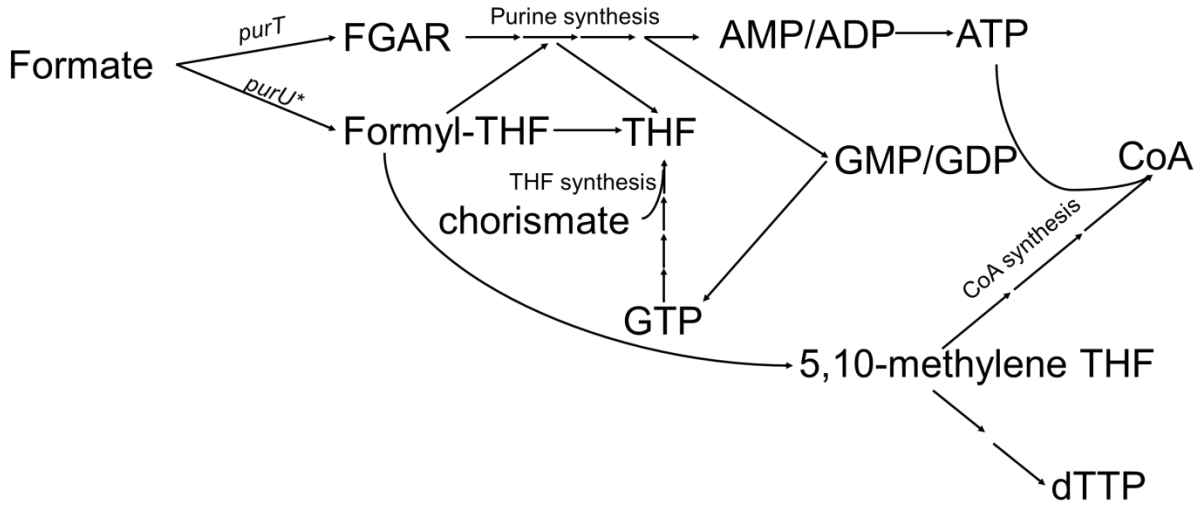


Figure 15 ^{13}C-labeling in nucleotides shows significant formate incorporation into purine synthesis, THF cycle, and CoA synthesis

(a) Labeling fractions of metabolites involved in purine synthesis. M+1and M+2 labeling fractions of NAD^+, $NADP^+$, GTP, AMP, ADP, ATP, formyl-THF, IMP, and dTTP for NOG26, JXG1, JXG2, JXG3, and JXG6 were compared against an unlabeled negative control. Cells were cultured in 4 g/L mM glucose, 10 mM [U-$^{13}C_1$]formate, and extracted from OD600 0.13 to 0.62. All samples were run in triplicate, and error bars are standard errors of the mean. Student's t-tests were run to confirm significant variation between strains and the negative control with * denoting $p<0.05$. (b) Diagram of potential paths for formate incorporation into nucleotides, one-carbon units, and CoA by *E. coli*. Italicized labels denote initial enzymes incorporating formate, although some reactions (*) are unfavorable under most conditions. Pathways denoted using multiple arrows are not too scale. FGAR denotes N^2-formyl-N^1-(5-phospho-β-D-ribosyl)glycinamide; THF, tetrahydrofolate; AMP, adenosine monophosphate; ADP, adenosine diphosphate; ATP, adenosine triphosphate; GMP, guanosine monophosphate; GDP, guanosine diphosphate; GTP, guanosine triphosphate; CoA, cofactor A; *purT*, phosphoribosylglycinamide formyltransferase 2; *purU*, formyltetrahydrofolate deformylase.

Absolute metabolite concentration estimates support reactivation of gluconeogenesis and higher succinate production in engineered strains

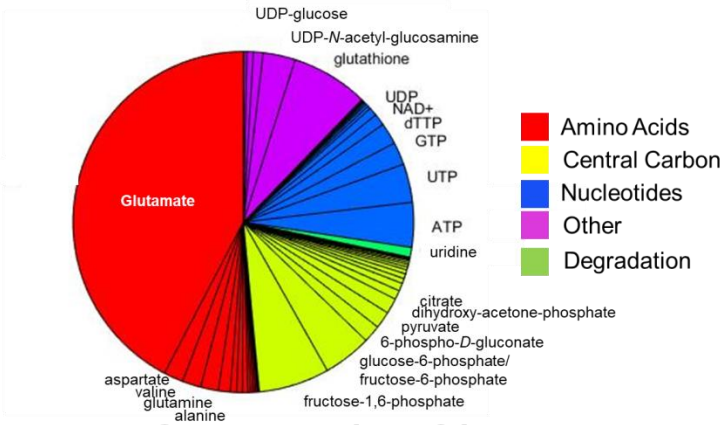
In addition to the ^{13}C-formate labeling experiments, we utilized [U-$^{13}C_6$] glucose to estimate absolute intracellular metabolite concentrations using a reference *E. coli* strain, NCM3722. While we were unable to account for the full suite of metabolites measured in the Park *et al.* study, we did measure the majority of metabolites involved in the GNG-NOG cycle, as well as amino acids, cofactors, and nucleotides [3]. All absolute metabolite concentrations and standard errors can be found in Appendix D.

We plotted the relative distribution of broad metabolite types for NCM3722, NOG26, and engineered strains JXG1, JXG2, and JXG3 (**Figure 16**). We did attempt to estimate concentrations for our triple knockout strain, JXG6, but found significantly lower concentrations compared to our other engineered strains and NOG26. We did adjust for optical density, but the lower concentrations were likely due to an inaccurate measurement of OD, as overestimating the OD would result in underestimating the metabolite concentrations. Thus, while we calculated the concentrations in JXG6, we did not include them in the final table. The ^{13}C-formate labeling experiment, on the other hand, did not rely optical density, but rather the relative labeling fractions from LC-MS; as such, this inaccuracy should not significantly impact the labeling fractions measured.

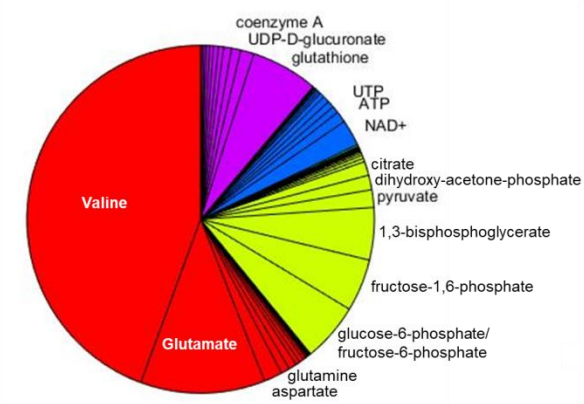
Amino acids composed the majority of measured metabolites, while metabolites involved in gluconeogenesis/glycolysis and non-oxidative glycolysis were the second largest category. Across the board, amino acids contributions increased while nucleotide synthesis decreased in NOG26 and the engineered strains compared to the reference strain. Compared to NOG26, the

engineered strains had higher central carbon contribution. The engineered strains also had valine as the largest amino acid contributor, while the reference strain was dominated by glutamate; as we did not wash our cells before extracting, it is likely that residue from overnight rich media cultures contaminated the samples, and contributed to the extraction. This may also be why we observed a larger overall pool of metabolites in our experiments compared to the Park *et al*. study.

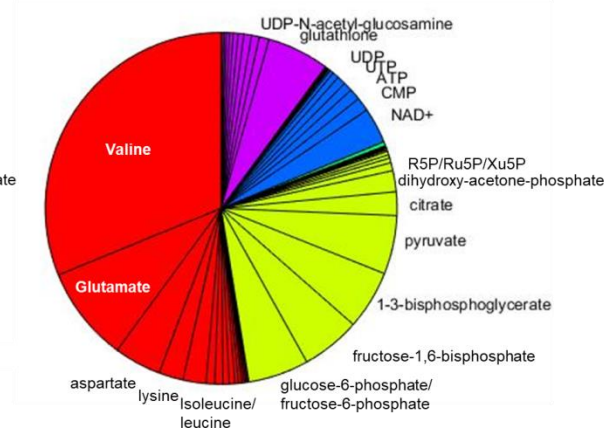
One notable metabolite was succinate, which we have highlighted beforehand as a significant fermentation product in our engineered strains. We did see higher succinate concentrations in JXG1, JXG2, and JXG3 versus NOG26. Another key metabolite was 1,3-bisphosphoglycerate (13BPG), a product that should be significantly lower in NOG26, as it lacks GAPDH [6]. We saw higher 13BPG concentrations in JXG1, JXG2, and JXG3 compared to NOG26, as expected. Of the engineered strains, JXG2 had the lowest concentrations of nucleotides and central carbon metabolites, while having the highest contribution by amino acids and other metabolites. JXG1 had the highest central carbon content. All three engineered strains had significantly higher total metabolites than NCM3722 and NOG26.



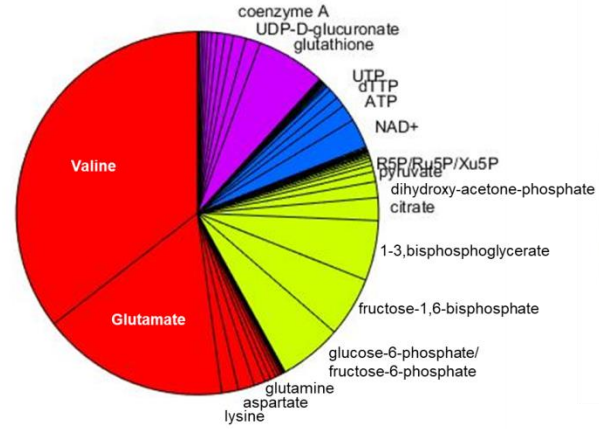
NCM3722(Ref.)
Total Pool: 0.2364 M



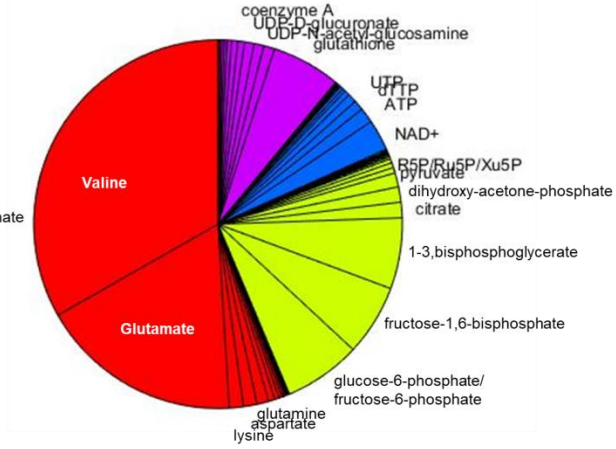
NOG26
Total Pool: 0.2264 M



JXG1
Total Pool: 0.3058 M



JXG2
Total Pool: 0.2949 M



JXG3
Total Pool: 0.2846 M

Figure 16 Relative distributions of intracellular absolute metabolite concentrations shows higher contribution to GNG/NOG cycle by engineered strains compared to NOG26 and NCM3722

Pie charts of different metabolite types for the reference strain NCM3722, NOG26, and engineered strains JXG1, JXG2, and JXG3. NCM3722 was grown in Gutnick minimal media with 4 g/L [U-$^{13}C_6$]glucose and used to estimate extracted samples from [U-$^{13}C_1$]formate tracing experiments. NCM3722 estimates were taken directly from [3]. The full table of absolute metabolite concentrations can be found in Appendix D.

3.3 Discussion

Formate is an intriguing feed substrate, as it can be produced from CO_2 and is cheaper than typical sugar feedstocks [105]. However, formate is a one carbon metabolite and essential anabolic reactions are energetically expensive, making growth on formate difficult. There have been several studies that have successfully grown *E. coli* on only one-carbon compounds [64,66]; however, these pathways required a large suite of heterologous genes and could only grow in the presence of oxygen or CO_2. Introduction of heterologous pathways is challenging, as introduction of large vectors may be unstable and may provide a repressive protein burden on cells. The lack of flexibility regarding gaseous access is relevant in an industrial setting, as it is common to manipulate oxygen access as a metabolic switch between prioritizing growth and prioritizing production. The GNG-NOG cycle was constructed with mostly native genes, minimizes carbon loss via NOG, and is flexible in its gaseous culture conditions. Acetate is a versatile product that is used in the production of therapeutics [106,107], flavorants [108,109], and bulk chemicals [110,111], to list a few applications. A strain with a fully-developed GNG-NOG cycle would be able to produce acetate from one-carbon compounds such as formate or carbon dioxide with minimal carbon loss, and could then be effectively coupled with another organism that consumed acetate to produce a more complex product. Bioproduction using multiple bioreactors in series is a common setup [112,113], and would be one potential application of the GNG-NOG cycle.

Through plasmid engineering and CRISPR knockouts, we engineered six different strains of *E. coli* capable of both gluconeogenesis and NOG. Using HPLC to analyze media samples, we identified acetate production and formate consumption by several strains, the most promising of which were JXG2 and JXG3, which consumed formate consistently at 4 g/L glucose, 10 mM

formate under microaerobic conditions. Upon reducing glucose concentration to 3 mM and spiking in a low amount of glucose and formate daily, we were able to increase formate uptake, but at the expense of acetate production, with JXG2 being the only strain having acetate production exceed succinate production. We were able to confirm growth, formate consumption and acetate production under both microaerobic and anaerobic conditions. We found an increase in gluconeogenic flux in JXG2 compared to NOG26 using ^{13}C-formate, as well as increased labeling in NOG metabolites. Thus, we believed that JXG2 would be a good candidate to continue GNG-NOG cycle optimization.

While we were able to engineer a strain capable of incorporating formate into acetate, we observed low flux through the GNG-NOG cycle, with less than 1% of carbons throughout the cycle originating from formate. We identified several formate sinks that may be diverting flux away from the desired pathway. This included the increased production of succinate in many of our engineered strains, as well as the uptake of formate into purine and CoA synthesis as well as the THF-cycle, which we identified from feeding ^{13}C-formate.

From our media experiments, we identified succinate production to be elevated in our engineered strains, compared to expected production by NOG26. This may be due to the oxidative stress associated with formate uptake, with Ahn *et al*. identifying a role for the glyoxylate shunt in reducing electron buildup [114]. We saw the majority of our strains preferentially produce succinate over acetate under microaerobic conditions (**Figure 10**). Three of our strains, JXG3, JXG5, JXG6, had the *arcA* regulator knocked out with the intention of increasing microaerobic *pfl* flux; unfortunately, it seems that these strains also had increased succinate production under anaerobic conditions as a result of TCA cycle dysregulation. The one exception was JXG2, which produced more acetate than succinate under all growth conditions. In future experiments,

it may be advisable to add an electron acceptor such as nitrate to decrease the reliance on glyoxylate shunt [72]. In addition, we may seek to attenuate the glyoxylate shunt and/or TCA cycle under microaerobic conditions, in order to increase acetate titers.

Using NOG26, a strain incapable of gluconeogenesis, as a baseline, we found that ^{13}C-formate was more highly incorporated by engineered strains into gluconeogenesis. Particularly, we observed increased M+2 hexose phosphates by engineered strains compared to NOG26. However, only the JXG2 strain also had significantly higher labeling in measured NOG metabolites, and had the highest labeling in DHAP/GAP and S7P. Considering both the increased labeling throughout the pathway as well as the higher formate consumption across multiple microaerobic experiments, we believe JXG2 turned the GNG-NOG cycle.

One interesting result from the formate labeling experiment was the high diversion of labeled formate into purine and CoA synthesis, as well as the THF pathway. The labeling fractions in these metabolites far exceeded the expected natural abundance, and indicated a high flux of formate to these pathways. There are several potential pathways for the entry of formate into these metabolites, but we were unable to resolve the exact mechanism. We saw in JXG1, JXG2, and JXG3, higher labeling fractions in numerous nucleotides compared to the unlabeled negative control. The higher labeling fractions in nucleotides, as well as amino acids, indicated higher overall formate consumption by the engineered strains. These pathways act as a major formate sink for our strains, and may need to be altered in order to facilitate GNG-NOG cycle flux. It may also be necessary to supplement a tertiary carbon source, such as glycine, that will contribute to purine synthesis and allow formate to proceed through GNG-NOG cycle.

Reaction thermodynamics can inform researchers as to the viability of certain pathways and potential bottlenecks to address. Using FBA, we identified potentially beneficial genetic modifications. While our simulations were done manually through the COBRA toolbox, further optimization could allow for a more comprehensive simulation, iterating through all possible genetic knockouts or modifications to identify optimal conditions. Further thermodynamic analysis to determine reaction fluxes and free energies of the engineered strains would accelerate pathway design by revealing important target reactions, especially considering the energetically-expensive nature of gluconeogenesis [3,115,116]. Recent studies have shown the need to examine pathways at the individual reaction level, rather than assuming a rate-determining step [1,2,52]. For the GNG-NOG cycle, reactions such as *pfl*, PDH, GAPDH, and *ppsA*, come to mind as impactful reactions that may act as bottlenecks.

Thermodynamics has also informed researchers about the balance between protein burden and reaction favorability [1]. A 2019 Park *et al*. study found that several glycolytic reactions in *E. coli* existed near-equilibrium, reflecting an excess of produced enzymes [1]. While the higher protein production placed a higher burden on cells, Park *et al*. found that the excess enzymes helped cells adjust more rapidly to changing nutrient conditions [1]. As we spiked in glucose and formate at regular intervals during several experiments, it would be interesting to identify near-equilibrium reactions in our engineered strains that may serve a similar purpose to poise cells upon reintroduction of carbon sources. In addition, we may also identify an excessive burden from the pTWGP and pTWG3PY plasmids, which overexpress their genes upon IPTG induction. This protein cost may be the reason behind the slow growth of JXG4 and JXG5, which may require adjustments to promote growth, such as using a weaker promoter.

3.4 Future Work

We do believe there is a significant amount of work to be done in order to optimize the pathway and test all of our strains. We were unable to examine our triple knockout strains under microaerobic conditions, and would like to do so in order to confirm formate uptake. We would also want to conduct time course experiments under anaerobic conditions, and allow cells to reach higher cellular density, if possible. For microaerobic experiments, it would be beneficial to run cultures in chemostats instead of our current tube setups, in order to better control oxygen exposure and ensure microaerobic conditions. For our triple knockout strains and JXG4, we found growth to be slow, and would like to optimize this. This would likely involve evolution when grown on minimal media containing formate and low glucose over multiple generations, which may also increase formate uptake. After further optimizing strains, we would like to repeat the ^{13}C-formate tracing experiment, as well as tracing with [3,4-$^{13}C_2$]glucose, which would confirm that NOG, and not EMP glycolysis, was being utilized [6]. Other potential genetic modifications include knockout of *pfkB* and the other two FDHs.

It would be beneficial to repeat the ^{13}C-formate labeling experiment, to confirm more confidently the turning of the GNG-NOG cycle through additional replicates. We would also look to elucidate the pathway taken by formate in order to enter purine and THF metabolism, as this appears to be a major formate sink in our strains. This would likely involve generating a library of single knockout strains, targeting *purT*, *purU*, *folE*, and other relevant genes, and examining formate labeling under each condition. We would then knock out the largest contributors, and supply any necessary nutrients to make up for potential deficiencies.

While we employed different media conditions and genetic modifications to optimize pathway formate uptake, we did not explore strain evolution strategies, which have had success in optimizing heterologous pathways [6,66,117-119]. Typically, a pre-existing strain will be grown on the desired nutrient source (in our case, formate), and allowed to evolve over multiple generations. Repeated mutagenesis can result in previously-unknown beneficial mutations to adapt to the new nutrient conditions. Growth on formate while decreasing glucose concentration in media may result in strains capable of higher formate uptake and increased GNG-NOG cycle usage.

Ultimately, the goal of our engineered pathway would be to utilize carbon dioxide, not formate, as the starting substrate. Without the introduction of heterologous machinery, FDH remains the most likely candidate to drive CO_2 assimilation. While *fdhF* has been truncated in our engineered strains, *fdhN* remains in the genome. While the functionality and thermodynamic viability of the *fdhN* reverse reaction has yet to be extensively assessed, its nitrate-inducible nature makes it an attractive option for the assimilation of CO_2 [90].

3.5 Methods

Culturing conditions of engineered strains

Upon finalizing the construction of our strains, we wished to identify ideal culturing conditions facilitate the usage of the GNG-NOG cycle. For all experiments, strains were grown overnight in rich LB liquid medium with ampicillin and spectinomycin, to select for pPL274 and the pTW plasmids respectively. Minimal media was then seeded with the liquid cultures to approximately the same optical densities. In addition to the M9 media base and carbon source(s), 0.5 mM IPTG, 100 µg/mL ampicillin, 100 µg/mL spectinomycin, and a 1% vitamin mix were added to all

minimal media experiments. All ODs were measured by the BioTek PowerWave HT Microplate Spectrophotometer in 96 well plates. Pathway and blank corrections were conducted using the BioTek Gen5 Data Analysis Software prior to recording results.

For JCL16, NOG26, JXG1, and JXG3, the average rates of glucose and formate consumption under aerobic conditions were examined. OD600, and glucose, acetate, and formate levels were measured periodically over 96 hours. Three different media conditions were examined: 10 g/L of glucose only, 10 g/L glucose with 20 mM formate, and 10 mM acetate only. Acetate only conditions were used to confirm the gluconeogenic capabilities of our engineered strains. Experiments were done in 15 mL conical tubes with a culture volume of 10 mL.

After measuring the formate and glucose consumption rates, we sought to examine the effect of altering media conditions. We conducted these media experiments with NOG26 and JXG1-4. We noticed significant growth inhibition at 20 mM of formate, and thus lowered the concentration to 10 mM. We first lowered the concentration of fed glucose to 4 g/L, in order to measure consumption at the new formate concentration with an excess of glucose. We also lowered the glucose concentration to 3 mM (about 0.54 g/L) for all strains to starve them of glucose and increase formate dependence. All cultures were conducted in 50 mL conical tubes with 40 mL working volume. The tubes were semi-tightly capped with a parafilm cover, which, combined with measurements at higher cell density, simulated a microaerobic environment. All media experiments were run in duplicate and averaged. Standards of known concentrations were run before HPLC measurements in order to determine a correction factor to adjust for run-to-run variation.

After engineering the triple knockout strains, we sought to examine their ability to grow anaerobically alongside our previously engineered strains. To conduct anaerobic experiments, overnight cultures were added to degassed media and Hungate tubes in an anaerobic chamber. Carbon sources included 4 g/L glucose and 10 mM formate. 50 mM of sodium nitrate was also added to act as an electron acceptor [72]. Glucose, acetate, and formate were measured at three time points to minimize potential exposure to oxygen.

Analyzing pathway feasibility via Flux Balance Analysis

Due to the lack of knowledge surrounding NOG and the energetically-expensive nature of gluconeogenesis, it may be thermodynamically challenging to establish a functioning pathway. In order to assess the viability of the pathway shown in Figure 3, *in silico* analysis was performed. The COnstraint-Based Reconstruction and Analysis (COBRA) Toolbox is a MATLAB add-on that can provide quantitative estimates of biochemical systems using constraint-based modeling [120]. Using Flux Balance Analysis (FBA), a stoichiometric matrix balancing metabolites and metabolic reactions is created and linear programming is used to calculate flux values of all reactions depending on constraints and the objective function, a single reaction that is to be maximized or minimized [120].

We employed the *E. coli* core model supplied by COBRA Toolbox developers to model the proposed pathways under various conditions, including various gene knockout and addition combinations, to determine its feasibility and likely limiting reactions. Upon confirming the flux constraints on each gene was set as desired, one of four objective functions was set: maximizing biomass production (indicative of maximizing growth rate), acetate export, or *pfl* flux, and minimizing glucose flux required to sustain growth. From the generated flux values, we identified key reactions to the desired pathway.

While the toolbox covered much of central carbon metabolism, several updates were made to account for reactions key to the pathway. Two reactions relevant to the pathway, *xpk* and *fdh*, were added to the model. The former is an essential heterologous reaction in NOG while the latter may divert fed formate away from the desired pathway. The other change made to the base model was to make *pfl* a reversible reaction, as new evidence has suggested [72].

In addition, from the original Lin *et al*. study, NOG26 had a suite of central carbon genes knocked out or mutated to facilitate flux through NOG [6]. Knockout of glucose-6-phosphate dehydrogenase (*zwf*), phosphogluconate dehydratase (*edd*) and Entner-Doudoroff aldolase (*eda*) inhibited the Entner-Doudoroff (ED) pathway, while knockout of phosphofructokinase I (*pfkA*), glyceraldehyde-3-phosphate dehydrogenase (*gapA*), erythorse-4-phosphate dehydrogenase (*gapB*), phosphoglycerate kinase (*pgk*) and mutation of pyruvate kinase I (*pykF*) inhibited Embden-Meyerhof-Parnas (EMP) glycolysis [6]. Due to their necessity to our target pathway, *gapB* and *pgk* functionality were left unchanged in the model (except to simulate NOG26), while the other listed enzymes had their flux boundaries set to zero in order to simulate genetic impairment.

Different nutrient and gene knockout combinations were run through the model to examine optimal conditions for the pathway. Taking a suggested glucose feeding rate of -18.5 mmol gDW^{-1} hr^{-1} as a baseline nutrient uptake level, an uptake rate for formate of -10 mmol gDW^{-1} hr^{-1} was assumed [120]. Negative values indicate cellular intake, while positive values indicate export. Aerobic, microaerobic, and anaerobic conditions, with access to glucose and formate or only glucose were examined for each gene knockout scenario (**Figure C3**). Microaerobic conditions were set by limiting the oxygen uptake rate to -10 mmol gDW^{-1} hr^{-1}, which is 1% of the aerobic lower limit [121]. In order to confirm cell viability according to the model, the lower

limit of the growth rate was set to 0.1 hr^{-1}, which equates to a doubling time of approximately 6.9 hours, nearly two times slower than NOG26. While relatively low, growth below the lower limit is still possible. The maximum growth rate of each scenario was also determined.

Plasmid construction

Having done FBA, we identified key genetic modifications that would potentially drive the target pathway. The first round of modifications focused on increasing expression of key genes; for the sake of simplicity, vector overexpression was done instead of genome knock-in. The finished NOG26 strain contains two plasmids overexpressing genes necessary for growth on glucose-only media. The first, pPL274, is 12 kbp long and contains genes involved in NOG and glucose transport/phosphorylation, reducing the strain's reliance on acetate [6]. The second, pTW371, is 7 kbp long and contains additional copies of transketolases I and II (*tkt1* and *tkt2*) and a multiple cloning site (MCS) for additional pathway engineering [6]. All genes on plasmids are IPTG-inducible. Because of the availability of a MCS on pTW371 and the higher likelihood of stable transformation due to its smaller size, it was selected for additional cloning.

The pTWGP plasmid was transformed into the JXG1, JXG3, and JXG5 strains. The starting backbone was generated by pTW371 linearized by PCR. From *B. subtilis* and JCL16 *E. coli* respectively, *gapB* and *pgk* were extracted by PCR, and generated fragments with overhangs were circularized by Gibson Assembly.

The pTWG3PY plasmid was transformed into JXG2, JXG4, and JXG6. In order to assemble pTWG3PY, an intermediate vector, pTWGPPY, was assembled first by Gibson Assembly. The pTWGP plasmid was linearized by NdeI and BglII restriction digestion directly downstream of the *pgk* gene. The *ppsA* and *yfiD* genes were extracted from JCL16 *E. coli* via PCR, with the

fragments and linearized pTWGP plasmid being assembled into pTWGPPY. Upon confirming pTWGPPY assembly, we reused the BglII cut site (which was regenerated in pTWGPPY) to linearize pTWGPPY, and extracted *pfl* from JCL16 *E. coli* via PCR, which was assembled into pTWG3PY by Gibson Assembly.

After isolating produced plasmids, NOG-capable strains were made chemically-competent using Zymo Research Mix and Go! E. coli Transformation kit and transformed with finalized plasmids. Undesired plasmids were removed through multiple generations of growth without the selection pressure of the corresponding antibiotic. Strains without plasmids were confirmed via colony PCR. All overnight cultures were grown at 37^{o}C on liquid and solid LB media with the necessary antibiotics added. Ampicillin was added to a working concentration of 100 µg/mL; kanamycin: 50 µg/mL; spectinomycin: 100 µg/mL.

Scar-less CRISPR-Cas9 genetic knockout

Genetic knockouts were conducted via the two-plasmid system developed by Jiang *et al.* [122]. The first plasmid, the "pTarget" contained the gene-specific guide RNA, which attached to the Cas9 protein encoded by the second plasmid, "pCas". The pCas vector also contained the λ-Red recombinase system, which repaired the double-stranded break performed by Cas9 via homologous recombination [123]. The "donor DNA" consisted of 500 bp of genomic DNA immediately before and after the target gene, replacing the spliced gene for a scar-less knockout. For the purposes of this project, the donor DNA was cloned onto the pTarget vector via Gibson Assembly, due to its smaller size.

No changes were made to pCas, as all gene-specific elements are located on the pTarget plasmid. To alter pTargets for different genes, the PAM sequence was added on primers, with overhangs

binding directly before and after the existing PAM sequence. In order to prevent transformation with the incorrect pTarget, the plasmid was linearized into two fragments through PCR, purified via gel extraction, and then assembled via Gibson Assembly. Next, the donor DNA was extracted by PCR with overhangs added. The assembled pTargets were linearized by restriction digest, treated with phosphatase, and then the donor DNA was added by Gibson Assembly. The pTargets with donor DNA added were named pTargetTs.

To perform the knockouts, each target strain was co-transformed with pCas and the desired pTargetT. Cas9 and recombinase are produced constitutively, while the guide RNA was induced upon adding arabinose (10 mM). All strains were grown at 30^{o}C, in order to retain the temperature-sensitive pCas. Colonies were selected for using selection markers corresponding to both plasmids, and screened for the knockout using colony PCR. As the donor DNA could exist both on the pTarget and the genome, primers landing outside of the donor DNA were designed to confirm the knockouts. For subsequent knockouts, the pCas was retained while the pTargetT was removed via an IPTG-inducible guide RNA located on pCas targeting the pTargetT. After all desired knockouts were performed, the pTargetT was again removed by IPTG-induction, while the pCas was cured from cells by overnight culture at 37^{o}C due to its temperature-sensitive origin of replication.

Media sampling and intracellular metabolite extraction

We first examined media samples of relevant strains under different media conditions. 200 µL was taken for OD600 measurements, while 200 µL was taken for HPLC measurement. HPLC samples were centrifuged at 13,000 RPM for 5 minutes at 4^{o}C, and the supernatant was

transferred to HPLC vials. For anaerobic cultures, media samples were taken using needles through the Hungate tube rubber stopper.

For extraction of intracellular metabolites, cells were grown to exponential phase at 37^{o}C and then extracted. Tubes were removed from the incubator, 3 mL of the culture was transferred to a 0.45 µm PVDF filter paper, and then the filter paper was submerged upside down in 10 cm 6-well plates containing extraction solvent. We utilized an extraction solvent with a volumetric ratio of 40:40:20 Acetonitrile:Methanol:H_2O. This procedure was done as quickly as possible, typically on a time scale between 10-20 seconds. We then incubated samples at -20^{o}C for 20 minutes. The samples were then transferred to pre-chilled 1.5 mL tubes and centrifuged at 13,000 RPM for 5 minutes at 4^{o}C to separate cell debris. The supernatant was transferred to HPLC tubes.

To determine the metabolic fate of formate, 10 mM [U-$^{13}C_1$]formate was fed to cells alongside 4 g/L of unlabeled glucose. Cultures were grown in 15 mL conical tubes with 10 mL working volume with a parafilm cover. After reaching OD600 of 0.1-0.6, cells were extracted using extraction solvent [124].

To measure absolute metabolite concentrations, a reference strain NCM3722 was cultured with 4 g/L [U-$^{13}C_6$]glucose in Gutnick minimal media [3]. After extracting using the same method as the labeled formate experiment, a portion of the extract was mixed with the strain extracts such that the OD*volume ratio was equal for both extracts. As our strain extracts were grown on ^{13}C-formate, which has one labeled carbon, and we had data from cells fed only labeled formate, we determined labeling patterns likely belonging to NCM3722 (typically >M+2, unless metabolites were small) and those belonging to NOG26 and the engineered strains (typically ≤M+2). As we

had already normalized for OD, we used the ratio of NCM3722:Strain labeling to estimate the absolute metabolite concentrations.

LC-MS and HPLC methods

Media samples were measured using an Agilent 1260 Infinity II HPLC system. The mobile phase was 14 mM H_2SO_4, with samples running at 0.7 mL/min for 20 minutes. An injection volume of 10 µL was used. Glucose was measured using the Refractive Index (RI) detector, while organic acids were measured using the UV/VIS detector. Standard curves for glucose, formate, acetate, succinate, lactate, ethanol, and pyruvate were created using six standards at different known concentrations.

To measure intracellular metabolites, extracted samples were measured with the Thermo-Fisher Vanquish UHPLC coupled with the Q-Exactive Orbitrap Mass Spectrometer. The Vanquish system allows for the usage of two columns sequentially, which allowed us to run samples on both the XBridge BEH Amide XP Column 130Å and the SeQuant ZIC-pHILIC column. Solvent A was 95:5 water:acetonitrile mixed with 20 mM ammonium acetate and ammonium hydroxide, while Solvent B was just acetonitrile, adapted from [125]. An increasing gradient of the organic solvent was introduced throughout a 21 minute method, running at 0.15 mL/min, also adapted from [125]. An injection volume of 10 µL was used, with the Q-Exactive mass spectrometer analyzing in the positive and negative mode from m/z of 65-975 with a resolution of 140,000. In order to facilitate efficient running of samples, columns were used in tandem, with sequential samples alternating columns, with the unused column allowed to regenerate and flush out the previous sample. Some controls were run on both columns to correct for column-to-column variation.

To determine absolute metabolite concentrations and isotope labeling fractions, data files were imported to the Metabolic Analysis and Visualization Engine (MAVEN). The mass spectra and chromatographic data were analyzed for ^{13}C labeling, and confirmed against known retention times. Several metabolites such as isoleucine and leucine or ribose-5-phosphate and ribulose-5-phosphate and xylulose-5-phosphate, were unable to be resolved from each other, and were lumped together. All metabolites that fell under narrow m/z windows and were grouped together are listed in the relevant figures.

3.6 Conclusion

While other studies have shown that the reverse *pfl* flux can be significant and can sustain bacterial growth on formate when fed alongside acetate [72], we have tried to facilitate a pathway capable of converting formate to acetate, a primary metabolite that is connected to many useful secondary metabolites. By utilizing non-oxidative glycolysis, we can maximize carbon yield towards acetate [6]. One benefit of our pathway is the fact that the vast majority of machinery is native to most microbes, allowing for increased flexibility as to our chassis of choice. Another benefit is the ability to utilize the GNG-NOG cycle under both microaerobic and anaerobic conditions, while maximizing growth rates under fully aerobic conditions is possible. Finally, NOG efficiently converts carbons to acetate, losing minimal carbons as CO_2. We were able to show that formate was consumed by our engineered strains and was incorporated into the GNG-NOG cycle by the JXG2 strain. This strain was shown to have consistently increasing formate uptake under microaerobic conditions, while also maintaining acetate as the primary export product. In addition, JXG2 was the only strain to have significantly higher labeling of both hexose phosphates and NOG metabolites compared to NOG26 when fed ^{13}C-formate.

This study highlights the utility of *in silico* predictive tools for pathway engineering, as well as the power of stable isotope tracing coupled with high-resolution LC-MS. This included Flux Balance Analysis to confirm pathway viability, as well as determining the favorability of the overall pathway using net metabolite balances. We also used [U-$^{13}C_1$]formate and [U-$^{13}C_6$]glucose to determine metabolic fate of formate and absolute intracellular concentrations respectively. The techniques used in this study are just a small sample of tools available for pathway design, with one notable example being thermodynamic analysis of metabolic reactions to develop robust metabolic control strategies [126].

While we have demonstrated the possibility of this pathway, there is a large amount of optimization yet to be done. We hope to continue engineering this pathway and improving formate-to-acetate conversion, with some of the future work to be done outlined previously. With the current toolbox in synthetic biology and metabolomics continually expanding, we hope to optimize the GNG-NOG cycle so that it can become a viable option towards the production of value-added chemicals from one-carbon compounds, while being relatively facile to implement in other microorganisms.

Appendix A1: Table 1 Abbreviations

Pathways

AA: amino acid synthesis
ED: Entner-Doudoroff
EMP: Embden-Meyerhof-Parnas
FA: fatty acid synthesis or catabolism
Glu: glutamate synthesis
oxPPP: oxidative pentose phosphate pathway
Pta-Acka: phosphotransacetylase-acetate kinase pathway
TCA: tricarboxylic acid cycle

Metabolites

Anteiso-C17:0: 14-methylhexadecanoate
ATP: adenosine triphosphate
BCAA: branched-chain amino acid
BCFA: branched-chain fatty acid
BCKA: branched-chain α-ketoacids
CDP: cytidine diphosphate
CoA: coenzyme A
DHAP: dihydroxyacetone phosphate
DMS: dimethyl-succinate
DNA: deoxyribonucleic acid
dTTP: deoxythymidine triphosphate
FBP: fructose-1,6-bisphosphate
f^5C: 5-formylcytidine
F6P: fructose-6-phosphate
GAP: glyceraldehyde-3-phosphate
GSH: glutathione
G6P: glucose-6-phosphate
Iso-C16:0: 14-methylpentadecanoate
Iso-C17:0: 5-methylhexadecanoate
KDPG: 2-dehydro-3-deoxy-phosphogluconate
MetO: methylene oxide
NADH: nicotinamide adenine dinucleotide
NADPH: nicotinamide adenine dinucleotide phosphate
OAA: oxaloacetateIMP: inosine monophosphate
PA: pantothenate
PEP: phosphoenolpyruvate
PPA: 4-phosphopantothenate
p-τm^5U: 5-taurinomethyluridine 5'-monophoshate
RE-024: fosmetpantotenate
RNA: ribonucleic acid
R5P: ribose-5-phosphate
S7P: sedoheptulose-7-phosphate

THF: tetrahydrofolate
tRNA: transfer ribonucleic acid
UMP: uridine monophosphate
X5P: xylulose -5-phosphate
αKG: α-ketoglutarate
2HG: *(R)*-2-hydroxyglutarate
23BPG: 2,3-bisphosphoglycerate
3-HIB: 3-hydroxy-isobutyrate
3PG: 3-phosphoglycerate
6PG: 6-phosphogluconate

Others

ACL: ATP citrate lyase
BCAT1/BCAT2: *branched chain amino acid transaminase 1/2*
CRISPR: clustered regularly interspaced short palindromic repeats
GDH: glutamate dehydrogenase
iBMK: immortalized baby mouse kidney cells
IDH1/IDH2: isocitrate dehydrogenase 1/2
ME1: malic enzyme 1
mTORC1: mammalian target of rapamycin complex 1
PDH: pyruvate dehydrogenase
PEPCK: phosphoenolpyruvate carboxykinase
PP2A: serine/threonine-protein phosphatase 2A
P4HA: prolyl-4-hydroxylase
SHMT2: serine hydroxymethyltransferase 2
TPI: Triose-phosphate isomerase
UK5099: 2-Cyano-3-(1-phenyl-1H-indol-3-yl)-2-propenoic acid
αKDH: α-ketoglutarate dehydrogenase

Appendix A2: Figure 4 and Figure 6 Abbreviations

Metabolites

F6P: fructose-6-phosphate
FBP: fructose-1,6-bisphosphate
GAP: glyceraldehyde-3-phosphate
DHAP: dihydroxyacetone phosphate
13BPG: 1,3-bisphosphoglycerate
3PG: 3-phosphoglycerate
2PG: 2-phosphoglycerate
PEP: phosphoenolpyruvate
PYR: pyruvate
AcCoA: acetyl-CoA
FOR: formate
AcP: acetylphosphate
Ac: acetate
S7P: sedoheptulose-7-phosphate
E4P: erythrose-4-phosphate
R5P: ribose-5-phosphate
Ru5P: ribulose-5-phosphate
Xu5P: xylulose-5-phosphate

Enzymes

PTS: phosphotransferase system
GALP: galactose permease
GLF: UDP-galactopyranose mutase
fbp: fructose-1,6-bisphosphatase
glpX: fructose-1,6-bisphosphatase class 2
pfkA/B: phosphofructokinase A/B
fba: fructose-1,6-bisphosphate aldolase
tpi: triosephosphate isomerase
gapA/B: glyceraldehyde-3-phosphate dehydrogenase A/B
pgk: phosphoglycerate kinase
pgm: phosphoglycerate mutase
eno: enolase
ppsA: phosphoenolpyruvate synthetase
pykA/F: pyruvate kinase A/F
pfl: pyruvate formate lyase
PDH: pyruvate dehydrogenase
yfiD: autonomous glycyl radical cofactor
arcA: aerobic respiration control protein
FDH: formate dehydrogenase
pta: phosphate acetyltransferase
acka: acetate kinase
tal: transaldolase

tkt1/2: transketolase 1/2
rpi: ribose-5-phosphate isomerase
rpe: ribulose-5-phosphate epimerase
xpk: phosphoketolase.

Appendix B1: Materials used

JCL16 and NOG26 strains were provided in cooperation with the Liao research group at UCLA. The JCL16 wild-type strain is derived from the BW25113 *E. coli* strain. The *B. subtilis* st. 168 used to extract *gapB* was donated by the Tang research group at UCLA. The pTarget and pCas plasmids were also provided in cooperation with the Liao research group at UCLA.

Assembly of all plasmids was done via Gibson Assembly using New England Biolabs' NEBuilder HiFi DNA Assembly Master Mix. Prior to Gibson Assembly, fragments were amplified using PCR and purified via New England Biolabs' Gel Extraction Kit. Assemblies were first transformed into XL1-Blue competent cells, and then extracted using New England Biolabs' Monarch Plasmid Miniprep Kit. Assembled plasmids were confirmed via colony PCR, restriction digestion, and sequencing. All used primers can be found in Appendix B2. Confirmation of assembled plasmids can be found in Appendix C1.

LB and M9 minimal media bases and spectinomycin were from Sigma-Aldrich. Kanamycin, ampicillin, carbenicillin, arabinose, IPTG, filter papers, extraction solvents, and culture tubes were from Fisher Scientific. LB and minimal media plates were from Teknova. All PCR and cloning materials, including restriction enzymes, master mixes, DNA ladders, etc., were from New England Biolabs. The Vitamin Supplement was from the American Type Culture Collection, and is composed of 2 mg/L folic acid, 10 mg/L pyridoxine hydrochloride, 5 mg/L riboflavin, 2 mg/L biotin, 5 mg/L thiamine, 5 mg/L nicotinic acid, 5 mg/L calcium pantothenate, 0.1 mg/L Vitamin B12, 5 mg/L p-aminobenzoic acid, 5 mg/L thioctic acid, and 900 mg/L monopotassium phosphate. [U-$^{13}C_1$]formate and [U-$^{13}C_6$]glucose were from Cambridge Isotope Laboratories. All DNA sequencing was done through Laragen. All primers were synthesized by Integrated DNA Technologies.

Appendix B2: Primer Sequences

Table B1: Table of primer sequences and their uses

Primer	Sequence	Use
Primers to construct pTWGP		
pTW371Frag1F	TTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTC	Extracts 1/3 of pTW371 for Gibson Assembly of pTWGP, Forward
pTW371Frag1R	ATGGTATATCTCCTGCATGCTCAAAGCAGAGCTTCGACGT	Extracts 1/3 of pTW371 for Gibson Assembly of pTWGP, Reverse
pWGPF2F	GTCGTCAAACACGTCGAAGCTCTGCTTTGAGCATGCAGGAGATATACCATGTCTAAATTT	Extracts 1/3 of pTW371 for Gibson Assembly of pTWGP with overhang for gapB_BS insert, Forward
pTW371Frag2R	ATCGCTACTTTTACCTTCATTACTATATCTCCTTCTTATACTTAA	Extracts 1/3 of pTW371 for Gibson Assembly of pTWGP, Reverse
pTW371Frag3F	AAGAGCGCGCTAAGAAGTAACATATGGCAGATCTCAATTG	Extracts 1/3 of pTW371 for Gibson Assembly of pTWGP, Forward
pTW371Frag3R	GACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAA	Extracts 1/3 of pTW371 for Gibson Assembly of pTWGP with overhang for pgk, Reverse
gapBFragF	TAGTTAAGTATAAGAAGGAGATATAGTAATGAAGGTAAAAGTAGCGAT	Extracts gapB (B. subtilis) for Gibson Assembly on pTW371 with NdeI and BglII cuts to make pTWGP, Forward
gapBFragR	GTCATCTTAATTACAGACATGGTATATCTCCTGCATGCTTATACAGCAGACGGATGTT	Extracts gapB (B. subtilis) for Gibson Assembly on pTW371 with NdeI and BglII cuts to make pTWGP, Reverse
pgkFragF	AACATCCGTCTGCTGTATAAGCATGCAGGAGATATACCATGTCTGTAATTAAGATGAC	Extracts pgk for Gibson Assembly on pTW371 with NdeI and BglII cuts to make pTWGP, Forward
pgkFragR	CAATTGAGATCTGCCATATGTTACTTCTTAGCGCGCTCTT	Extracts pgk for Gibson Assembly on pTW371 with NdeI and BglII cuts to make pTWGP, Reverse
Primers to construct pTWGPPY		
pgkv3R	TGACGAGCCATTGTTGGACATGGGATCCTCCTTTAAATATGTTACTTCTTAGCGCGCTCT	New primer to add overhang for ppsA for Gibson Assembly on pTWGP to replace to make pTWGPPY, Reverse primer on pgk
ppsAFtrunc	*acttacagca*tatttaaaggaggatcccATGTCCAACAATGGCTCGTCACCG	Extracts ppsA for Gibson Assembly on pTWGPPY, Forward
ppsARtrunc	*tcattgcaat*GGTATATCTCCTGCATGCTTATTTCTTCAGTTCAGCCAGGCTT	Extracts ppsA for Gibson Assembly on pTWGPPY, Reverse
yfiDv3F	CTGGCTGAACTGAAGAAATAAGCATGCAGGAGATATACCATGATTACAGGTATCCAGATT	New primer to add overhang for ppsA for Gibson Assembly on pTWGPPY, Forward primer on yfiD
yfiDFragR	GCCGATATCCAATTGAGATCTTACAGGCTTTCAGTAAAGG	New primer to add overhang for ppsA for Gibson Assembly on pTWGPPY, Reverse primer on yfiD

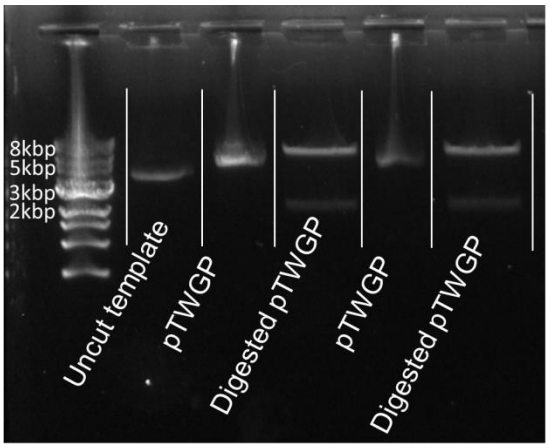
Primers to construct pTWG3PY		
yfiDgib2F	ACCTTCACTCAATCTATGTAAGCATGCAGG AGATATACCATGATTACAGGTATCCAGATT	Extracts yfiD for Gibson Assembly on pTWGP and pTWGPPY to construct pTWG2PY and pTWG3PY, Forward
yfiDgib2R	GCCGATATCCAATTGAGATCTTACAGGCTT TCAGTAAAGGTACGA	Extracts yfiD for Gibson Assembly on pTWGP and pTWGPPY to construct pTWG2PY and pTWG3PY, Reverse
pflB4gib	ACCTTTACTGAAAGCCTGTAAGCATGCAGG AGATATACCATGTCCGAGCTTAATGAAAAG	Extracts pfl for Gibson Assembly on pTWGPPY to construct pTWG3PY, Forward
pflAFgib	GGCCGGCCGATATCCAATTGAGATCTTACA TAGATTGAGTGAAGGTACGA	Extracts pfl for Gibson Assembly on pTWGPPY to construct pTWG3PY, Reverse
Primers to construct pTargetTs		
pTg1R	TCGCGGAACTTCGGAATAGGAACTTCATGA GCTCAGCCAATCGACTGGCGAGCGGCATCT	Extracts 1/2 of pTargets, pairs with 1F primers
pTg2F	AGATGCCGCTCGCCAGTCGATTGGCTGAGC TCATGAAGTTCCTATTCCGAAGTTCCGCGA	Extracts 1/2 of pTargets, pairs with 2R primers
pTgarcA2R	GCTATTTCTAGCTCTAAAACAGGCAAAATT CAGTCCCGTGACTAGTATTATACCTAGGAC	Extracts 1/2 of pTargetFAA, with arcA gRNA as part of overhang, pairs with pTg2F, Reverse
pTgarcA1F	GTCCTAGGTATAATACTAGTCACGGGACTG AATTTTGCCTGTTTTAGAGCTAGAAATAGC	Extracts 1/2 of pTargetFAA, with arcA gRNA as part of overhang, pairs with pTG1R, Forward
arcAafF	GTTTTTTTTGACGGTGGTAAAGCCGAGTTTG CTACCTAAATTGCCAACTAA	Extracts 500 bp downstream of arcA gene for Gibson Assembly to form pTargetTAA, Forward
arcAafRgb	CTCGAGTAGGGATAACAGGGTAATAGCCCT TTACTTAGCTTAAAAAAGGC	Extracts 500 bp downstream of arcA gene for Gibson Assembly to form pTargetTAA, Reverse
arcAb4Fgb	GAGTCGACCTGCAGAAGCTTAGATCACTAA GTACGCCATTCTGCTGATTG	Replacement for arcAb4F, Extracts 500 bp upstream of arcA gene for Gibson Assembly to form pTargetTAA, Forward
arcAb4R	TTAGTTGGCAATTTAGGTAGCAAACTCGGC TTTACCACCGTCAAAAAAAAC	Extracts 500 bp upstream of arcA gene for Gibson Assembly to form pTargetTAA, Reverse
pTargTpykA1NF	GTCCTAGGTATAATACTAGTATACACGGAT TTTGGGCCCCGTTTTAGAGCTAGAAATAGC	Extracts 1/2 of pTargetTPYK, with pykA gRNA as part of overhang, pairs with pTg1R, Forward
pTargTpykA2NR	GCTATTTCTAGCTCTAAAACGGGGCCCAAA ATCCGTGTATACTAGTATTATACCTAGGAC	Extracts 1/2 of pTargetTPYK, with pykA gRNA as part of overhang, pairs with pTg2F, Reverse
pykAb4F	GAGTCGACCTGCAGAAGCTTAGATCTGTAC TCCGATGATATCGTGCTGCA	Extracts 500 bp upstream of pykA gene for Gibson Assembly of pTargetTPC, Forward
pykAb4R	GTTTTCGCCGCATCCGGCAACGTACGTAAT ACTCCGTTGACTGAAACAAC	Extracts 500 bp upstream of pykA gene for Gibson Assembly of pTargetTPC, Reverse

pykAafF	GTTGTTTCAGTCAACGGAGTATTACGTACG TTGCCGGATGCGGCGAAAAC	Extracts 500 bp downstream of pykA gene for Gibson Assembly of pTargetTPC, Forward
pykAafR	CTCGAGTAGGGATAACAGGGTAATACATTC ATCCAGTCGGTACGTCAGGG	Extracts 500 bp downstream of pykA gene for Gibson Assembly of pTargetTPC, Forward
pTgaceE2R	GCTATTTCTAGCTCTAAAACAGCACAG ATCGCCAATACGCACTAGTATTATACC TAGGAC	Extracts 1/2 of pTargetFAE, with aceE gRNA as part of overhang, pairs with pTg2F, Reverse
pTgaceE1F	GTCCTAGGTATAATACTAGTGCGTATT GGCGATCTGTGCTGTTTTAGAGCTAGA AATAGC	Extracts 1/2 of pTargetFAE, with aceE gRNA as part of overhang, pairs with pTG1R, Forward
aceEafF	TTATTAGATAGATAAGGAATAACCCGA GGTAAAAGAATAATGGCTATCGA	Extracts 500 bp downstream of aceE gene for Gibson Assembly to form pTargetTAE, Forward
aceEafRgb	CTCGAGTAGGGATAACAGGGTAATATA GACACTTTGTCACCCACG	Extracts 500 bp downstream of aceE gene for Gibson Assembly to form pTargetTAE, Reverse
aceEb4Fgb	GAGTCGACCTGCAGAAGCTTAGATCTA GAGCTGGCGCAGCAGTCTGGCG	Extracts 500 bp upstream of aceE gene for Gibson Assembly to form pTargetTAE, Forward
aceEb4R	TCGATAGCCATTATTCTTTTACCTCGGG TTATTCCTTATCTATCTAATAA	Extracts 500 bp upstream of aceE gene for Gibson Assembly to form pTargetTAE, Reverse
Primers for confirmation via colony PCR or sequencing		
PTW371AF	GCGGATAACAAGATACTGAG	Sequences tkt2 of pTW371, Forward
PTW371AR	GACTTAAGCATTATGCGGCC	Sequences tkt2 of pTW371, Reverse
PPL274AF	GATACTGAGCACATCAGCAG	Sequences xpk of pPL274, Forward
PPL274AR	CTAGTCGCGTGACTAGAC	Sequences xpk of pPL274, Reverse
pTargSeqF	CCGCATATGCTGGATCCTTG	Sequences 20bp gRNA sequence, Forward
pTargSeqR	AACTCGAGTAGGGATAACAG	Sequences 20bp gRNA sequence, Reverse
pCasSeqF	AGACATTGCCGTCACTGCGTCTTTT	Sequences Recombinase on pCas, Forward
pCasSeqR	TCATCGCCATTGCTCCCCAAATACA	Sequences Recombinase on pCas, Reverse
specseqF	AGTATCGACTCAACTATCAG	Sequences spectinomycin resistance gene on pTargets, Forward
specseqR	TGGACAAATTCTTCCAACTG	Sequences spectinomycin resistance gene on pTargets, Reverse
arcAseqF	GAACCTGAAGCTGAATATGG	Sequences arcA on genome, Forward
arcAseqR	GTGTAACGATTCAGCCAATG	Sequences arcA on genome, Reverse
pykAseq15F	CGAATGGCACTCCCTATGTTAATCG	Sequences pykA on genome, Forward
pykAseq15R	AGCGATGGCAATGATGGCTGAGTTG	Sequences pykA on genome, Reverse
aceEseq15F	CCTCCTTGCGTGAGGCGATTCAACGTCTCG	Sequences aceE on genome, Forward
aceEseq15R	CTGCAAACGGCGCCGGAACTTCCATAGAAG	Sequences aceE on genome, Reverse
PPL274BF	ATGCCTTCGCGCCGAGACTT	Sequences tkt2 on pPL274, Forward
PPL274BR	TCAAAGCAGAGCTTCGACGTGTTTG	Sequences tkt2 on pPL274, Reverse
fdhNseqF	ATGGACGTCAGTCGCAGACAATTTTTTAAA	Sequences fdhN on genome, Forward

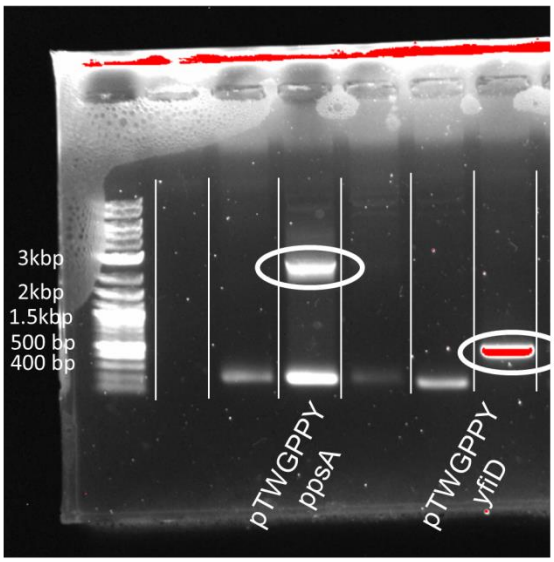
	ATCTG	
fdhNseqR	TTATATCCCTTCTTCACTCTCTTTTTTCGC	Sequences fdhN on genome, Reverse
fdhOseqF	TCATTCAGCTTTCTTTTCCGTTGTCTTGCG	Sequences fdhO on genome, Forward
fdhOseqR	ATGCAGGTCAGCAGAAGGCAGTTCTTTAAG	Sequences fdhO on genome, Reverse
PTWseqF	ACGGCCGCATAATCGAAATTAATACG	Sequences outside MCS of pTW371, to confirm entire insert, Forward
PTWseqR	GTGGCAGCAGCCTAGGTTAATTAAG	Sequences outside MCS of pTW371, to confirm entire insert, Reverse
fdhFB4seq	ATCTTCCGAGGCGTTTTCTAACAG	Sequences fdhF on genome, can be paired with fdhFB4seq, fdhFB4R and fdhFAFR, Forward
fdhFAFseq	CTTGCCGAAGAGATGTCATTACTG	Sequences fdhF on genome, can be paired with fdhFB4seq. Reverse
fdhFgeneF	TTACGCCAGTGCCGCTTCGC	Sequences fdhF on genome, starting on gene, Forward
fdhFgeneR	ATGAAAAAAGTCGTCACGGTTTGCCC	Sequences fdhF on genome, starting on gene, Reverse
gapAseqF	TCGACAAACGCTGGTATACCTCAGG	Sequences gapA on genome, Forward
gapAseqR	ATATCGCCCATGCTAATTGAAAGCG	Sequences gapA on genome, Reverse
pgkseqF	GTAATCAACGTCTTCCGGCTGGGTG	Sequences pgk on genome, Forward
pgkseqR	CATGAACGTTCACTGCAATCGCTC	Sequences pgk on genome, Reverse

Appendix C: Supplementary Figures

pTWGP



pTWGPPY



pTWG3PY

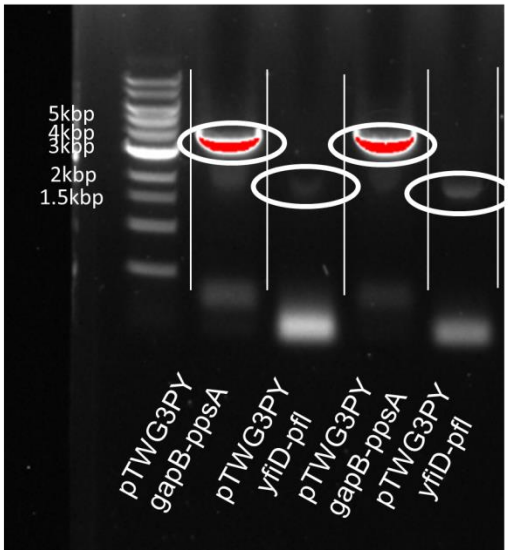


Figure C1 Confirmation of assembled plasmids by gel electrophoresis

Gel images confirming construction of pTWGP, pTWGPPY, and pTWG3PY. For pTWGP, uncut pTW371 (lane 2) was compared against uncut pTWGP (lanes 3 and 5), and pTWGP digested by NotI and NdeI (lanes 4 and 6), which should produce fragments of 7.7 kbp and 2.2 kbp. Uncut vectors traveled further than expected due to supercoiling. For pTWGPPY, *ppsA* and *yfiD* genes were amplified from isolated plasmids via PCR, producing expected bands of 2.4 kbp and 400 bp in lanes 4 and 7 respectively. For pTWG3PY, we confirmed the insertion of *gapB* from *B. subtilis*, *pgk*, and *ppsA* (lanes 2 and 4) as well as the insertion of *yfiD* and *pfl* (lanes 3 and 5) via PCR on isolated plasmids. All plasmids were also confirmed via Sanger sequencing.

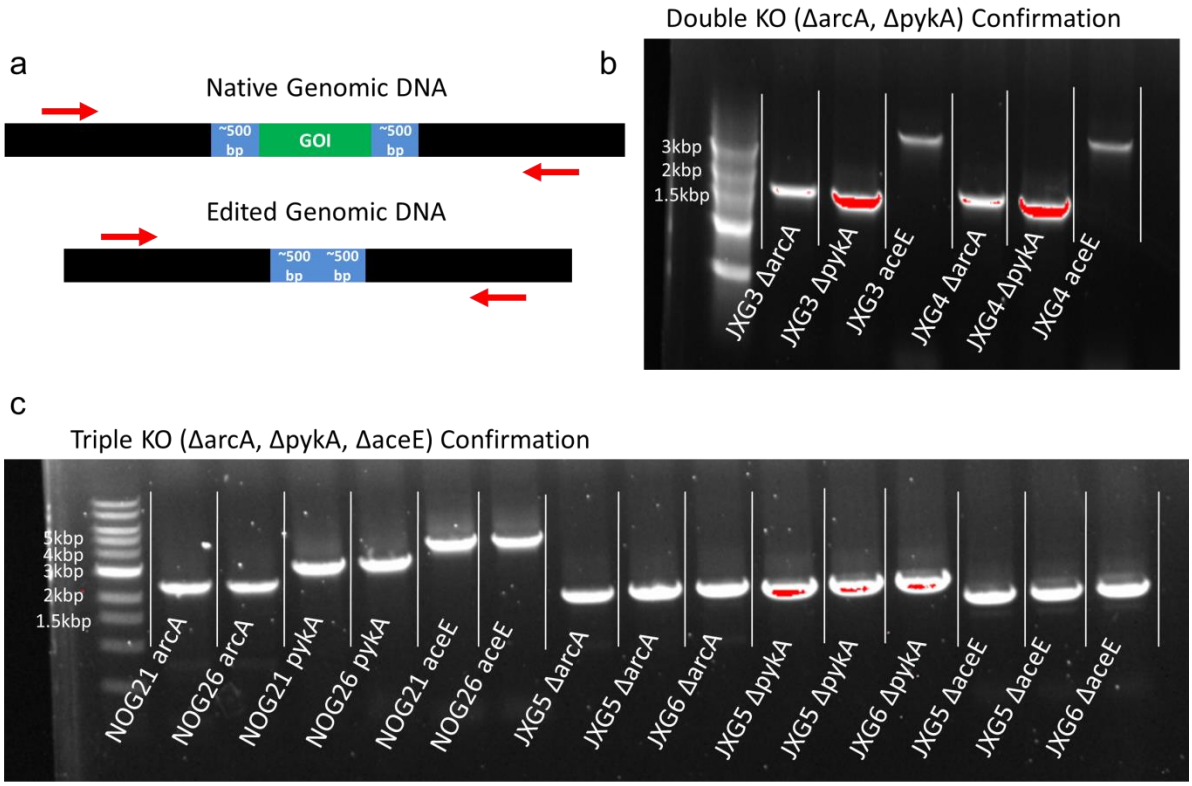


Figure C2 Confirmation of *arcA*, *pykA*, and *aceE* knockouts in JXG3, JXG4, JXG5, and JXG6 via gel electrophoresis

(a) Graphic detailing design of primers to confirm genetic knockouts. Due to the fact that 500 bp flanking the genes of interest were included on pTargetTs, we designed primers to bind 800 bp outside the gene of interest, which should produce a band of 1.6 kbp upon successful knockout. **(b)** Confirmation of double knockout in JXG3 and JXG4. For primers flanking *arcA* and *pykA*, we saw bands slightly above the 1.5 kbp ladder band, indicating successful knockout. Retention of *arcA* and *pykA* would have produced bands around 2.4 and 3 kbp. The *aceE* gene remained intact, as shown in lanes 4 and 7. **(c)** Confirmation of triple genetic knockouts in JXG5 and JXG6. The intact *arcA*, *pykA*, and *aceE* from NOG26 and its precursor NOG21 were shown in lanes 2-7. Two JXG5 clones and one JXG6 clone were tested via PCR, with each of the three target genes producing bands around 1.6 kbp.

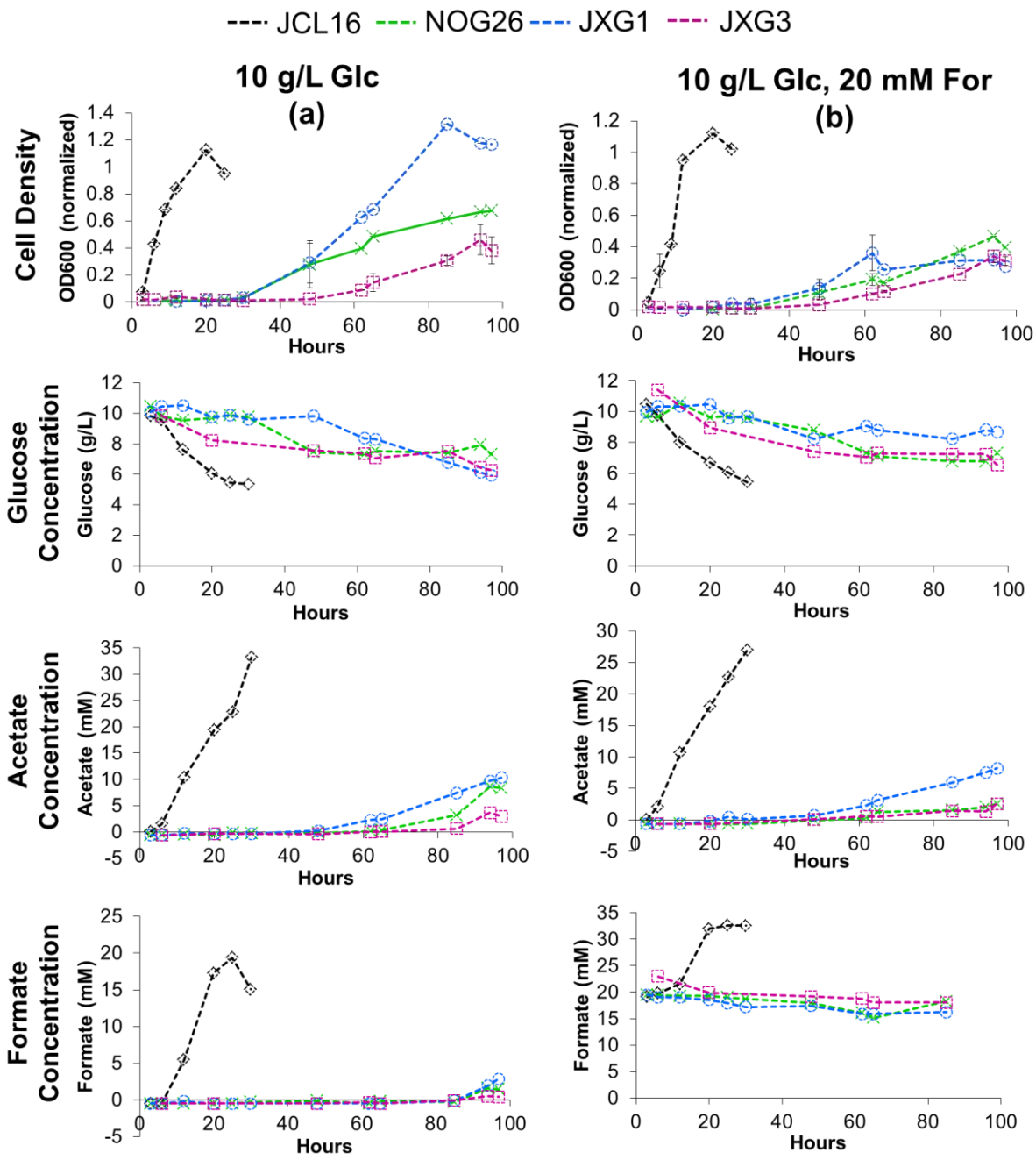


Figure C3 Aerobic growth suggests similarly low levels of formate consumption and acetate production in engineered strains when fed 10 g/L glucose

(a) OD600, glucose concentration (g/L), acetate concentration (mM), and formate concentration (mM) in media over time of JCL16, a wild-type *E. coli* strain, NOG26, JXG1, and JXG3, when fed 10 g/L glucose in M9 minimal media under aerobic conditions. OD600 values were normalized by blank and pathway correction. **(b)** OD600, and glucose, acetate, and formate concentrations in media over time of JCL16, NO26, 26GP, and 26APGP when fed 10 g/L glucose, 20 mM formate in M9 minimal media under aerobic conditions.

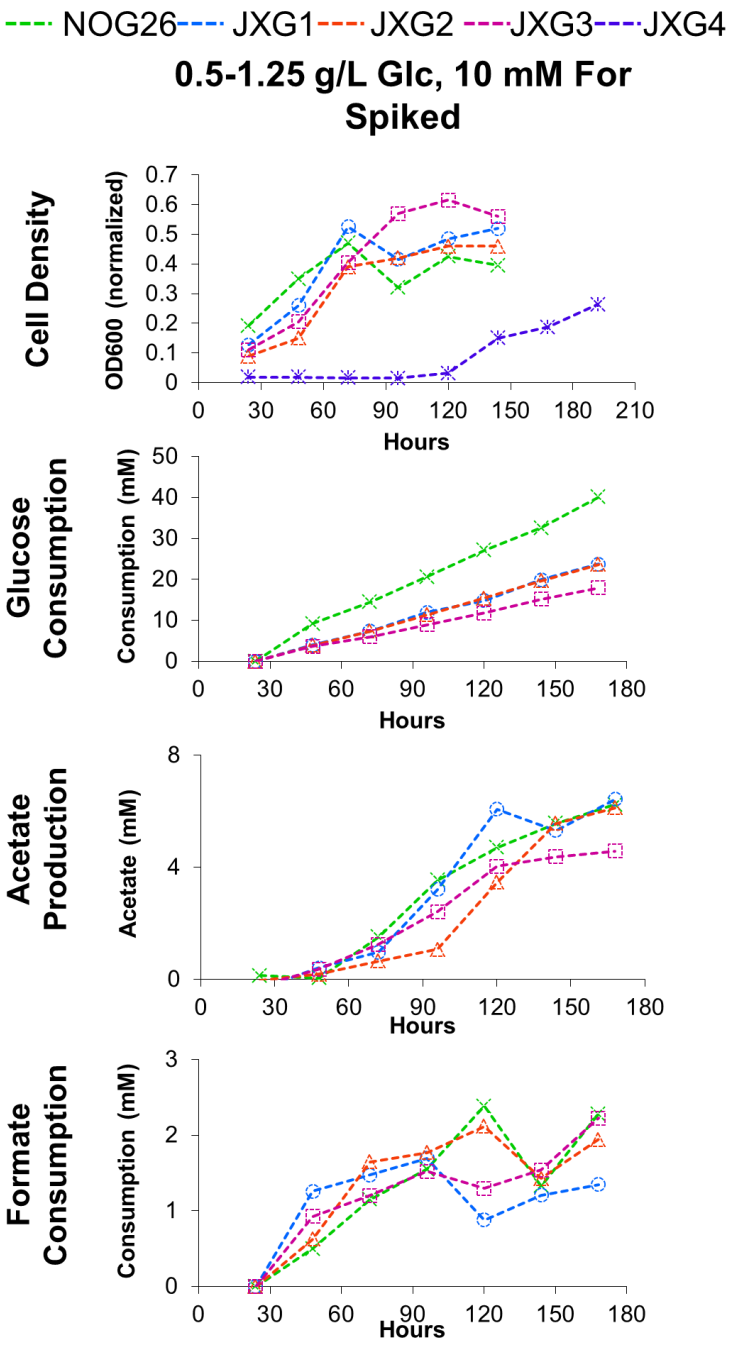


Figure C4 Feeding glucose and formate reflective of daily measured consumption rates decreased acetate production but did not significantly increase formate consumption

OD600, glucose consumed (mM), acetate produced (mM), and formate consumed (mM) in media over time of NOG26, JXG1, JXG2, JXG3, and JXG4, when fed 0.5-1.25 g/L glucose, 10 mM formate in M9 minimal media with 0.5-1.25 g/L glucose and 0.75 mM formate spiked in daily. Consumption was again normalized by a correction factor, but the daily spiked amount was added to the total consumption.

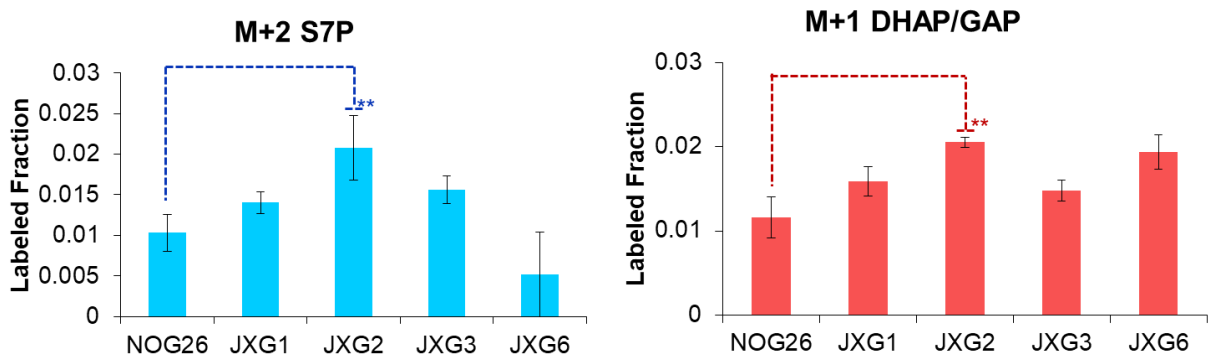


Figure C5 M+2 S7P and M+1 DHAP/GAP support the turning of the GNG-NOG cycle by JXG2

M+2 sedoheptulose-7-phosphate (S7P) and M+1 dihydroxy-acetone-phosphate (DHAP) and glyceraldehyde-3-phosphate (GAP) in NOG26, JXG1, JXG2, JXG3, and JXG6, with the engineered strain compared against NOG26. Cells were cultured in 4 g/L mM glucose, 10 mM [U-$^{13}C_1$]formate, and extracted from OD600 0.13 to 0.62. All samples were run in triplicate, and error bars are standard errors of the mean. Two-sided student's t-tests were run to confirm significant variation between NOG26 and engineered strain, with dashed lines and ** denoting $p \approx 0.1$.

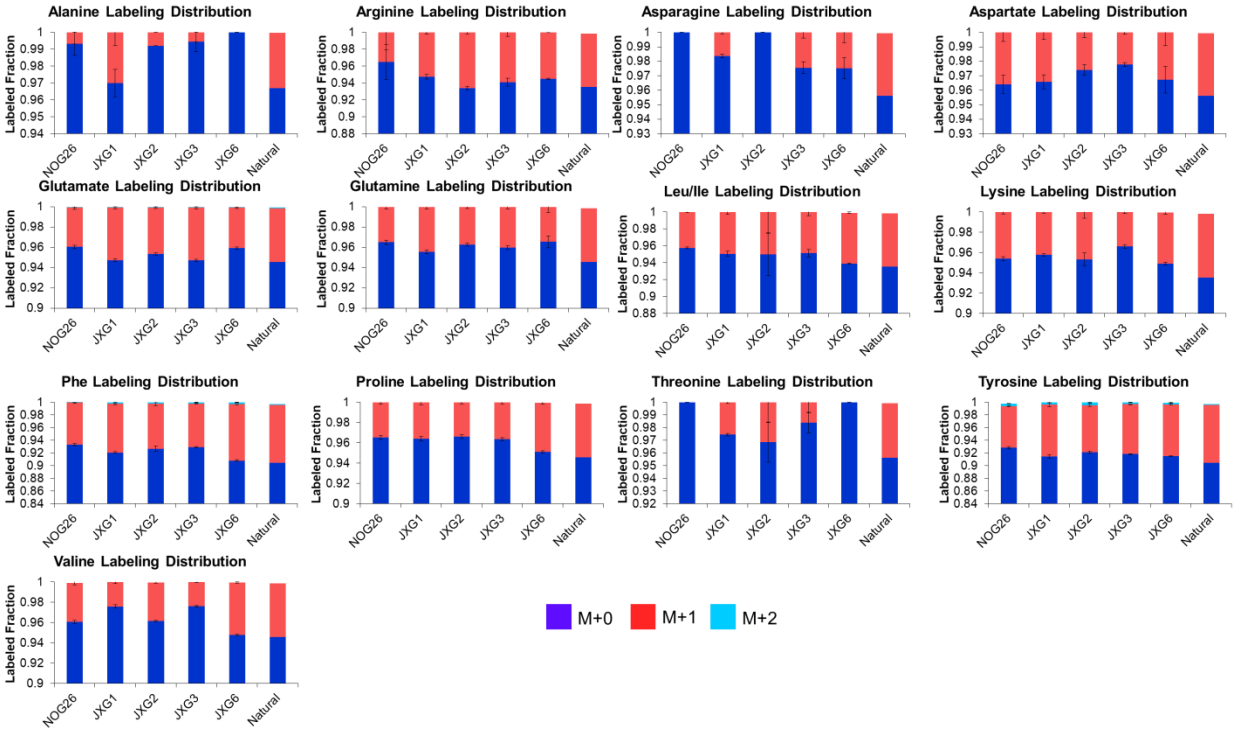


Figure C6 Isotope labeling fractions of all amino acids after feeding [U-$^{13}C_1$]formate

M+0, M+1, and M+2 fractions of all detectable amino acids in NOG26, JXG1, JXG2, JXG3, and JXG6, as well as the expected natural abundance fractions. Cells were cultured in 4 g/L mM glucose, 10 mM [U-$^{13}C_1$]formate, and extracted from OD600 0.13 to 0.62. All samples were run in triplicate, and error bars are standard errors of the mean.

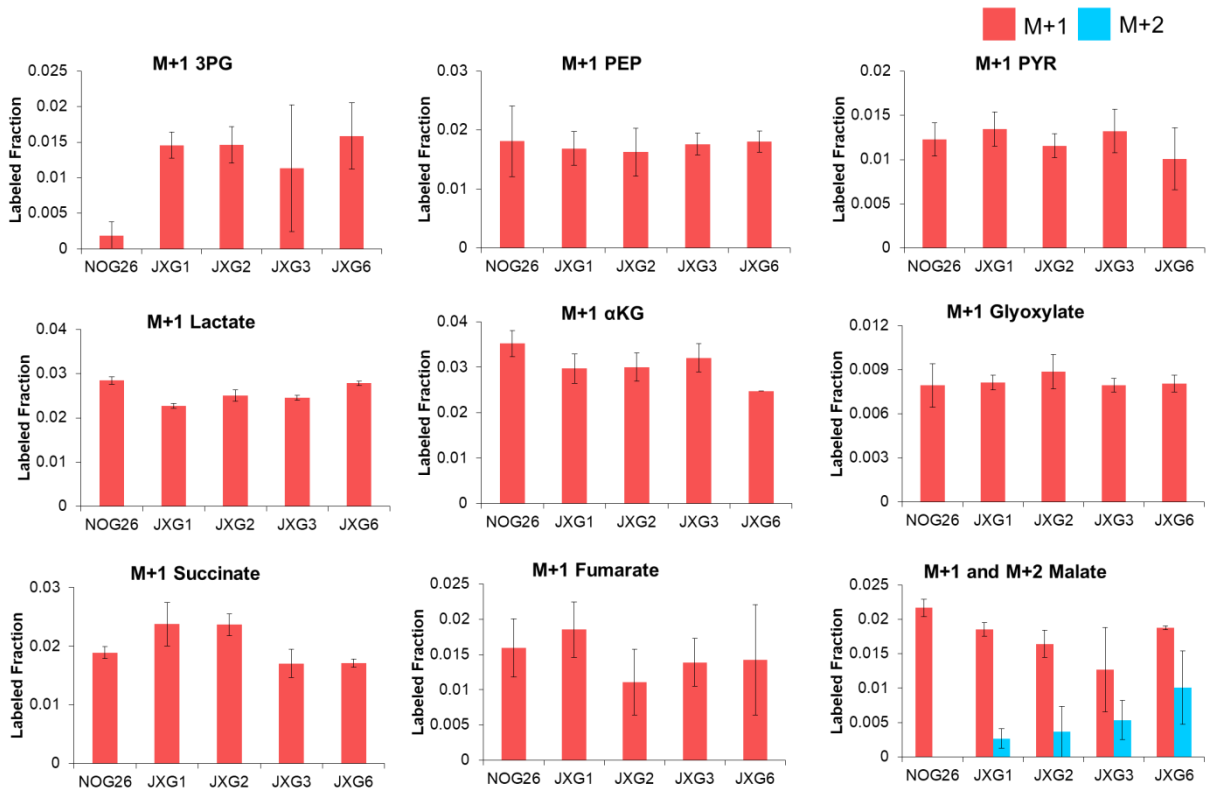


Figure C7 Isotope labeling fractions in central carbon metabolites with negligible differences between strains

M+1 and M+2 fractions of 3-phosphoglycerate (3PG), phosphoenolpyruvate (PEP), pyruvate (PYR), lactate, α-ketoglutarate (αKG), glyoxylate, succinate, fumarate, and malate in NOG26, JXG1, JXG2, JXG3, and JXG6. Cells were cultured in 4 g/L mM glucose, 10 mM [U-$^{13}C_1$]formate, and extracted from OD600 0.13 to 0.62. All samples were run in triplicate, and error bars are standard errors of the mean. Two-sided student's t-tests were run to determine significant variation between NOG26 and engineered strain, but all metabolites above showed insignificant variation.

Appendix D: Absolute Metabolite Concentrations

Table D1: Table of absolute metabolite concentrations and standard errors by strain

Strain	NCM3722 (Ref. Strain)	NOG26		JXGN1		JXGN2		JXGN3	
Metabolite Name	Concentration(M)	Concentration(M)	Standard Error	Concentration(M)	Standard Error	Concentration(M)	Standard Error	Concentration(M)	Standard Error
1,3-BPG+2,3-BPG	1.63E-03	1.10E-02	2.14E-03	1.66E-02	9.83E-04	1.58E-02	1.05E-03	1.77E-02	5.76E-04
2-PG+3-PG	4.28E-04	8.68E-05	5.32E-04	6.18E-05	4.11E-04	2.74E-05	4.73E-04	1.16E-05	1.88E-03
6-phosphogluconate	3.77E-03	2.72E-04	4.11E-05	4.76E-04	5.62E-05	1.35E-04	1.87E-06	1.52E-04	7.04E-06
acetyl-CoA	6.06E-04	1.44E-04	2.23E-05	4.96E-05	3.41E-06	6.34E-04	3.13E-05	5.09E-04	4.25E-05
acetylphosphate	1.07E-03	4.81E-04	4.31E-05	1.30E-03	2.97E-04	1.09E-03	1.01E-04	1.26E-03	4.65E-05
aconitate	1.61E-05	2.91E-05	2.75E-06	1.05E-04	1.17E-05	3.43E-05	6.35E-06	2.73E-05	4.95E-06
adenine	1.47E-06	6.39E-07	1.88E-07	1.21E-06	2.87E-07	1.17E-06	7.28E-08	1.10E-06	3.89E-08
adenosine	1.31E-07	1.40E-07	1.40E-09	1.45E-07	1.40E-08	4.57E-08	1.51E-08	4.78E-08	1.02E-08
ADP	5.55E-04	2.95E-04	7.27E-05	4.55E-04	5.73E-05	3.42E-04	1.58E-05	3.39E-04	4.97E-06
a-ketoglutarate	4.43E-04	1.97E-04	3.77E-05	1.96E-04	2.99E-05	4.66E-04	1.61E-05	3.76E-04	1.23E-05
alanine	2.55E-03	9.34E-04	1.98E-04	1.72E-03	1.08E-04	1.58E-03	1.16E-04	1.63E-03	2.09E-04
AMP	2.81E-04	5.13E-04	1.17E-04	1.05E-03	5.92E-05	3.03E-04	4.37E-05	3.42E-04	3.41E-05
arginine	5.69E-04	2.26E-04	2.88E-05	5.53E-04	1.71E-04	4.26E-04	1.99E-05	3.76E-04	2.95E-05
asparagine	5.11E-04	9.56E-05	2.44E-06	8.72E-05	1.04E-06	2.55E-05	1.03E-05	3.86E-05	1.83E-06
aspartate	4.23E-03	3.89E-03	9.45E-04	1.32E-02	3.21E-03	4.30E-03	5.19E-04	3.55E-03	2.44E-04
ATP	9.63E-03	2.36E-03	5.28E-04	3.15E-03	5.20E-04	4.43E-03	3.95E-04	3.82E-03	1.46E-04
citrate+isocitrate	1.96E-03	2.94E-03	2.66E-04	6.63E-03	1.11E-03	5.98E-03	4.00E-04	3.89E-03	2.15E-04
citrulline	1.35E-03	1.32E-04	1.68E-05	4.71E-04	2.66E-04	2.35E-04	2.50E-05	2.64E-04	8.04E-06
CMP	3.60E-04	1.43E-03	6.14E-04	3.89E-03	1.37E-03	2.94E-04	1.02E-05	7.84E-04	5.12E-04
CoA	1.37E-03	1.65E-03	4.80E-04	2.18E-03	3.21E-04	2.02E-03	9.33E-05	2.39E-03	1.76E-04
CTP	2.73E-03	3.79E-04	5.49E-05	7.84E-04	1.26E-04	6.49E-04	7.07E-05	6.16E-04	5.82E-05
cyclic-AMP	3.52E-05	8.05E-05	1.75E-05	4.02E-05	2.23E-06	2.62E-04	2.24E-04	3.25E-05	1.19E-05
dATP	1.55E-05	1.93E-06	1.20E-06	7.94E-06	1.47E-06	2.11E-06	9.99E-07	4.58E-06	2.55E-06
dCTP	3.45E-05	2.58E-06	6.33E-07	5.86E-06	4.54E-07	3.97E-06	8.71E-07	7.20E-06	1.11E-07
deoxyribose-5-phosphate	3.03E-04	2.51E-05	8.41E-06	5.66E-05	5.46E-06	2.68E-05	4.78E-06	3.82E-05	4.00E-06
DHAP+GAP	3.33E-03	3.06E-03	8.79E-04	5.92E-03	1.06E-03	4.19E-03	4.52E-05	3.82E-03	2.35E-04
dTDP	3.78E-04	2.87E-04	7.05E-05	3.08E-04	3.38E-05	4.40E-04	3.13E-05	3.31E-04	1.11E-05
dTTP	4.62E-03	1.39E-03	3.37E-04	1.85E-03	2.27E-04	2.64E-03	1.50E-04	2.50E-03	1.10E-04
FAD	1.73E-04	1.02E-04	3.32E-05	1.70E-04	2.55E-05	3.44E-04	1.99E-04	2.12E-04	3.17E-05
FMN	5.37E-05	1.71E-04	7.28E-06	1.27E-04	2.36E-05	1.95E-04	1.61E-05	2.15E-04	3.19E-05
fructose-1,6-bisphosphate	1.52E-02	1.10E-02	2.14E-03	1.66E-02	9.83E-04	1.58E-02	1.05E-03	1.77E-02	5.76E-04
F6P/G6P	1.04E-02	1.23E-02	2.68E-03	1.70E-02	1.86E-03	1.65E-02	1.14E-03	1.92E-02	1.03E-03
fumarate	2.88E-04	2.99E-05	1.04E-05	5.32E-05	6.72E-06	8.04E-05	9.40E-06	4.76E-05	9.72E-06
GDP	6.76E-04	9.74E-05	1.67E-05	1.71E-04	3.46E-05	1.02E-04	1.53E-05	1.04E-04	5.59E-06
gluconate	4.16E-05	7.85E-04	8.71E-05	8.71E-04	3.57E-04	8.16E-04	4.72E-05	6.07E-04	2.37E-04
glucosamine-6-phosphate	1.15E-03	1.12E-03	2.66E-04	1.58E-03	4.70E-05	1.49E-03	2.43E-04	1.70E-03	1.42E-05
glutamate	9.60E-02	2.63E-02	5.54E-03	2.68E-02	2.64E-03	4.93E-02	2.77E-03	5.05E-02	2.42E-03
glutamine	3.81E-03	1.84E-03	4.18E-04	2.44E-03	3.00E-04	2.95E-03	9.51E-05	2.91E-03	4.77E-05
glutathione	1.66E-02	1.42E-02	3.34E-03	1.74E-02	2.45E-03	1.82E-02	6.76E-04	1.72E-02	1.00E-03
glutathione disulfide	2.37E-03	6.84E-04	2.02E-04	1.03E-03	7.25E-05	1.09E-03	1.33E-04	1.07E-03	4.92E-05
glyoxylate*	3.80E-06	3.93E-06	3.38E-07	2.34E-05	4.74E-06	2.28E-06	1.76E-07	3.44E-06	1.97E-07
GMP	2.37E-05	3.50E-05	1.85E-05	6.26E-05	1.19E-05	9.27E-06	1.21E-06	1.09E-05	2.46E-06
GTP	4.87E-03	5.13E-04	2.89E-04	1.11E-04	2.15E-05	4.57E-04	9.52E-05	7.76E-04	3.60E-04
guanine	1.88E-04	3.21E-05	1.96E-07	9.77E-05	3.26E-05	6.29E-05	0.00E+00	7.69E-05	4.07E-06
guanosine	1.62E-06	1.33E-06	1.56E-07	3.99E-06	2.12E-06	4.48E-07	3.00E-07	5.62E-07	4.78E-07
histidine	6.76E-05	4.74E-05	1.17E-05	9.98E-04	1.33E-04	4.08E-05	1.31E-05	4.83E-05	5.28E-06
IMP	2.72E-04	6.45E-05	7.05E-06	2.71E-04	2.27E-04	2.71E-04	9.71E-06	1.68E-04	8.79E-05
isoleucine+leucine	3.03E-04	3.82E-04	4.50E-05	6.45E-03	7.56E-04	6.86E-04	7.30E-05	7.77E-04	4.44E-05
lysine	4.05E-04	1.67E-03	4.67E-04	6.88E-03	3.40E-04	4.36E-03	5.92E-04	3.38E-03	1.80E-04
malate	1.68E-03	1.11E-04	3.90E-05	3.78E-04	1.78E-05	3.43E-04	2.49E-05	2.76E-04	2.84E-05
methionine	1.45E-04	7.94E-06	2.88E-06	1.88E-05	5.44E-06	8.93E-06	1.12E-06	1.54E-05	5.07E-06
N-acetyl-glucosamine 1/6 phosphate	8.19E-05	2.06E-05	3.41E-06	1.72E-05	1.54E-06	2.78E-05	6.36E-06	2.47E-05	1.24E-05
NAD+	2.55E-03	5.51E-03	1.44E-03	1.00E-02	7.98E-04	7.71E-03	4.97E-04	7.78E-03	4.96E-04
NADH	8.36E-05	3.81E-05	9.62E-06	6.80E-05	6.93E-06	4.91E-05	3.57E-06	4.58E-05	2.52E-06
NADP+	2.08E-06	1.44E-05	1.04E-05	3.49E-06	6.16E-07	4.49E-06	2.84E-07	5.39E-06	4.43E-07
NADPH	1.21E-04	3.88E-05	9.76E-06	4.13E-05	9.00E-06	5.35E-05	4.01E-06	6.65E-05	6.54E-06
p-hydroxybenzoate	5.22E-05	4.46E-04	2.89E-05	5.52E-04	2.46E-05	6.46E-04	8.43E-06	5.20E-04	6.13E-06
phenylalanine	1.82E-05	5.73E-04	8.37E-05	1.59E-03	1.27E-04	8.67E-04	1.38E-04	5.88E-04	9.78E-05
phenylpyruvate	8.98E-05	4.61E-04	3.05E-05	5.97E-04	2.88E-05	6.10E-04	5.60E-05	6.80E-04	2.76E-05
phosphoenolpyruvate	1.84E-04	1.57E-04	3.51E-05	2.02E-04	2.75E-05	1.61E-04	9.43E-06	1.95E-04	4.99E-06
proline	3.85E-04	5.27E-05	1.24E-05	4.44E-04	6.51E-05	1.10E-04	1.07E-05	8.45E-05	5.43E-06
pyruvate	3.66E-03	3.72E-03	3.27E-04	1.65E-02	3.01E-03	2.66E-03	1.82E-04	3.50E-03	2.16E-04

R5P/Ru5P/Xu5P	1.08E-03	7.25E-04	8.22E-05	2.26E-03	2.08E-04	1.72E-03	2.24E-04	1.39E-03	3.66E-05
S7P	8.82E-04	6.83E-04	1.59E-04	9.66E-04	9.83E-05	8.09E-04	8.17E-05	1.00E-03	4.13E-05
serine	1.13E-03	1.89E-04	7.34E-06	8.41E-04	3.99E-04	1.97E-04	4.01E-05	2.47E-04	7.41E-05
sn-glycerol-3-phosphate	4.90E-05	5.26E-05	2.69E-05	3.53E-05	5.02E-06	2.51E-04	8.26E-05	4.18E-05	4.91E-06
succinate	5.69E-04	3.82E-04	5.67E-05	9.99E-04	1.01E-04	4.23E-04	1.12E-04	3.97E-04	1.10E-05
succinyl-CoA	2.33E-04	5.31E-05	1.51E-05	7.97E-05	1.39E-05	8.61E-05	1.69E-05	1.20E-04	5.03E-05
threonine	1.26E-03	1.73E-04	3.02E-05	4.82E-04	6.87E-05	2.20E-04	1.72E-05	2.65E-04	1.79E-05
tryptophan	1.21E-05	3.17E-06	9.07E-07	1.77E-05	2.84E-06	5.34E-06	4.50E-07	5.31E-06	7.69E-07
tyrosine	2.89E-05	9.07E-04	1.47E-04	2.28E-03	1.53E-04	1.39E-03	6.88E-05	1.09E-03	1.56E-04
UDP	1.79E-03	1.04E-03	2.66E-04	2.46E-03	2.41E-04	1.28E-03	9.83E-05	1.37E-03	6.61E-05
UDP-D-glucose	2.50E-03	1.05E-03	2.71E-04	1.38E-03	5.86E-05	2.20E-03	2.33E-04	1.55E-03	2.75E-05
UDP-D-glucosamine	5.66E-04	2.80E-03	6.34E-04	2.16E-03	1.10E-04	3.82E-03	4.04E-04	2.40E-03	7.68E-05
UDP-N-acetyl-glucosamine	9.24E-03	1.92E-03	5.26E-04	2.77E-03	8.72E-05	3.36E-03	3.94E-04	2.75E-03	2.31E-05
UTP	8.29E-03	1.46E-03	2.82E-04	2.55E-03	3.02E-04	2.62E-03	6.31E-05	2.22E-03	1.75E-04
uridine	2.09E-03	4.26E-04	8.75E-05	1.18E-03	3.47E-06	4.17E-04	3.09E-05	3.98E-04	2.79E-05
valine	4.02E-03	1.01E-01	2.20E-03	9.54E-02	3.03E-03	1.05E-01	6.74E-03	9.46E-02	4.37E-03
Total Sum	2.38E-01	2.27E-01		3.07E-01		2.96E-01		2.85E-01	
GNG/NOG	3.79E-02	4.33E-02		7.74E-02		5.87E-02		6.59E-02	
TCA/GS	5.80E-03	3.89E-03		8.51E-03		8.04E-03		5.64E-03	
DNA/RNA synthesis	2.04E-02	6.22E-03		1.29E-02		8.00E-03		8.01E-03	
Amino Acids	1.17E-01	1.38E-01		1.61E-01		1.71E-01		1.60E-01	
Energetic Cofactors	1.89E-02	9.48E-03		1.52E-02		1.38E-02		1.35E-02	
Other Metabolites	2.17E-02	1.61E-02		2.01E-02		2.03E-02		1.98E-02	

References

1. Park JO, Tanner LB, Wei MH, Khana DB, Jacobson TB, Zhang Z, Rubin SA, Li SHJ, Higgins MB, Stevenson DM, et al.: **Near-equilibrium glycolysis supports metabolic homoestasis and energy yield**. *Nature* 2019, **15**: 1001-1008.

2. Jacobson TB, Adamczyk PA, Stevenson DM, Regner M, Ralph J, Reed JL, Amador-Noguez D: **(2)H and (13)C metabolic flux analysis elucidates in vivo thermodynamics of the ED pathway in Zymomonas mobilis**. *Metab Eng* 2019, **54**:301-316.

3. Park JO, Rubin SA, Xu YF, Amador-Noguez D, Fan J, Shlomi T, Rabinowitz JD: **Metabolite concentrations, fluxes and free energies imply efficient enzyme usage**. *Nature Chemical Biology* 2016, **12**:482-489.

4. Jacobson TB, Korosh TK, Stevenson DM, Foster C, Maranas C, Olson DG, Lynd LR, Amador-Noguez D: **In Vivo Thermodynamic Analysis of Glycolysis in Clostridium thermocellum and Thermoanaerobacterium saccharolyticum Using (13)C and (2)H Tracers**. *mSystems* 2020, **5**:e00736.

5. Bogorad IW, Lin TS, Liao JC: **Synthetic non-oxidative glycolysis enables complete carbon conservation**. *Nature* 2013, **502**:693-697.

6. Lin PP, Jaeger AJ, Wu TY, Xu SC, Lee AS, Gao F, Chen PW, Liao JC: **Construction and evolution of an Escherichia coli strain relying on nonoxidative glycolysis for sugar catabolism**. *Proc Natl Acad Sci U S A* 2018, **115**:3538-3546.

7. Hui S, Ghergurovich JM, Morscher RJ, Jang C, Teng X, Lu W, Esparza LA, Reya T, Le Z, Yanxiang Guo J, et al.: **Glucose feeds the TCA cycle via circulating lactate**. *Nature* 2017, **551**:115-118.

8. Hackett SR, Zanotelli VR, Xu W, Goya J, Park JO, Perlman DH, Gibney PA, Botstein D, Storey JD, Rabinowitz JD: **Systems-level analysis of mechanisms regulating yeast metabolic flux**. *Science* 2016, **354**:6311.

9. McBrayer SK, Mayers JR, DiNatale GJ, Shi DD, Khanal J, Chakraborty AA, Sarosiek KA, Briggs KJ, Robbins AK, Sewastianik T, et al.: **Transaminase Inhibition by 2-Hydroxyglutarate Impairs Glutamate Biosynthesis and Redox Homeostasis in Glioma**. *Cell* 2018, **175**:101-116.

10. Momcilovic M, Bailey ST, Lee JT, Fishbein MC, Braas D, Go J, Graeber TG, Parlati F, Demo S, Li R, et al.: **The GSK3 Signaling Axis Regulates Adaptive Glutamine Metabolism in Lung Squamous Cell Carcinoma**. *Cancer Cell* 2018, **33**:905-921.

11. Templeton N, Smith KD, McAtee-Pereira AG, Dorai H, Betenbaugh MJ, Lang SE, Young JD: **Application of C-13 flux analysis to identify high-productivity CHO metabolic phenotypes**. *Metab Eng* 2017, **43**:218-225.

12. Park JO, Liu N, Holinski KM, Emerson DF, Qiao K, Woolston BM, Xu J, Lazar Z, Islam MA, Vidoudez C et al.: **Synergistic substrate cofeeding stimulates reductive metabolism**. *Nat Metab* 2019, **1**:643-651.

13. Xie XQ, Garg A, Khosla C, Cane DE: **Elucidation of the Cryptic Methyl Group Epimerase Activity of Dehydratase Domains from Modular Polyketide Synthases Using a Tandem Modules Epimerase Assay**. *J Am Chem Soc* 2017, **139**:9507-9510.

14. Lis AV, Schneider K, Weber J, Keasling JD, Jensen MK, Klein T: **Exploring small-scale chemostats to scale up microbial processes: 3-hydroxypropionic acid production in S. cerevisiae**. *Microb Cell Fact* 2019, **18**:50.

15. Chen WW, Freinkman E, Wang T, Birsoy K, Sabatini DM: **Absolute Quantification of Matrix Metabolites Reveals the Dynamics of Mitochondrial Metabolism**. *Cell* 2016, **166**:1324-1337.

16. Cordova LT, Cipolla RM, Swarup A, Long CP, Antoniewicz MR: **C-13 metabolic flux analysis of three divergent extremely thermophilic bacteria: Geobacillus sp LC300, Thermus thermophilus HB8, and Rhodothermus marinus DSM 4252**. *Metab Eng* 2017, **44**:182-190.

17. Metallo CM, Gameiro PA, Bell EL, Mattaini KR, Yang J, Hiller K, Jewell CM, Johnson ZR, Irvine DJ, Guarente L, et al.: **Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia**. *Nature* 2011, **481**:380-384.

18. Lee WD, Mukha D, Aizenshtein E, Shlomi T: **Spatial-fluxomics provides a subcellular-compartmentalized view of reductive glutamine metabolism in cancer cells**. *Nat Commun* 2019, **10**:1351.

19. Buescher JM, Antoniewicz MR, Boros LG, Burgess SC, Brunengraber H, Clish CB, DeBerardinis RJ, Feron O, Frezza C, Ghesquiere B, et al.: **A roadmap for interpreting (13)C metabolite labeling patterns from cells**. *Curr Opin Biotechnol* 2015, **34**:189-201.

20. Latendresse M, Malerich JP, Travers M, Karp PD: **Accurate atom-mapping computation for biochemical reactions**. *J Chem Inf Model* 2012, **52**:2970-2982.

21. Nilsson R, Jain M: **Simultaneous tracing of carbon and nitrogen isotopes in human cells**. *Mol Biosyst* 2016, **12**:1929-1937.

22. Kreuzer-Martin HW, Ehleringer JR, Hegg EL: **Oxygen isotopes indicate most intracellular water in log-phase Escherichia coli is derived from metabolism**. *Proc Natl Acad Sci U S A* 2005, **102**:17337-17341.

23. Liu L, Shah S, Fan J, Park JO, Wellen KE, Rabinowitz JD: **Malic enzyme tracers reveal hypoxia-induced switch in adipocyte NADPH pathway usage**. *Nat Chem Biol* 2016, **12**:345-352.

24. Chen L, Zhang Z, Hoshino A, Zheng HD, Morley M, Arany Z, Rabinowitz JD: **NADPH production by the oxidative pentose-phosphate pathway supports folate metabolism**. *Nat Metab* 2019, **1**:404-415.

25. Kirkwood JS, Miranda CL, Bobe G, Maier CS, Stevens JF: **18O-Tracer Metabolomics Reveals Protein Turnover and CDP-Choline Cycle Activity in Differentiating 3T3-L1 Pre-Adipocytes**. *PLoS One* 2016, **11**:e0157118.

26. Gorenstein DG, Lee YG, Kar D: **Kinetic Isotope Effects in Reactions of Aryl-O-18-2,4-Dinitrophenyl Dibenzyl Phosphate and Aryl-O-18-2,4-Dinitrophenyl Phosphate - Evidence for Monomeric Metaphosphate**. *J Am Chem Soc* 1977, **99**:2264-2267.

27. Mirica LM, McCusker KP, Munos JW, Liu HW, Klinman JP: **O-18 kinetic isotope effects in non-heme iron enzymes: Probing the nature of Fe/O-2 intermediates**. *J Am Chem Soc* 2008, **130**:8122-8123.

28. Neinast MD, Jang C, Hui S, Murashige DS, Chu Q, Morscher RJ, Li X, Zhan L, White E, Anthony TG, et al.: **Quantitative Analysis of the Whole-Body Metabolic Fate of Branched-Chain Amino Acids**. *Cell Metab* 2019, **29**:417-429.

29. Spinelli JB, Yoon H, Ringel AE, Jeanfavre S, Clish CB, Haigis MC: **Metabolic recycling of ammonia via glutamate dehydrogenase supports breast cancer biomass**. *Science* 2017, **358**:941-946.

30. Liu X, Cooper DE, Cluntun AA, Warmoes MO, Zhao S, Reid MA, Liu J, Lund PJ, Lopes M, Garcia BA, et al.: **Acetate Production from Glucose and Coupling to Mitochondrial Metabolism in Mammals**. *Cell* 2018, **175**:502-513.

31. Fan J, Ye J, Kamphorst JJ, Shlomi T, Thompson CB, Rabinowitz JD: **Quantitative flux analysis reveals folate-dependent NADPH production**. *Nature* 2014, **510**:298-302.

32. Lanfermann I, Krings U, Schopp S, Berger RG: **Isotope labelling experiments on the formation pathway of 3-hydroxy-4,5-dimethyl-2(5H)-furanone from l-isoleucine in cultures of Laetiporus sulphureus**. *Flavour Frag J* 2014, **29**:233-239.

33. Kloehn J, Saunders EC, O'Callaghan S, Dagley MJ, McConville MJ: **Characterization of Metabolically Quiescent Leishmania Parasites in Murine Lesions Using Heavy Water Labeling**. *PLos Pathog* 2015, **11**:e1004683.

34. Xiao G, Chan LN, Klemm L, Braas D, Chen Z, Geng H, Zhang QC, Aghajanirefah A, Cosgun KN, Sadras T, et al.: **B-Cell-Specific Diversion of Glucose Carbon Utilization Reveals a Unique Vulnerability in B Cell Malignancies**. *Cell* 2018, **173**:470-484.

35. Morscher RJ, Ducker GS, Li SH, Mayer JA, Gitai Z, Sperl W, Rabinowitz JD: **Mitochondrial translation requires folate-dependent tRNA methylation**. *Nature* 2018, **554**:128-132.

36. Ducker GS, Chen L, Morscher RJ, Ghergurovich JM, Esposito M, Teng X, Kang Y, Rabinowitz JD: **Reversal of Cytosolic One-Carbon Flux Compensates for Loss of the Mitochondrial Folate Pathway**. *Cell Metab* 2016, **24**:640-641.

37. Sun RC, Fan TWM, Deng P, Higashi RM, Lane AN, Le AT, Scott TL, Sun QS, Warmoes MO, Yang Y: **Noninvasive liquid diet delivery of stable isotopes into mouse models for deep metabolic network tracing**. *Nat Commun* 2017, **8**:1646.

38. Chen JH, Shahnavas S, Singh N, Ong WY, Walczyk T: **Stable iron isotope tracing reveals significant brain iron uptake in adult rats**. *Metallomics* 2013, **5**:167-173.

39. Crown SB, Marze N, Antoniewicz MR: **Catabolism of Branched Chain Amino Acids Contributes Significantly to Synthesis of Odd-Chain and Even-Chain Fatty Acids in 3T3-L1 Adipocytes**. *PLoS One* 2015, **10**:e0145850.

40. Elbaum D, Beconi MG, Monteagudo E, Di Marco A, Quinton MS, Lyons KA, Vaino A, Harper S: **Fosmetpantotenate (RE-024), a phosphopantothenate replacement therapy for pantothenate kinase-associated neurodegeneration: Mechanism of action and efficacy in nonclinical models**. *PLoS One* 2018, **13**:e0192028.

41. Chen Y, Nielsen J: **Energy metabolism controls phenotypes by protein efficiency and allocation**. *Proc Natl Acad Sci USA* 2019, **116**:17592-17597.

42. Yang C, Ko B, Hensley CT, Jiang L, Wasti AT, Kim J, Sudderth J, Calvaruso MA, Lumata L, Mitsche M, et al.: **Glutamine oxidation maintains the TCA cycle and cell survival during impaired mitochondrial pyruvate transport**. *Mol Cell* 2014, **56**:414-424.

43. Seim GLB, E. C.; John, S. V.; Yeo, F. J.; Johnson, A. R.; Eisenstein, R. S.; Pagliarini, D. J.; Fan, J. : **Two-stage metabolic remodelling in macrophages in response to lipopolysaccharide and interferon-γ stimulation**. *Nat Metab* 2019, **1**:731-742.

44. Kukurugya MA, Mendonca CM, Solhtalab M, Wilkes RA, Thannhauser TW, Aristilde L: **Multi-omics analysis unravels a segregated metabolic flux network that tunes co-utilization of sugar and aromatic carbons in Pseudomonas putida**. *J Biol Chem* 2019, **294**:8464-8479.

45. Czajka JJ, Nathenson JA, Benites VT, Baidoo EEK, Cheng Q, Wang Y, Tang YJ: **Engineering the oleaginous yeast Yarrowia lipolytica to produce the aroma compound beta-ionone**. *Microb Cell Fact* 2018, **17**:136.

46. Elia I, Rossi M, Stegen S, Broekaert D, Doglioni G, van Gorsel M, Boon R, Escalona-Noguero C, Torrekens S, Verfaillie C, et al.: **Breast cancer cells rely on environmental pyruvate to shape the metastatic niche**. *Nature* 2019, **568**:117-121.

47. Miller RA, Shi YJ, Lu WY, Pirman DA, Jatkar A, Blatnik M, Wu H, Cardenas C, Wan M, Foskett JK, et al.: **Targeting hepatic glutaminase activity to ameliorate hyperglycemia**. *Nat Med* 2018, **24**:518-524.

48. Wallace M, Green CR, Roberts LS, Lee YM, McCarville JL, Sanchez-Gurmaches J, Meurs N, Gengatharan JM, Hover JD, Phillips SA, et al.: **Enzyme promiscuity drives branched-chain fatty acid synthesis in adipose tissues**. *Nat Chem Biol* 2018, **14**:1021-1031.

49. Nofal M, Zhang K, Han S, Rabinowitz JD: **mTOR Inhibition Restores Amino Acid Balance in Cells Dependent on Catabolism of Extracellular Protein**. *Mol Cell* 2017, **67**:936-946.

50. Palm W, Park Y, Wright K, Pavlova NN, Tuveson DA, Thompson, CB: **The Utilization of Extracellular Proteins as Nutrients Is Suppressed by mTORC1**. *Cell* 2015, **162**:259-270.

51. Flamholz A, Noor E, Bar-Even A, Liebermeister W, Milo R: **Glycolytic strategy as a tradeoff between energy yield and protein cost**. *Proc Natl Acad Sci USA* 2013, **110**:10039-10044.

52. Enjalbert B, Millard P, Dinclaux M, Portais JC, Letisse F: **Acetate fluxes in Escherichia coli are determined by the thermodynamic control of the Pta-AckA pathway**. *Sci Rep* 2017, **7**:42135.

53. Tanner LB, Goglia AG, Wei MH, Sehgal T, Parsons LR, Park JO, White E, Toettcher JE, Rabinowitz JD: **Four Key Steps Control Glycolytic Flux in Mammalian Cells**. *Cell Syst* 2018, **7**:49-62.

54. International Energy Agency: **Global Energy and CO2 Status Report**. *IEA* 2018, 1-28.

55. Yang D, Zhu Q, Chen C, Liu H, Liu Z, Zhao Z, Zhang X, Liu S, Han B: **Selective electroreduction of carbon dioxide to methanol on copper selenide nanocatalysts**. *Nat Commun* 2019, **10**:677.

56. Li D, Huang L, Liu T, Liu J, Zhen L, Wu J, Feng Y: **Electrochemical reduction of carbon dioxide to formate via nano-prism assembled CuO microspheres**. *Chemosphere* 2019, **237**:124527.

57. Jeletic MS, Mock MT, Appel AM, Linehan JC: **A cobalt-based catalyst for the hydrogenation of CO2 under ambient conditions**. *J Am Chem Soc* 2013, **135**:11533-11536.

58. Arandiyan H, Wang Y, Scott J, Mesgari S, Dai H, Amal R: **In Situ Exsolution of Bimetallic Rh-Ni Nanoalloys: a Highly Efficient Catalyst for CO2 Methanation**. *ACS Appl Mater Interfaces* 2018, **10**:16352-16357.

59. Hermann BG, Blok K, Patel MK: **Producing bio-based bulk chemicals using industrial biotechnology saves energy and combats climate change**. *Environ Sci Technol* 2007, **41**:7915-7921.

60. Erickson B, Nelson JE, Winters P: **Perspective on opportunities in industrial biotechnology in renewable chemicals**. *Biotechnol J* 2012, **7**:176-185.

61. Beller M, Bornscheuer UT: **CO(2) fixation through hydrogenation by chemical or enzymatic methods**. *Angew Chem Int Ed Engl* 2014, **53**:4527-4528.

62. Ragsdale SW, Pierce E: **Acetogenesis and the Wood-Ljungdahl pathway of CO2 fixation**. *Biochimica Et Biophysica Acta-Proteins and Proteomics* 2008, **1784**:1873-1898.

63. Court SJ, Waclaw B, Allen RJ: **Lower glycolysis carries a higher flux than any biochemically possible alternative**. *Nat Commun* 2015, **6**:8427.

64. Gleizer S, Ben-Nissan R, Bar-On YM, Antonovsky N, Noor E, Zohar Y, Jona G, Krieger E, Shamshoum M, Bar-Even A, et al.: **Conversion of Escherichia coli to Generate All Biomass Carbon from CO2**. *Cell* 2019, **179**:1255-1263.

65. Bang J, Lee SY: **Assimilation of formic acid and CO2 by engineered Escherichia coli equipped with reconstructed one-carbon assimilation pathways**. *Proc Natl Acad Sci U S A* 2018, **115**:9271-9279.

66. Kim SJ, Yoon J, Im DK, Kim YH, Oh MK: **Adaptively evolved Escherichia coli for improved ability of formate utilization as a carbon source in sugar-free conditions**. *Biotechnol Biofuels* 2019, **12**:207.

67. Kim S, Lindner SN, Aslan S, Yishai O, Wenk S, Schann K, Bar-Even A: **Growth of E. coli on formate and methanol via the reductive glycine pathway**. *Nat Chem Biol* 2020, **16**:538-545.

68. Biswas R, Yamaoka M, Nakayama H, Kondo T, Yoshida K, Bisaria VS, Kondo A: **Enhanced production of 2,3-butanediol by engineered Bacillus subtilis**. *Appl Microbiol Biotechnol* 2012, **94**:651-658.

69. Cheng KK, Wu J, Lin ZN, Zhang JA: **Aerobic and sequential anaerobic fermentation to produce xylitol and ethanol using non-detoxified acid pretreated corncob**. *Biotechnol Biofuels* 2014, **7**:166.

70. Avila-Bolivar B, Garcia-Cruz L, Montiel V, Solla-Gullon J: **Electrochemical Reduction of CO2 to Formate on Easily Prepared Carbon-Supported Bi Nanoparticles**. *Molecules* 2019, **24**:2032.

71. Pinske C, Sargent F: **Exploring the directionality of Escherichia coli formate hydrogenlyase: a membrane-bound enzyme capable of fixing carbon dioxide to organic acid**. *Microbiologyopen* 2016, **5**:721-737.

72. Zelcbuch L, Lindner SN, Zegman Y, Vainberg Slutskin I, Antonovsky N, Gleizer S, Milo R, Bar-Even A: **Pyruvate Formate-Lyase Enables Efficient Growth of Escherichia coli on Acetate and Formate**. *Biochemistry* 2016, **55**:2423-2426.

73. Kirkpatrick C, Maurer LM, Oyelakin NE, Yoncheva YN, Maurer R, Slonczewski JL: **Acetate and formate stress: opposite responses in the proteome of Escherichia coli**. *J Bacteriol* 2001, **183**:6466-6477.

74. Sah S, Aluri S, Rex K, Varshney U: **One-carbon metabolic pathway rewiring in Escherichia coli reveals an evolutionary advantage of 10-formyltetrahydrofolate synthetase (Fhs) in survival under hypoxia**. *J Bacteriol* 2015, **197**:717-726.

75. Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD: **Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli**. *Nature Chemical Biology* 2009, **5**:593-599.

76. Siddiquee KA, Arauzo-Bravo MJ, Shimizu K: **Effect of a pyruvate kinase (pykF-gene) knockout mutation on the control of gene expression and metabolic fluxes in Escherichia coli**. *FEMS Microbiol Lett* 2004, **235**:25-33.

77. Wagner AF, Schultz S, Bomke J, Pils T, Lehmann WD, Knappe J: **YfiD of Escherichia coli and Y06I of bacteriophage T4 as autonomous glycyl radical cofactors reconstituting the catalytic center of oxygen-fragmented pyruvate formate-lyase**. *Biochem Biophys Res Commun* 2001, **285**:456-462.

78. Wyborn NR, Messenger SL, Henderson RA, Sawers G, Roberts RE, Attwood MM, Green J: **Expression of the Escherichia coli yfiD gene responds to intracellular pH and reduces the accumulation of acidic metabolic end products**. *Microbiology-Sgm* 2002, **148**:1015-1026.

79. Gerth U, Krieger E, Zuhlke D, Reder A, Volker U, Hecker M: **Stability of Proteins Out of Service: the GapB Case of Bacillus subtilis**. *J Bacteriol* 2017, **199**:e00148.

80. Tastensen JB, Schonheit P: **Two distinct glyceraldehyde-3-phosphate dehydrogenases in glycolysis and gluconeogenesis in the archaeon Haloferax volcanii**. *FEBS Lett* 2018, **592**:1524-1534.

81. Buffing MF, Link H, Christodoulou D, Sauer U: **Capacity for instantaneous catabolism of preferred and non-preferred carbon sources in Escherichia coli and Bacillus subtilis**. *Sci Rep* 2018, **8**:11760.

82. Bar-Even A: **Formate Assimilation: The Metabolic Architecture of Natural and Synthetic Pathways**. *Biochemistry* 2016, **55**:3851-3863.

83. Noor E, Bar-Even A, Flamholz A, Reznik E, Liebermeister W, Milo R: **Pathway Thermodynamics Highlights Kinetic Obstacles in Central Metabolism**. *PLos Comput Biol* 2014, **10**: e1003483.

84. Himo F, Eriksson LA: **Catalytic mechanism of pyruvate formate-lyase (PFL). A theoretical study**. *J Am Chem Soc* 1998, **120**:11449-11455.

85. Knappe J, Blaschkowski HP, Grobner P, Schmitt T: **Pyruvate formate-lyase of Escherichia coli: the acetyl-enzyme intermediate**. *Eur J Biochem* 1974, **50**:253-263.

86. Kao KC, Tran LM, Liao JC: **A global regulatory role of gluconeogenic genes in Escherichia coli revealed by transcriptome network analysis**. *J Biol Chem* 2005, **280**:36079-36087.

87. Sauer U, Eikmanns BJ: **The PEP-pyruvate-oxaloacetate node as the switch point for carbon flux distribution in bacteria**. *Fems Microbiology Reviews* 2005, **29**:765-794.

88. Trotter EW, Rolfe MD, Hounslow AM, Craven CJ, Williamson MP, Sanguinetti G, Poole RK, Green J: **Reprogramming of Escherichia coli K-12 Metabolism during the Initial Phase of Transition from an Anaerobic to a Micro-Aerobic Environment**. *PLoS One* 2011, **6**:e25501.

89. Pedraz L, Blanco-Cabra N, Torrents E: **Gradual adaptation of facultative anaerobic pathogens to microaerobic and anaerobic conditions**. *Faseb Journal* 2020, **34**:2912-2928.

90. Wang HN, Gunsalus RP: **Coordinate regulation of the Escherichia coli formate dehydrogenase fdnGHI and fdhF genes in response to nitrate, nitrite, and formate: Roles for NarL and NarP**. *J Bacteriol* 2003, **185**:5076-5085.

91. Sawers G, Suppmann B: **Anaerobic Induction of Pyruvate Formate-Lyase Gene-Expression Is Mediated by the Arca and Fnr Proteins**. *J Bacteriol* 1992, **174**:3474-3478.

92. Zhu JF, Shalel-Levanon S, Bennett G, San KY: **The yfiD protein contributes to the pyruvate formate-lyase flux in an Escherichia coli arcA mutant strain**. *Biotechnol Bioeng* 2007, **97**:138-143.

93. Levanon SS, San KY, Bennett GN: **Effect of oxygen on the Escherichia coli ArcA and FNR regulation systems and metabolic responses**. *Biotechnol Bioeng* 2005, **89**:556-564.

94. Nikel PI, Zhu J, San KY, Galvagno MA, Mendez BS, Bennett GN: **Metabolic flux analysis of Escherichia coli ArcA and CreB mutants reveals shared control of carbon catabolism under microaerobic conditions of growth**. *J Bacteriol* 2009, **191**:5538-5548.

95. Alexeeva S, de Kort B, Sawers G, Hellingwerf KJ, de Mattos MJT: **Effects of limited aeration and of the ArcAB system on intermediary pyruvate catabolism in Escherichia coli**. *J Bacteriol* 2000, **182**:4934-4940.

96. Alexeeva S, Hellingwerf KJ, de Mattos MJT: **Requirement of ArcA for redox regulation in Escherichia coli under microaerobic but not anaerobic or aerobic conditions**. *J Bacteriol* 2003, **185**:204-209.

97. Zhao CH, Lin Z, Dong HJ, Zhang YP, Li Y: **Reexamination of the Physiological Role of PykA in Escherichia coli Revealed that It Negatively Regulates the Intracellular ATP Levels under Anaerobic Conditions**. *Appl Environ Microbiol* 2017, **83**:e00316.

98. Murarka A, Clomburg JM, Moran S, Shanks JV, Gonzalez R: **Metabolic analysis of wild-type Escherichia coli and a pyruvate dehydrogenase complex (PDHC)-deficient derivative reveals the role of PDHC in the fermentative metabolism of glucose**. *J Biol Chem* 2010, **285**:31548-31558.

99. Toya Y, Nakahigashi K, Tomita M, Shimizu K: **Metabolic regulation analysis of wild-type and arcA mutant Escherichia coli under nitrate conditions using different levels of omics data**. *Mol Biosyst* 2012, **8**:2593-2604.

100. Li M, Ho PY, Yao S, Shimizu K: **Effect of lpdA gene knockout on the metabolism in Escherichia coli based on enzyme activities, intracellular metabolite concentrations and metabolic flux analysis by 13C-labeling experiments**. *J Biotechnol* 2006, **122**:254-266.

101. Kotlarz D, Garreau H, Buc H: **Regulation of the amount and of the activity of phosphofructokinases and pyruvate kinases in Escherichia coli**. *Biochimica Et Biophysica Acta* 1975, **381**:257-268.

102. Forster AH, Gescher J: **Metabolic Engineering of Escherichia coli for Production of Mixed-Acid Fermentation End Products**. *Front Bioeng Biotechnol* 2014, **2**:16.

103. Iuchi S, Lin EC: **arcA (dye), a global regulatory gene in Escherichia coli mediating repression of enzymes in aerobic pathways**. *Proc Natl Acad Sci USA* 1988, **85**:1888-1892.

104. Rodriguez A, Martinez JA, Flores N, Escalante A, Gosset G, Bolivar F: **Engineering Escherichia coli to overproduce aromatic amino acids and derived compounds**. *Microb Cell Fact* 2014, **13**:126.

105. Cotton CA, Claassens NJ, Benito-Vaquerizo S, Bar-Even A: **Renewable methanol and formate as microbial feedstocks**. *Curr Opin Biotechnol* 2020, **62**:168-180.

106. Kobayashi M, Kakizono T, Nagai S: **Enhanced Carotenoid Biosynthesis by Oxidative Stress in Acetate-Induced Cyst Cells of a Green Unicellular Alga, Haematococcus pluvialis**. *Appl Environ Microbiol* 1993, **59**:867-873.

107. Kang CD, Lee JS, Park TH, Sim SJ: **Comparison of heterotrophic and photoautotrophic induction on astaxanthin production by Haematococcus pluvialis**. *Appl Microbiol Biotechnol* 2005, **68**:237-241.

108. Yan S, Dong D: **Improvement of caproic acid production in a Clostridium kluyveri H068 and Methanogen 166 co-culture fermentation system**. *AMB Express* 2018, **8**:175.

109. Leone S, Sannino F, Tutino ML, Parrilli E, Picone D: **Acetate: friend or foe? Efficient production of a sweet protein in Escherichia coli BL21 using acetate as a carbon source**. *Microb Cell Fact* 2015, **14**:106.

110. Xu X, Xie M, Zhao Q, Xian M, Liu H: **Microbial production of mevalonate by recombinant Escherichia coli using acetic acid as a carbon source**. *Bioengineered* 2018, **9**:116-123.

111. Yang X, Yuan Q, Zheng Y, Ma H, Chen T, Zhao X: **An engineered non-oxidative glycolysis pathway for acetone production in Escherichia coli**. *Biotechnol Lett* 2016, **38**:1359-1365.

112. Doll K, Ruckel A, Kampf P, Wende M, Weuster-Botz D: **Two stirred-tank bioreactors in series enable continuous production of alcohols from carbon monoxide with Clostridium carboxidivorans**. *Bioprocess Biosyst Eng* 2018, **41**:1403-1416.

113. Tapia F, Wohlfarth D, Sandig V, Jordan I, Genzel Y, Reichl U: **Continuous influenza virus production in a tubular bioreactor system provides stable titers and avoids the "von Magnus effect"**. *PLoS One* 2019, **14**:e0224317.

114. Ahn S, Jung J, Jang IA, Madsen EL, Park W: **Role of Glyoxylate Shunt in Oxidative Stress Response**. *J Biol Chem* 2016, **291**:11928-11938.

115. Xu J, Martien J, Gilbertson C, Ma J, Amador-Noguez D, Park JO: **Metabolic flux analysis and fluxomics-driven determination of reaction free energy using multiple isotopes**. *Curr Opin Biotechnol* 2020, **64**:151-160.

116. Hollinshead WD, Rodriguez S, Martin HG, Wang G, Baidoo EEK, Sale KL, Keasling JD, Mukhopadhyay A, Tang YJJ: **Examining Escherichia coli glycolytic pathways, catabolite repression, and metabolite channeling using Delta pfk mutants**. *Biotechnol Biofuels* 2016, **9**:212.

117. Mohamed ET, Mundhada H, Landberg J, Cann I, Mackie RI, Nielsen AT, Herrgard MJ, Feist AM: **Generation of an E. coli platform strain for improved sucrose utilization using adaptive laboratory evolution**. *Microb Cell Fact* 2019, **18**:116.

118. Luo H, Yang L, Kim SH, Wulff T, Feist AM, Herrgard M, Palsson BO: **Directed Metabolic Pathway Evolution Enables Functional Pterin-Dependent Aromatic-Amino-Acid Hydroxylation in Escherichia coli**. *ACS Synth Biol* 2020, **9**:494-499.

119. Wen Z, Ledesma-Amaro R, Lin J, Jiang Y, Yang S: **Improved n-Butanol Production from Clostridium cellulovorans by Integrated Metabolic and Evolutionary Engineering**. *Appl Environ Microbiol* 2019, **85**:e02560.

120. Orth JD, Thiele I, Palsson BO: **What is flux balance analysis?** *Nat Biotechnol* 2010, **28**:245-248.

121. Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, Haraldsdottir HS, Wachowiak J, Keating SM, Vlasov V, et al.: **Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0**. *Nat Protoc* 2019, **14**:639-702.

122. Jiang Y, Chen B, Duan C, Sun B, Yang J, Yang S: **Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system**. *Appl Environ Microbiol* 2015, **81**:2506-2514.

123. Pyne ME, Moo-Young M, Chung DEA, Chou CP: **Coupling the CRISPR/Cas9 System with Lambda Red Recombineering Enables Simplified Chromosomal Gene Replacement in Escherichia coli**. *Appl Environ Microbiol* 2015, **81**:5103-5114.

124. Bennett BD, Yuan J, Kimball EH, Rabinowitz JD: **Absolute quantitation of intracellular metabolite concentrations by an isotope ratio-based approach**. *Nat Protoc* 2008, **3**:1299-1311.

125. Wang L, Xing X, Chen L, Yang L, Su X, Rabitz H, Lu W, Rabinowitz JD: **Peak Annotation and Verification Engine for Untargeted LC-MS Metabolomics**. *Anal Chem* 2019, **91**:1838-1846.

126. Moreno-Sanchez R, Saavedra E, Rodriguez-Enriquez S, Olin-Sandoval V: **Metabolic control analysis: a tool for designing strategies to manipulate metabolic pathways**. *J Biomed Biotechnol* 2008, **2008**:597913.