

2017-09-12

# Genome-scale architecture of small molecule regulatory networks and the fundamental trade-off betw...

*This work was made openly accessible by BU Faculty. Please [share](#) how this access benefits you. Your story matters.*

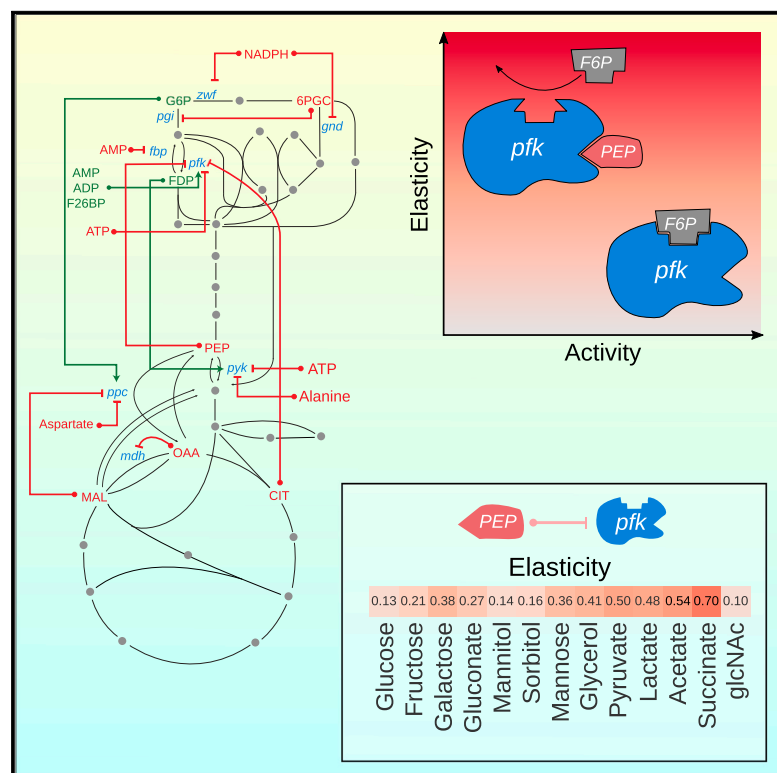
Version	
Citation (published version):	Ed Reznik, Dimitris Christodoulou, Joshua E Goldford, Emma Briars, Uwe Sauer, Daniel Segrè, Elad Noor. 2017. "Genome-Scale Architecture of Small Molecule Regulatory Networks and the Fundamental Trade-Off between Regulation and Enzymatic Activity.. Cell Rep, Volume 20, Issue 11, pp. 2666 - 2677.

<https://hdl.handle.net/2144/24310>

*Boston University*

## Genome-Scale Architecture of Small Molecule Regulatory Networks and the Fundamental Trade-Off between Regulation and Enzymatic Activity

### Graphical Abstract



### Authors

Ed Reznik, Dimitris Christodoulou, Joshua E. Goldford, Emma Briars, Uwe Sauer, Daniel Segrè, Elad Noor

### Correspondence

reznike@mskcc.org (E.R.),  
noor@imsb.biol.ethz.ch (E.N.)

### In Brief

Reznik et al. report a computational pipeline for the genome-scale reconstruction of metabolic small molecule regulatory networks. They describe general design principles underlying small molecule regulation, and they prove a fundamental trade-off between the activity of a metabolic enzyme and the extent to which it is regulated.

### Highlights

- Automatic reconstruction of genome-scale small molecule regulatory networks
- A fundamental trade-off between regulatory capacity and enzymatic activity
- Condition-dependent regulatory capacity of small molecule interactions in *E. coli*
- Identification of conserved small molecule regulation across kingdoms of life



# Genome-Scale Architecture of Small Molecule Regulatory Networks and the Fundamental Trade-Off between Regulation and Enzymatic Activity

Ed Reznik,<sup>1,2,7,8,\*</sup> Dimitris Christodoulou,<sup>3,4,8</sup> Joshua E. Goldford,<sup>5</sup> Emma Briars,<sup>5</sup> Uwe Sauer,<sup>3</sup> Daniel Segrè,<sup>2,5,6</sup> and Elad Noor<sup>3,8,9,\*</sup>

<sup>1</sup>Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

<sup>2</sup>Department of Biomedical Engineering, Boston University, Boston, MA, USA

<sup>3</sup>Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland

<sup>4</sup>Systems Biology Graduate School, Zurich 8057, Switzerland

<sup>5</sup>Bioinformatics Program, Boston University, Boston, MA, USA

<sup>6</sup>Department of Biology, Boston University, Boston, MA, USA

<sup>7</sup>Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

<sup>8</sup>These authors contributed equally

<sup>9</sup>Lead Contact

\*Correspondence: [reznike@mskcc.org](mailto:reznike@mskcc.org) (E.R.), [noor@imsb.biol.ethz.ch](mailto:noor@imsb.biol.ethz.ch) (E.N.)

<http://dx.doi.org/10.1016/j.celrep.2017.08.066>

## SUMMARY

Metabolic flux is in part regulated by endogenous small molecules that modulate the catalytic activity of an enzyme, e.g., allosteric inhibition. In contrast to transcriptional regulation of enzymes, technical limitations have hindered the production of a genome-scale atlas of small molecule-enzyme regulatory interactions. Here, we develop a framework leveraging the vast, but fragmented, biochemical literature to reconstruct and analyze the small molecule regulatory network (SMRN) of the model organism *Escherichia coli*, including the primary metabolite regulators and enzyme targets. Using metabolic control analysis, we prove a fundamental trade-off between regulation and enzymatic activity, and we combine it with metabolomic measurements and the SMRN to make inferences on the sensitivity of enzymes to their regulators. Generalizing the analysis to other organisms, we identify highly conserved regulatory interactions across evolutionarily divergent species, further emphasizing a critical role for small molecule interactions in the maintenance of metabolic homeostasis.

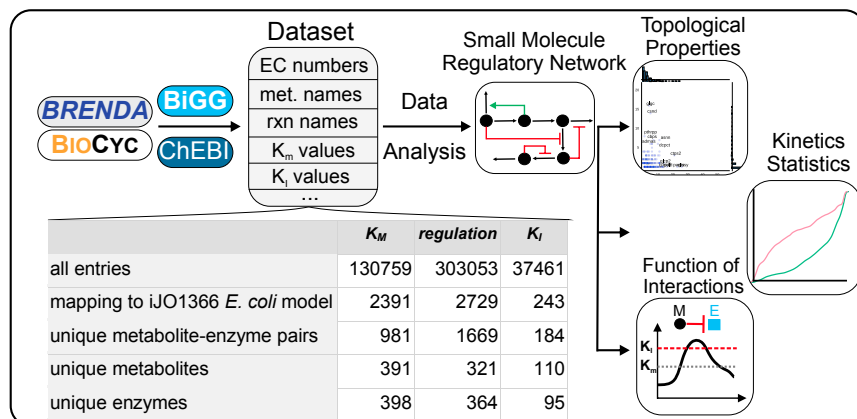
## INTRODUCTION

Despite nearly a century of accumulated experimental data on the identity, biophysical nature, and structural basis of small molecule regulation of enzymes, there has been little progress in elucidating at genome-scale the regulation of enzymes by small molecules. Such interactions, e.g., allosteric regulation or competitive inhibition, play an essential role in homeostasis and in fast adaptations to abrupt environmental changes. For instance, feedback inhibition of amino acid biosynthesis path-

ways by end products preserves anabolic resources when sufficient levels of amino acids are available (Stryer et al., 2002). In contrast, the ultrasensitive feedforward activation of PEP carboxylase by fructose bisphosphate in glycolysis enables *Escherichia coli* to rapidly import glucose following periods of carbon starvation (Xu et al., 2012a).

Recently, it has been established that properly accounting for the activation/inhibition of enzymes by endogenous small molecules can lead to metabolic models that explain experimental data better (Chandra et al., 2011; Hackett et al., 2016; Khodayari and Maranas, 2016; Link et al., 2013; Xu et al., 2012a), facilitate engineering of novel metabolic pathways (Chen et al., 2015; He et al., 2016), and improve our understanding of metabolic phenomena in health and disease (Christofk et al., 2008). So far, high-throughput experimental assays for discovering small molecule regulatory interactions have been technically limited (Feng et al., 2014; Li et al., 2013; Nikolaev et al., 2016; Orsak et al., 2012; Reinhard et al., 2015; Savitski et al., 2014), while hybrid approaches that integrate experimental data with computational models are not scalable and typically focus on central carbon metabolism (Hackett et al., 2016; Link et al., 2013, 2014; Schueler-Furman and Wodak, 2016).

An alternative strategy for studying small molecule regulation is to leverage the vast record of biochemical studies to informatively reconstruct a small molecule regulatory network (SMRN) (Alam et al., 2017). Such an approach would produce a network of interactions between enzymes and metabolites/small molecules (terms we use interchangeably here) that mirrors the native interactions of metabolites as substrates for enzymes and could be naturally integrated with genome-scale metabolic models (GSMMs, e.g., Biochemical, Genetic and Genomic [BiGG] models; King et al., 2016), which are in wide use today. An informatic approach would likely cover a larger swath of metabolism, including peripheral and rarely studied pathways, by aggregating experimental data from many reports. Furthermore, reports of small molecule/enzyme regulation from separate publications could provide additional, independent evidence for such an



**Figure 1. Framework for Reconstruction and Analysis of the Small Molecule Regulatory Network**

The BRENDA and BioCyc database were mined for each reaction taking place in *E. coli*. The identified entries (including data on EC numbers, enzyme names, activating or inhibiting small molecule-enzyme interactions, metabolite names, and  $K_M$  and  $K_I$  values) were stored and then matched by EC number to reactions in the most recent genome-scale reconstruction of *E. coli*, iJO1366. This dataset was searched and analyzed for regulatory small molecules, yielding a comprehensive SMRN. The SMRN was used as the primary resource for the remainder of the analysis. We analyzed the topological properties of the SMRN, evaluated the similarities and differences in the kinetic properties of reactions and interactions, and used published metabolite concentration data in order to evaluate the functional role of inhibitory small molecule-enzyme interactions.

interaction. Finally, an informatically reconstructed SMRN would also offer a window to distilling how the critical regulatory components of the cell, i.e., the regulating metabolites and the regulated enzymes, fit into the broader hierarchy of processes controlling metabolic flux, from thermodynamics to transcriptional regulation.

Here, we report a computational framework for investigating small molecule regulation across the complete metabolism of an organism. Using *E. coli* as a model, we assemble a genome-scale SMRN by mining the Braunschweig Enzyme Database (BRENDA) (Chang et al., 2009, 2015) and BioCyc database (Caspi et al., 2016). The resulting atlas of small molecule regulation captures widespread inhibition and activation of metabolic enzymes by endogenous metabolites. Overlaying this network onto a genome-scale metabolic model of *E. coli* enables a direct comparison between the topology of metabolism and its regulatory scaffolding. Integrating the SMRN with experimentally determined metabolite concentrations and binding affinities exposes how cells balance between the dual roles of small molecules (i.e., as substrates or inhibitors) as well as their condition-dependent contribution to metabolic flux regulation. Finally, by a natural extension of our approach, we compare the incidence of small molecule regulatory interactions across phylogenetic taxa spanning all kingdoms of life, revealing a handful of canonical regulatory interactions that permeate the metabolism of widely divergent species.

## RESULTS

### Assembling an SMRN

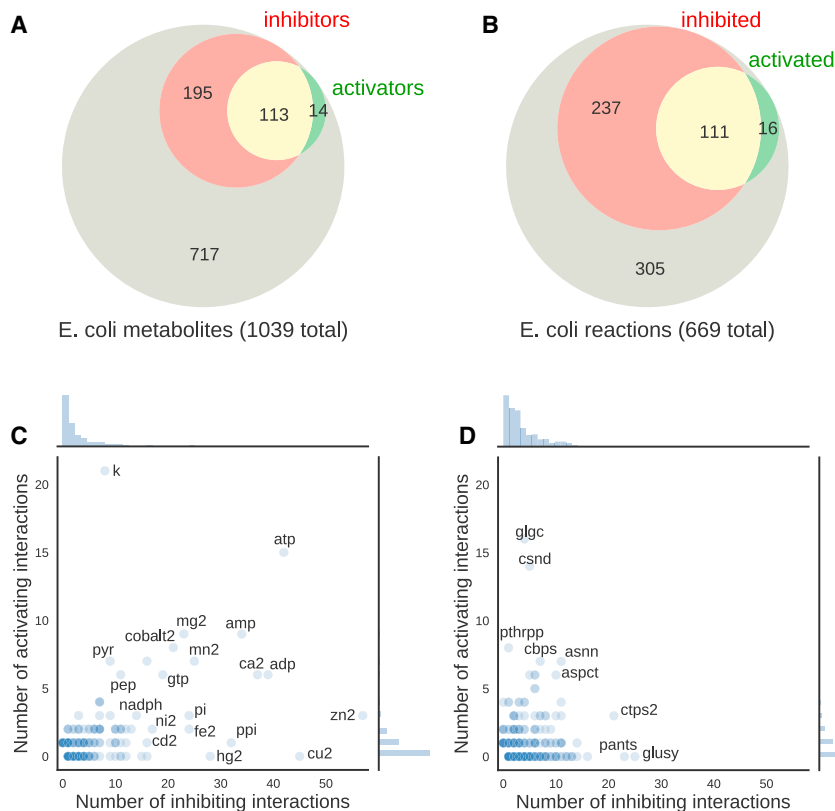
In contrast to the proliferation (Orth et al., 2011; Thiele et al., 2009) and automated reconstruction (Henry et al., 2010) of genome-scale metabolic models, no analogous computational pipeline is available for the analysis of enzyme regulation by small molecules. Here we describe a pipeline for mining existing data on small molecule regulation from public repositories and computational tools for integrating it with a curated, genome-scale metabolic model (Figure 1).

Our general approach relies on the BRENDA and BioCyc database for data concerning the inhibition or activation of enzymes by small molecules in a particular organism. To facilitate comparison across species, information in these databases is organized along Enzyme Commission (EC) numbers, which functionally classify enzymes according to the reaction they catalyze. Thus, for every EC number, we obtained a list of possible regulating small molecules, the type of interaction (activation versus inhibition), and the interaction constant ( $K_I$ ), if available. In addition to gathering data on the presence of a small molecule interaction, we also compiled available information on the Michaelis-Menten constants for the substrates of each metabolic reaction. The computational pipeline is freely available for download at <https://github.com/eladnoor/small-molecule-regulation>.

Because of its well-defined genome, highly curated metabolic network, and heavily studied metabolism, we decided to focus on the model bacterium *E. coli*. The assembled data described above was mapped onto a genome-scale metabolic reconstruction of *E. coli* (Orth et al., 2011), producing an SMRN (Table S1). Importantly, our computational framework can be extended with minimal effort to reconstruct and analyze the SMRN of other organisms besides *E. coli*, provided that a genome-scale metabolic model is available and there is sufficient data in the BRENDA and BioCyc database.

### Landscape of Interactions in the *E. coli* SMRN

The computationally reconstructed *E. coli* SMRN contains 1,669 unique interactions between 321 unique endogenous metabolites and 364 unique enzymes (EC numbers) (Figure 1). The vast majority (83%) of these interactions are inhibitory. Of the ~700 unique EC numbers in the *E. coli* model, about half are regulated by at least one native metabolite in *E. coli*. Similarly, ~320 distinct native metabolites (of the total ~1,000) regulate at least one enzyme (Figures 2A and 2B). Figure 1 provides a tally of interactions and kinetic constants recovered from the BRENDA and BioCyc database. We additionally found that 325 of the 1,669 interactions in the SMRN are supported by 2 or



**Figure 2. Overview of Small Molecule Interactions from the BRENDA**

(A and B) Statistics of the computationally reconstructed *E. coli* SMRN, covering the proportions (A) of 1,039 *E. coli* metabolites that are inhibitors, activators, or both and (B) of 669 *E. coli* enzymatic reactions that are inhibited, activated, or both.

(C and D) Scatterplots depicting the number of activating and inhibiting interactions in which (C) each metabolite and (D) each reaction participates.

more independent literature references, which may be treated in the future as an informatic surrogate for the likelihood that the interaction is functionally relevant *in vivo*.

Certain metabolites and EC numbers participated in exceptionally high numbers of regulatory interactions (Figures 2C and 2D). In particular, the cofactors ATP, AMP, ADP, PI, PPI, NADPH, and GTP together with the metabolites cysteine, pyruvate, and phosphoenolpyruvate (PEP) were the most frequent metabolite regulators, participating in at least 15 interactions (both inhibitory and activating). Notably, ATP was found to regulate 57 different reactions, reflecting its important role as a global reporter of the energetic status of the cell (Atkinson, 1968). Although not strictly molecules, metal ions were found to comprise a significant fraction of the group of small molecule regulators. Interestingly, potassium was the most frequently reported activator, likely reflecting the ability of monovalent cations to activate a broad group of enzymes (Page and Di Cera, 2006). Divalent cations, on the other hand, were among the most recurrent inhibitors, with zinc having more than 50 reported interactions and copper, calcium, manganese, mercury, iron, and magnesium each reported to inhibit more than 20 different reactions.

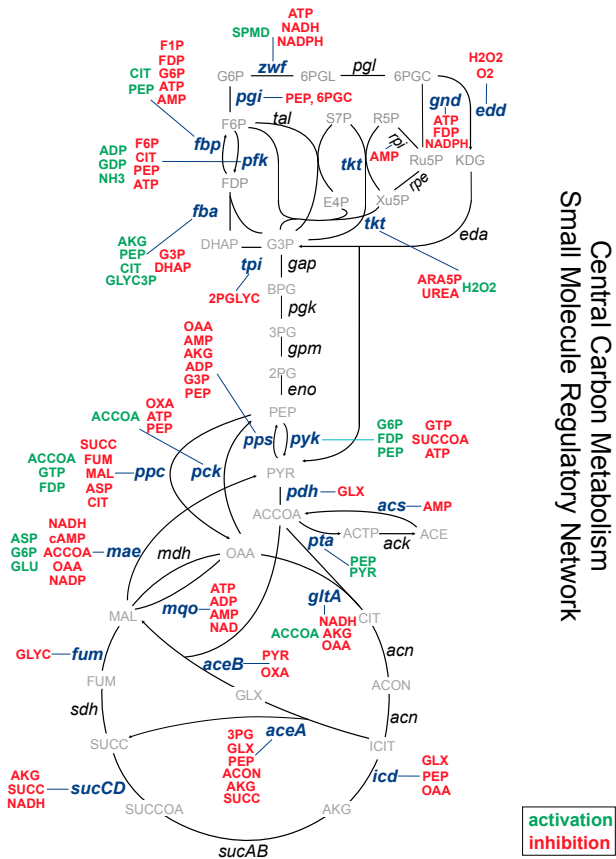
Small molecules can regulate enzymes for which they are not native substrates or products, potentially as a mechanism for long-distance signaling between metabolic pathways without direct connection via their reactants. Therefore, we investigated the distance covered by small molecule regulation (i.e., the shortest graph-distance between each regulatory metabolite and its targeted enzyme), using the genome-scale stoichiometric

network of *E. coli* as a scaffold. As described in the [Experimental Procedures](#), this distance corresponds to the number of reactions a metabolite must traverse in order to reach a target enzyme. We found that enzymes are typically regulated by metabolites that are in their close vicinity. Specifically, 17% of inhibitory interactions are also reactants of the corresponding enzyme, 25% are only one enzymatic step away, and 35% are two steps away. Activating interactions tend to have slightly longer range interactions, with only 8% of activators regulating enzymes that utilize them as substrates (Figure S1).

To obtain specific insight on how small molecule-enzyme interactions operate across different metabolic pathways, we used the genome-scale *E. coli* metabolic model (Orth et al., 2010) to classify each reaction according to functional metabolic subsystems. We found that most interactions in the *E. coli* SMRN seem to target seven main subsystems/pathways: cofactor biosynthesis, alternate carbon metabolism, nucleotide salvage pathway, arginine/proline metabolism, nucleotide biosynthesis, cell envelope biosynthesis, and glycolysis/gluconeogenesis (Figure S2). Interestingly, some canonically high-flux pathways (e.g., the tricarboxylic acid (TCA) cycle and the pentose phosphate pathway) were regulated by comparatively few metabolites. It is possible that such pathways are sparsely regulated by small molecules because they are mostly regulated transcriptionally, as has been recently suggested for TCA cycle genes (Chubukov et al., 2013; Gerosa et al., 2015).

Further mining of this rich dataset revealed pathway-specific preferences for the regulatory targets of certain small molecules. In particular, PEP, citrate, and AMP each activated 3 or more reactions in the glycolysis/gluconeogenesis pathway, suggesting that these three metabolites act as critical sensors controlling the overall rate and direction of glucose metabolism. Similarly, a group of nucleoside triphosphates, deoxynucleoside triphosphates, and adenosine appear to specifically inhibit enzymes in three pathways related to nucleotide metabolism (nucleotide salvage, purine/pyrimidine biosynthesis, and prosthetic group biosynthesis), perhaps reflecting a negative feedback loop for the maintenance of adequate levels of various nucleotides (Figure S2).





**Figure 3. Small Molecule Regulatory Network of *E. coli* Central Carbon Metabolism**

Depiction of the small molecule regulatory interactions in the central carbon metabolism of *E. coli*. Red metabolites are inhibitors and green metabolites are activators of the indicated reactions. Another view of this SMRN is given in Figure S3, showing clearly which reactions are inhibited, activated, or both.

### Design Principles in the Regulation of Central Carbon Metabolism

Central carbon metabolism, encompassing glycolysis, the pentose phosphate pathway, and the TCA cycle, provides all the energetic and biosynthetic precursors for the cell, and it is known to be highly transcriptionally, post-translationally, and allosterically regulated (Chubukov et al., 2014). The central metabolism of *E. coli* is also one of the few parts of metabolism where in vivo evidence is available to support the functional role of small molecule regulation, e.g., in order to induce flux reversal (Link et al., 2013). In silico efforts to model the response of central metabolism to nutrient perturbations, combined with experimental data, have highlighted the fact that our understanding of the intricate regulation of central metabolism is incomplete (Gerosa et al., 2015; Hackett et al., 2016; Kochanowski et al., 2013; Link et al., 2013; Xu et al., 2012a).

The majority of enzymes in *E. coli*'s central carbon metabolism are regulated (Figure 3; Figure S3), and they interact with more small molecules than average in the SMRN (Figure S1), reflecting the heavy research attention these pathways have historically

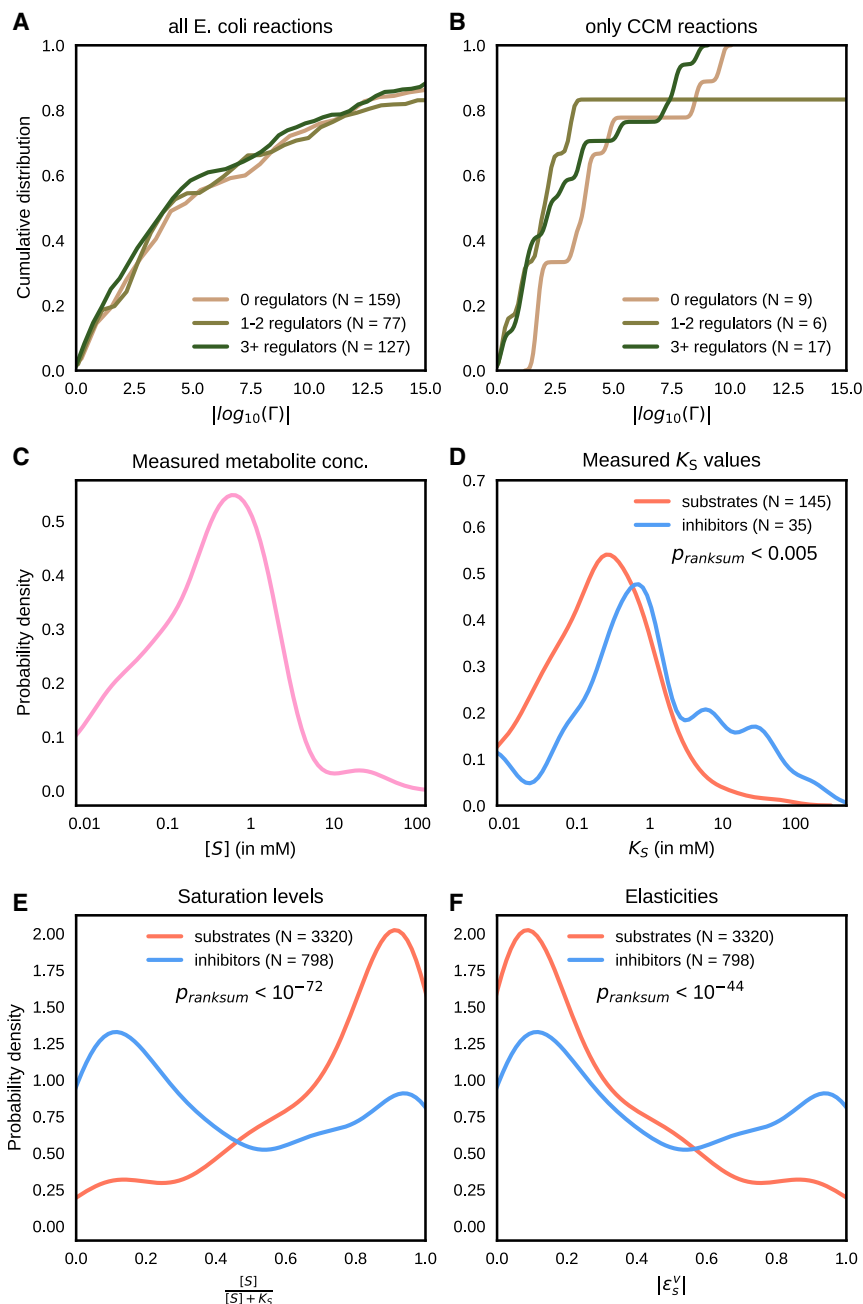
attracted. Interestingly, some of the enzymes in central metabolism are very heavily regulated, specifically those in upper glycolysis (e.g., *fbpase*, *pfk*, and *fba*), terminal glycolysis (*pck*, *ppc*, *pps*, and *pyk*), and the branching reactions of the TCA cycle (*mae*, *aceA*, and *icd*). Conversely, some metabolites seem to have a more central role in certain regions of central metabolism; PEP, for instance, regulates six reactions in glycolysis (*pfk*, *pgi*, *fbpase*, *fba*, *pps*, and *pyk*).

A glance at the structure of small molecule regulation in *E. coli*'s central metabolism strongly suggests that the distribution of regulatory interactions is non-random and has likely been shaped by evolution. What are the pressures selecting for regulatory interactions in *E. coli*'s SMRN? In this regard, the theoretical and experimental literature has proposed a variety of thermodynamic and economic arguments to explain patterns of SMR interactions in central carbon metabolism. Below, we evaluate the consistency of each hypothesis with data from the *E. coli* SMRN.

One frequently cited hypothesis is that small molecule regulation is concentrated in those reactions exhibiting a large drop in free energy (Stryer et al., 2002). To evaluate this possibility, we acquired thermodynamic data for most metabolic reactions using the component contribution method (Noor et al., 2013). Using reactions'  $\Delta G^\circ$  together with reaction stoichiometry and standard physiological metabolite concentrations of substrates or products, we calculated a reversibility index (denoted  $\Gamma$ ) quantifying the extent to which each reaction is thermodynamically reversible (Noor et al., 2012). Using two complementary methods, we did not find the distributions of  $\Gamma$  values for regulated/unregulated reactions in central carbon metabolism to be statistically different ( $p$  value < 0.3, Mann-Whitney U test;  $p$  value = 0.1, gene set enrichment analysis; Figure 4B). The same result was observed when repeating the analysis on all available reactions in *E. coli* ( $p$  value < 0.5, Mann-Whitney U test;  $p$  value = 0.25, gene set enrichment analysis; Figure 4A; Table S4). While many irreversible reactions in central carbon metabolism were indeed regulated by small molecules, a similar proportion of reversible reactions were also regulated. In particular, we found reactions, like 6-phosphogluconolactonase in the PP pathway (EC 3.1.1.31,  $\log_{10}(\Gamma) = 4.17$ ), that do not have any reported small molecule interactions yet exhibit a large drop in free energy. Similarly, many reversible reactions have several regulators, as in the case of succinyl-CoA synthetase (EC 6.2.1.5,  $\log_{10}(\Gamma) = 0.09$ ), which is inhibited by NADH and alpha-ketoglutarate, or glucose-6-phosphate isomerase (EC 5.3.1.9,  $\log_{10}(\Gamma) = 0.44$ ), which is inhibited by PEP and 6PGC. Taken together, our data do not support the hypothesis that thermodynamically irreversible reactions are more likely to be regulated by a small molecule.

A second hypothesis is that cells use small molecule regulation to conserve precious metabolic resources by preventing futile cycling. We observed several examples of interactions in the *E. coli* SMRN supporting this possibility. For example, *fbpase* and *pfk* catalyze opposing reactions, and their simultaneous operation leads to futile cycling (Daldal and Fraenkel, 1983). Two metabolites (citrate and PEP) serve as activators for *fbpase* and as inhibitors for *pfk*, thus curbing this futile cycle (Figure 3). A similar regulatory architecture can be found in the regulation of

Central Carbon Metabolism  
Small Molecule Regulatory Network



**Figure 4. Thermodynamics, Saturation, and Elasticity of Small Molecule Regulatory Interactions in *E. coli***

(A and B) Comparison of reversibility indices ( $\Gamma$ ) (Noor et al., 2012) between regulated and unregulated reactions.  $|\log_{10}(\Gamma)|$  reflects how much freedom the reactants' concentrations need for reversing the flux. For instance, a reaction whose  $|\log_{10}(\Gamma)| > 3$  would require a concentration range of at least  $1:10^3$  (e.g., 30  $\mu\text{M}$ –30 mM) in order to reverse its direction. Physiological constraints typically limit this range to  $10^3$ – $10^4$ . (A) shows the cumulative distribution of  $|\log_{10}(\Gamma)|$  for all reactions in the *E. coli* model for which an equilibrium constant could be computed using component contributions (Noor et al., 2013). This means that 60%–70% of reactions are reversible, with virtually no regard to whether they are regulated or not. The difference between the distribution of  $|\log_{10}(\Gamma)|$  for regulated and unregulated reactions is not significant (Mann-Whitney U test, p value < 0.5). In (B), we focus only on reactions in central carbon metabolism, where the difference is not significant either (p value < 0.3).

(C) The distribution of measured metabolite concentrations, where the peak value is slightly below  $\sim 1$  mM.

(D) The histograms of  $K_M$  and  $K_I$  values are significantly different (Mann-Whitney U test, p value < 0.005). For a more detailed comparison on a single-metabolite basis, see Figure S6.

(E and F) Conversion of measured binding constants to saturation levels using measurements of metabolite concentrations across 13 conditions further highlights the difference between substrates ( $K_M$  values) and inhibitors ( $K_I$  values) (Mann-Whitney U test, p value <  $10^{-72}$ ). When comparing the distributions of scaled elasticities (F), we find that inhibitors have significantly higher values (p value <  $10^{-44}$ ), and they seem to have a bimodal distribution that is split exactly at 0.5. Note that the absolute elasticity value for inhibitors is exactly equal to the saturation, therefore, the blue histogram is the same in (E) and (F). For substrates, however, the elasticity is equal to 1 minus the saturation, so the red histogram in (F) is the mirror image of the one in (E).

the four reactions metabolizing PEP: PEP itself activates pyruvate kinase but inhibits the reverse reaction PEP synthase. Moving beyond anecdotal observations, we used a genome-scale metabolic model (iJO1366; Orth and Palsson, 2012) to detect reactions that can lead to futile cycling. We identified 58 non-overlapping futile cycles, the majority of which comprise only two reactions. Combining this information with the SMRN, we found no statistically significant overrepresentation of small molecule regulation in reactions that take part in futile cycling (see the Supplemental Experimental Procedures). Our finding here needs careful interpretation: the fact that futile cycle reac-

tions are not more likely to be regulated does not necessarily mean that conservation of resources is not important for fitness. For example, it might be that many futile cycles are not regulated because they are in the periphery of metabolism and do not carry high flux, thus making their lack of regulation not very costly for the cell; or, perhaps, most of these futile cycles are avoided by preventing the co-expression of all cycle enzymes simultaneously. Indeed, the fact that *fbpase* and *pfk* are often co-expressed in *E. coli* might be a rare case due to the importance of glycolysis and the need for rapid adaptation of its flux direction.

Besides the prevention of futile cycles, conservation of resources also can be achieved by avoiding wasteful biosynthetic

overproduction. This could be implemented by control of supply and demand of amino acids or nucleotides. In particular, SMR interactions can prevent large fluctuations/instability in the concentrations of biosynthetic end products via feedback inhibition. It is, therefore, often hypothesized that allosteric regulation of the branching reactions from central metabolism leading to amino acid biosynthesis may achieve accurate supply/demand control (Hofmeyr and Rohwer, 2011; Hofmeyr and Cornish-Bowden, 2000) and, thus, prevent unnecessary waste. Interactions in the *E. coli* SMRN are consistent with this hypothesis: analysis of the SMRN revealed that 16 of the 20 amino acids regulate their own biosynthesis using a negative feedback loop, i.e., by inhibiting the first enzyme of their biosynthetic pathway (Figure S4). The remaining non-feedback-inhibitory amino acids (glycine, alanine, aspartate, and glutamate) are 4 of the 5 cheapest ones in terms of the energetic investment required to produce them (Akashi and Gojobori, 2002; Link et al., 2015).

In summary, we found that one of the common (thermodynamic) hypotheses regarding the incidence of small molecule regulation is not supported by data from the *E. coli* SMRN, and we found only anecdotal evidence for an enrichment of regulation in futile cycles. We did, however, find compelling support for an economic role of SMR interactions via feedback inhibition of biosynthesis pathways, at least for 16 of the 20 amino acids. The four exceptions to this rule, however, might prove to be interesting cases where having a regulatory interaction is more costly than the benefit it provides to the cell. In the following sections, we elaborate on this cost, and we quantify it using metabolic control analysis.

### How Is a Metabolite's Role as Regulator and Substrate Balanced?

Small molecule metabolites serve two fundamentally distinct roles in the cell: one, as substrates for metabolic reactions, and another, as regulatory molecules affecting the activity of enzymes and transcription factors (Gerosa and Sauer, 2011). How does the cell balance these two responsibilities, especially in bacteria that have no intracellular compartments that could offer spatial separation (Alam et al., 2017)? More specifically, are the cellular concentrations of these metabolites and the affinities of their interactions with the different enzymes in *E. coli* tuned such that they can inhibit some reactions while efficiently serving as substrates for others? To address this question, we gathered all reported Michaelis-Menten constants ( $K_M$ ) and inhibitory half-saturation constants ( $K_I$ ). We decided to use these  $K_M$  and  $K_I$  values as approximate indicators of a small molecule's metabolic operating point and regulatory operating point, respectively.

We first evaluated whether  $K_M$  and  $K_I$  values were quantitatively similar to each other ( $K_A$  values, or binding constants associated with activating interactions, were not available in the BRENDA and were thus excluded). As reported by others (Park et al., 2016), in general,  $K_I$  values tend to be higher than  $K_M$  values (Figure 4D; Mann-Whitney U test, p value < 0.005). A more informative approach to evaluating differences in  $K_M$  and  $K_I$  values is by a direct comparison with physiological metabolite concentrations. For example, if a metabolite's concentration is much higher than all of its associated  $K_M$  and  $K_I$  values, then

all interactions related to this metabolite are approximately fully saturated and any differences between  $K_M$  and  $K_I$  are not physiologically meaningful. Here, we quantify the level of saturation using the formula  $s/(s+K_S)$ , where  $K_S$  is a binding constant (either  $K_M$  or  $K_I$ ) and  $s$  is the concentration of the substrate or inhibitor. For a substrate, the saturation represents the relative activity of an enzyme compared to a case where  $s$  is very high and the enzyme is fully activated (assuming all other parameters are kept constant and the reaction obeys irreversible, mono-substrate kinetics). We use this definition to identify physiologically relevant differences between substrate and inhibitor affinities.

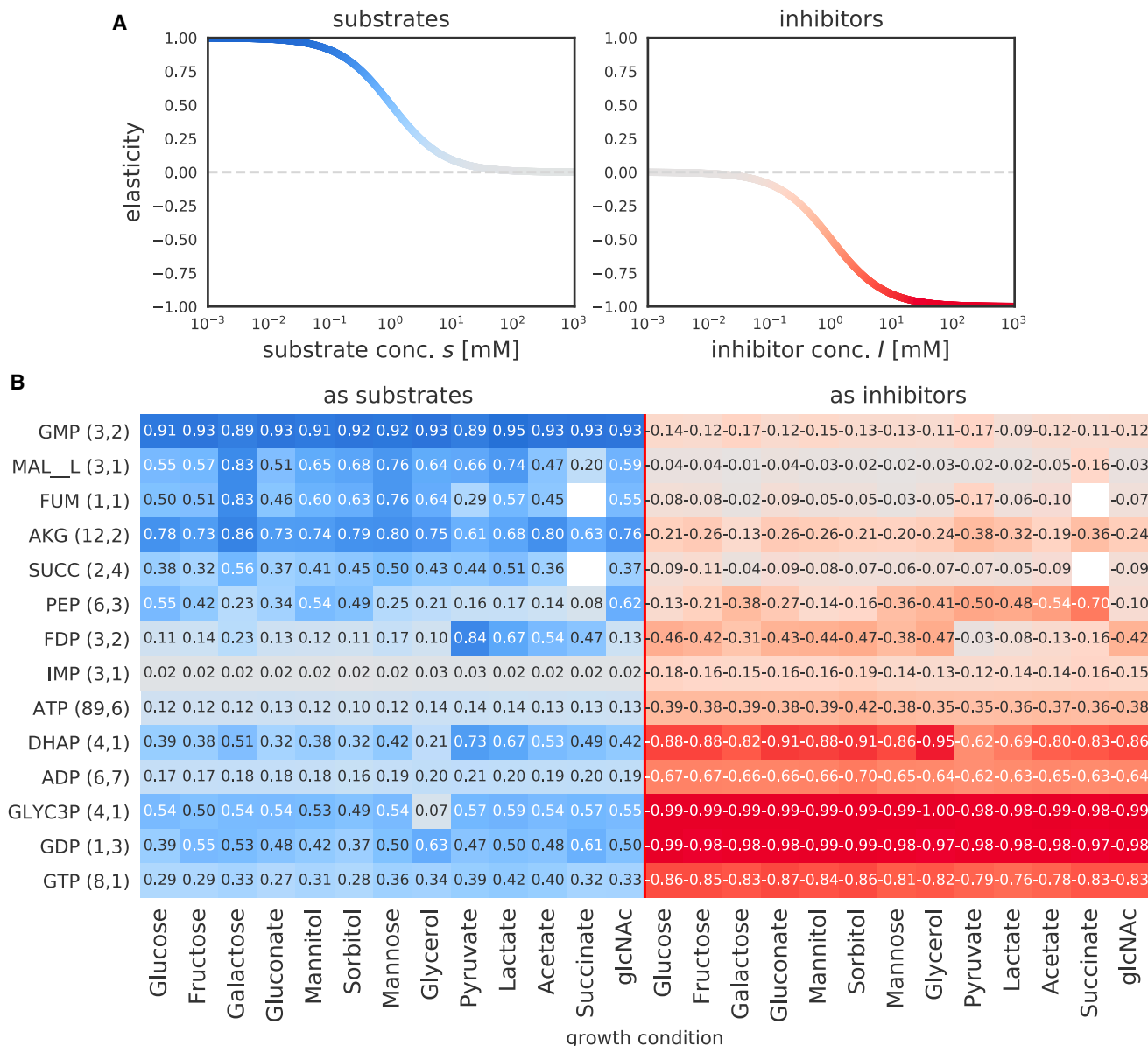
To apply the notion of saturation to our data, we obtained previously published metabolite concentrations in exponentially growing *E. coli* cultures on 13 different carbon sources (Kochanowski et al., 2017). For each unique binding constant/metabolite concentration pair, we calculated the saturation level (332 enzyme-inhibitor-condition triplets and 798 enzyme-substrate-condition triplets; Figure 4E). A comparison between substrate and inhibitor saturation levels yielded a significant difference in the saturation of inhibitor- and substrate-binding sites (Mann-Whitney U test, p value <  $10^{-72}$ ). This suggests that, at physiologically relevant concentrations of metabolites, the majority of substrate binding sites are at or near saturation, while inhibitor sites are occupied but largely far from being saturated, as was reported in the past (Bennett et al., 2009; Park et al., 2016).

### Quantifying the Metabolic Response to Small Molecules across Different Conditions and the Trade-Off between Regulation and Enzymatic Activity

The results presented in Figure 4E indicate that approximately one-third of the inhibitory interactions involved metabolites whose concentrations were higher than the associated  $K_I$  (saturation level at least 0.5). This suggests that many enzymes are operating below their maximal catalytic potential. Why would metabolic enzymes be poised at such a point, well below maximal activity? As we show below and in the Supplemental Experimental Procedures, one possible explanation is a fundamental trade-off that enzymes face: in order to be responsive to the abundance of metabolite regulators, enzymes must sacrifice some of their catalytic activity. Put another way, there is an inherent cost associated with small molecule regulation; to effectively regulate enzyme activity, an inhibitor must be at a sufficiently high concentration. Using theoretical arguments from metabolic control analysis (MCA), we prove that this trade-off between activity and regulation is valid for a general class of kinetic rate laws (e.g., competitive, noncompetitive, and uncompetitive inhibition), applying both to inhibitors and activators (see the Supplemental Experimental Procedures).

The relevant quantity in MCA that leads to the finding above is the scaled elasticity  $\epsilon_s^v$ , which quantifies how fluctuations in the concentration of a metabolite  $S$  affect the rate of reaction  $v$ . For substrates of reactions described by irreversible Michaelis-Menten kinetics, the scaled elasticity is given by  $\epsilon_s^v = (\partial v / \partial s) \cdot (s/v) = 1 - s/(s+K_M)$ , i.e., it is maximized (equal to 1) when the substrate-binding site is unsaturated ( $S \ll K_M$ ). In this regime, a fluctuation in the concentration of substrate leads to a linearly proportional change in the flux of the enzyme (for an isolated enzyme, like in an in vitro assay). On the other





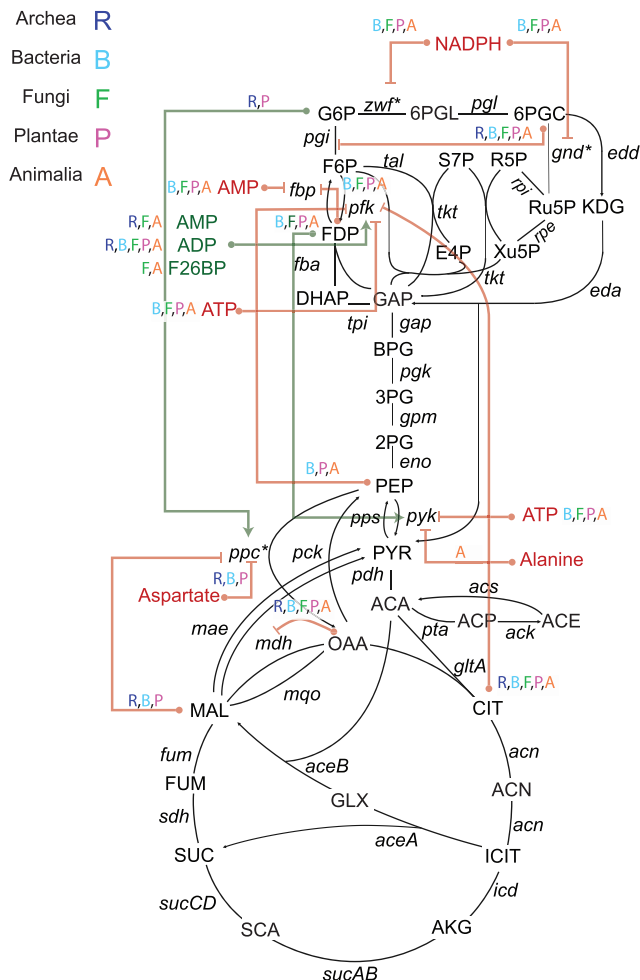
**Figure 5. Functional Role of Small Molecule-Enzyme Interactions in Central Metabolism**

(A) The opposing relationship between elasticity and concentration for a prototypical substrate or inhibitor. In general, substrates have high elasticity at low concentrations, while inhibitors have high elasticity at high concentrations.

(B) A heatmap of the median values of each metabolite's elasticity values across all enzymes that utilize it as a substrate (left-hand side) and across all enzymes inhibited by it (right-hand side). The different columns correspond to different growth conditions (batch growth on minimal media with single carbon sources, samples in mid-exponential phase). Saturation levels were calculated using the formula  $[S]/([S] + K_S)$ , where  $K_S$  is either the Michaelis-Menten coefficient ( $K_M$ ) or the inhibition constant ( $K_I$ ), and, in turn, elasticities were calculated as described in the text. The numbers next to each metabolite in parentheses count the number of different  $K_M$  or  $K_I$  values, respectively, that a metabolite has in our database (for different reactions). If a metabolite has more than one  $K_M$  or  $K_I$  value (i.e., for more than one enzyme), the median of all elasticities is shown. For more details, see Figure S5 and Tables S5, S6, S7, and S8.

hand, the elasticity of non-cooperative, non-competitive inhibitors  $I$  is described by  $e_I^r = (\partial v / \partial I) \cdot (I/v) = -I/(I + K_I)$  (see Figure S5). Counterintuitively, this means that a reaction is most sensitive to an inhibitor  $I$  when  $I$  is at high concentration and the enzyme is strongly inhibited ( $I \gg K_I$ ). Conversely, when the inhibitor concentration is low, its elasticity approaches zero

and the flux cannot respond to changes in  $I$ . Therefore, substrates and inhibitors are subject to opposite quantitative relationships describing their potential to regulate reactions (Figure 5A). Therefore, the saturation level has two complementary consequences: first, it affects the elasticity (how sensitive the reaction rate is to changes of the regulator level), and second,



**Figure 6. Small Molecule Regulation across Kingdoms of Life**  
 The BRENDA was mined for all reports of small molecule regulatory interactions across all species. These interactions were aggregated by unique metabolite-reaction pairs. For each interaction evident in at least 10 different organisms and supported by evidence from at least 10 different published studies, manual curation of the results followed. We identified the broad phylogenetic taxon within which the interaction was present. Nearly all conserved interactions are inhibitory, with three exceptions: the activation of phosphofructokinase by three metabolites (AMP, ADP, and fructose-2,6-bisphosphate), the activation of PEP carboxylase by G6P, and the activation of pyruvate kinase by FDP.

it determines the necessary amount of enzyme (because enzyme that is not used at its full capacity needs to be compensated by higher enzyme levels).

Thus, an alternative way of examining the difference between  $K_M$  and  $K_I$  data is by translating estimates of saturation into elasticities (which we treat as proxies for the metabolic response coefficient; see the Discussion and the Supplemental Experimental Procedures). To simplify the calculation and avoid dependencies between multiple parameters and metabolite concentrations, we assumed that all reactions follow irreversible Michaelis-Menten kinetics and all inhibitors are non-competitive. A detailed examination of elasticities, across the 13 different growth conditions, further revealed the regulatory contribution of a set of central

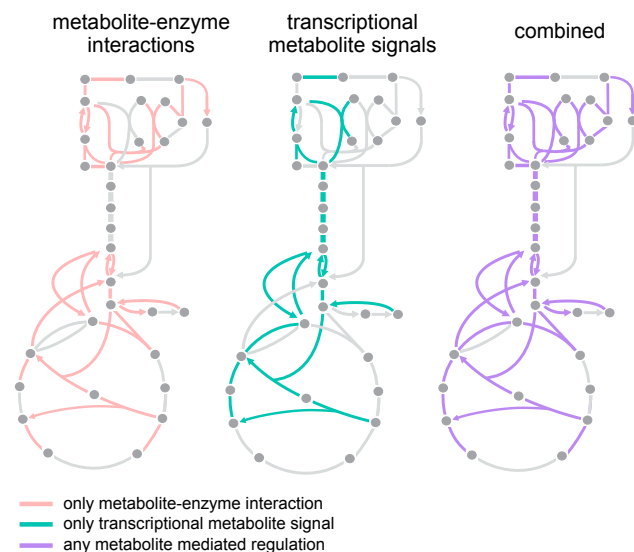
metabolites as substrates and inhibitors (Figures 4F and 5; Tables S5, S6, S7, and S8). Interestingly, metabolites like IMP, ATP, and ADP have very low elasticities as substrates (since they are typically at saturating levels), whereas others (GMP and AKG) have high substrate elasticities spanning from 0.7 to 0.9. As inhibitors, many metabolites are poised between elasticities of 0.2 and 0.5, whereas some metabolites (e.g., GDP and ADP) have high elasticities across all conditions. A higher elasticity could increase the flux response to changes in these metabolites, which is supported by prior work, e.g., it has been reported that ADP strongly regulates PRPP synthase (*prpps*, EC 2.7.6.1) as a form of biosynthetic feedback inhibition (Willemoës et al., 2000), as well as fructose-1-phosphate kinase (*fruk*, EC 2.7.1.56) as a mediator of end product inhibition (Buschmeier et al., 1985).

Notably, the elasticity of several highly connected metabolites in the SMRN (i.e., PEP and FDP) changes substantially between environmental conditions. FDP has high inhibitory elasticities when cells grow in a glycolytic mode (e.g., growth on glucose, fructose, and mannitol), consistent with the proposed role of FDP as a flux sensor in glycolytic conditions (Kochanowski et al., 2013). Interestingly, PEP as an inhibitor has high (absolute) elasticities on gluconeogenic carbon sources (e.g., pyruvate, acetate, and succinate), operating antisymmetrically to FDP. This regulatory design is critical for the adaptation to environmental conditions, for example, for the control of the flux through phosphofructokinase (*pfk*), which PEP is known to inhibit (Fenton and Reinhart, 2009). This inhibition is not needed though when cells grow on glycolytic carbon sources like glucose. Interestingly, we found that, on glycolytic carbons that can only support a slow growth rate, PEP still has high elasticities (e.g.,  $-0.38$  and  $-0.36$  for *pfk* when cells grow on galactose or mannose, respectively).

### Small Molecule Regulation across Kingdoms of Life

While some central pathways of metabolism are nearly ubiquitous, whole-cell metabolism varies substantially among different organisms. At the coarsest resolution, some phylogenetic taxa (e.g., bacteria and plants) can fix inorganic carbon while others (e.g., animals) cannot. On the other hand, the architecture of central carbon metabolism is broadly conserved across all kingdoms of life (Peregrín-Alvarez et al., 2003, 2009). However, there is little understanding of the extent to which small molecule regulatory interactions are conserved across evolutionarily distant taxa. Therefore, we analyzed all available data on small molecule activators and inhibitors available in the BRENDA, stratifying by the species in which the interaction was reported. We mined these data for recurrence of a regulatory interaction between a small molecule and EC number across different species. We then focused on analyzing interactions that were (1) evident in at least 10 different species and (2) supported by at least 10 different reports in the literature. A full list describing the 253 such interactions is available as Table S3.

Because of the high interest in modeling and understanding of flux through central carbon metabolism, we focused our efforts on understanding recurrent regulatory interactions in this pathway (Figure 6). After excluding small ions and non-endogenous metabolites, we identified 18 small molecule regulatory interactions evident broadly across several phylogenetic taxa,



**Figure 7. Combined Architecture of Direct Small Molecule and Indirect Transcriptional Regulation via Endogenous Metabolites in *E. coli***

A map of the reactions in central carbon metabolism that are regulated directly or indirectly by metabolite(s). On the left are reactions that are reported to have at least one metabolite-enzyme interaction. The middle diagram indicates reactions that are indirectly regulated by metabolites via transcription; in each case, the reaction is regulated by transcription factors that are recipients of metabolic signals (i.e., Cra-FDP and Crp-cAMP), as reported in Kochanowski et al. (2017). Some reactions, e.g., those in intermediate glycolysis, are regulated exclusively by transcription. The map on the right overlays small molecule and transcriptional regulation.

which converged on the regulation of a small number of enzymes: four nodes in glycolysis, phosphofructokinase (6 recurrent regulators), fructose biphosphatase (2 recurrent regulators), PEP carboxylase (3 recurrent regulators), and pyruvate kinase (3 recurrent regulators). Surprisingly, we found few conserved interactions in the TCA cycle, glyoxylate shunt, and the nonoxidative branch of the PP pathway. In line with its role as a committing step in glycolysis, phosphofructokinase was subject to negative feedback control by two metabolites far downstream in glucose catabolism, citrate and PEP, as well as regulation by energy-related cofactors ATP, ADP, and AMP. In contrast, pyruvate kinase was negatively regulated by ATP, but also it was subject to feedforward activation by fructose-1,6 diphosphate (FDP). The counterintuitive activation of pyruvate kinase by FDP has been shown (Xu et al., 2012b) to be important for the rapid response of yeast to changes in environmental glucose levels, by driving the accumulation of PEP for future phosphorylation of glucose in glucose-depleted conditions. Similarly, the inhibition of phosphofructokinase by PEP also has been shown to be of critical importance in dynamic perturbations in *E. coli* (Link et al., 2013).

Interestingly, several of the recurrent regulatory interactions we identified were evident in only a subset of phylogenetic taxa. In some cases, this was due to the absence of the enzyme in a taxon (e.g., PEP carboxylase is only present in archaea, bacteria, and plants and the pentose phosphate pathway is not pre-

sent in archaea). In other cases, small molecule regulation was simply different across taxa, with potentially interesting implications. For example, pyruvate kinase was inhibited by L-alanine in animals, but not in other phyla. In humans, this differential regulation plays a role in disease: the inhibition of one splice isoform of pyruvate kinase (PKM2) by L-alanine (but not the other, PKM1) contributes to the cancer-associated shift to aerobic glycolysis by promoting the shunting of glucose-derived carbon into biosynthetic pathways (Morgan et al., 2013).

## DISCUSSION

The regulatory action of small molecules on enzymes and other proteins ensures robust operation of metabolism upon dynamic changes. For central carbon metabolism, metabolites can directly (acting as effectors) or indirectly (acting as signals to transcription factors) regulate the flux of almost all enzymatic reactions (Figure 7). While our understanding of transcriptional and post-translational regulation of metabolism has benefited from advances in sequencing and mass spectrometric technologies, experimental challenges have hindered similar breakthroughs in our understanding of the regulation of enzyme activity by small molecule metabolites (Lindsley and Rutter, 2006). Our approach here has been to leverage the fragmented wealth of published biochemical data to generate an atlas of small molecule regulation informatically, i.e., without performing additional laboratory experiments. Our findings illustrate the dual architecture of small molecule regulation and the underlying metabolic network, allow us to test a common hypothesis about the connection between regulation and thermodynamics, and enable us to compare between metabolite concentrations and their respective binding affinities to target enzymes.

Here we report a resource of experimentally evaluated interactions between endogenous metabolites and enzymes. Importantly, the computational framework we developed is freely available (<https://github.com/eladnoor/small-molecule-regulation>), and it can be readily applied to reconstruct the SMRN of an arbitrary organism of choice, given a suitable genome-scale metabolic model and adequate data in the BRENDA and BioCyc database. Such a resource may guide future implementations of kinetic models and also guide experiments designed to identify novel and functionally relevant *in vivo* metabolite-protein interactions. Most implementations of kinetic models so far do not account for all known small molecule enzyme interactions but rather a subset of them (Khodayari and Maranas, 2016; Millard et al., 2017). Using our SMRN, such kinetic models could be expanded to include all relevant interactions, allowing for system-level evaluation of the topological properties of the system. Our SMRN could also be instrumental in the understanding of the functional role of different small molecule enzyme interactions, in combination with kinetic models. For example, an SMRN may prioritize metabolite-enzyme pairs to be included in a kinetic model based on the number of independent literature reports supporting their existence or based on the elasticity of the interactions in different conditions. In addition, the SMRN and its associated inhibitor constants can be used as prior information to parameterize dynamic models of metabolism, an approach that has proven to be successful in *E. coli* and yeast (Hackett et al., 2016; Link et al., 2013).

A critical shortcoming of our approach is the inherent biases of the BRENDA and Ecocyc data, i.e., the tendency for well-studied enzymes, pathways, and organisms to be overrepresented in these databases. Indeed, one factor in selecting *E. coli* as the organism of choice for reconstruction was the breadth of studies conducted on its metabolism. One approach we took to resolving this issue was identifying putatively high-confidence edges in the SMRN, i.e., those with at least 2 independent literature reports supporting the interaction (Table S2; Figure S7). This portion of the SMRN retains 325 (20%) of the total edges, a figure an order of magnitude larger than the number of edges included in typical kinetic models of metabolism. Furthermore, wherever possible, we performed our analyses on the exhaustively explored sub-network of central carbon metabolism, as well as on the full genome-scale network. For example, in the thermodynamic analysis, we indeed find a weak but statistically significant signal when testing for enrichment of small molecule regulation in irreversible reactions. However, when the analysis is expanded to the full network, we no longer find a statistically significant difference in small molecule regulation between reversible and irreversible reactions. This might suggest that the thermodynamic principle is most relevant for reactions with high fluxes like those in central carbon metabolism and does not apply more generally to the entire metabolic network.

Our theoretical analysis of a wide class of inhibiting and activating small molecules shows that, in general, there is a direct trade-off between the elasticity of an effector and the reduction in activity that is caused by its interaction with the enzyme. Mathematically, if  $\theta$  is the relative activity of the enzyme (e.g.,  $\theta = 1 - x/(x + K_i)$  in the case of non-competitive inhibition), then the absolute scaled elasticity is equal to  $|e_x^v| = 1 - \theta$  (see the Supplemental Experimental Procedures). We show that this relationship between activity and elasticity holds for activators and inhibitors alike. Therefore, using a small molecule effector to regulate a flux always comes with the cost of effectively lowering the activity of the enzyme. This cost of regulation might be the reason why *E. coli* does not have end product inhibition for 4 of its amino acid biosynthesis pathways. Perhaps, the cost of regulation for these metabolically cheap amino acids is larger than the energetic cost of overproduction.

To estimate the level of control imposed by a small molecule, we made several simplifications. First, we assumed that all substrates and inhibitors bind non-cooperatively and that inhibition is not competitive. Second, because a detailed and accurate representation of the kinetic form of each reaction rate law in *E. coli* was unavailable, we estimated elasticities assuming irreversible, mono-substrate kinetics. In the Supplemental Experimental Procedures, we explore the consequences of these assumptions. In particular, we find that substrate elasticities assuming irreversibility are generically upper bounds on reversible substrate elasticities. Furthermore, we show that the elasticity of a substrate in a bi-substrate reaction is (assuming identical kinetic rate constants where applicable) generically a lower bound on the corresponding elasticity, assuming a mono-substrate rate law. Relaxing these assumptions can both strengthen and weaken some of our conclusions, depending on the magnitude of their effect (which itself depends on detailed rate laws and parameters), such as those in Figure 4.

Perhaps more importantly, our results regarding elasticities must be treated carefully when making inferences on metabolic control. According to MCA, a high elasticity is a necessary but not sufficient condition for high control of a flux by a metabolite. We treat interactions with high elasticity as cases where high regulatory capacity is possible; in contrast, interactions with low elasticity are likely to have little regulatory capacity. A trade-off between enzyme activity and elasticity will therefore translate to a trade-off between activity and regulation. To be fully consistent with MCA, one would have to calculate all control coefficients for the inhibited reactions and apply MCA to calculate response coefficients, which quantify the level of flux control of a metabolite, something that is not in the scope of our work. A tractable path forward for decoding the control of flux by the SMRN is to integrate data on changes in *in vivo* metabolite abundance across conditions/perturbations. As we showed in Figure 5, doing so creates a map of the condition-dependent regulatory capacity (i.e., the elasticity) of small molecules, and it prioritizes interactions that appear particularly relevant in a subset of conditions. For canonical interactions with available data on binding affinities, this kind of analysis can be used as a first step in understanding the importance of a regulatory interaction, perhaps by focusing on those interactions with a particularly large change in elasticity between two conditions. For instance, we found that certain metabolites (e.g., fructose bisphosphate) may have a substantially higher inhibition capacity in a subset of conditions. Doing so will likely add to the rich composition of transcriptional, post-translational, and small molecule regulatory interactions that we know to control metabolic flux.

## EXPERIMENTAL PROCEDURES

### Assembly of the SMRN

Raw data were obtained from the BRENDA and BioCyc database. Scripts for parsing the obtained data; filtering unwanted values (i.e., for mutants or negative results); and mapping among BRENDA ligand IDs, Kyoto Encyclopedia of Genes and Genomes (KEGG) identifiers, and BiGG metabolite IDs were written in Python and can all be found on GitHub (<https://github.com/eladnoor/small-molecule-regulation>, available under an MIT license). Likewise, all computations and analyses done for this paper can be found in the same repository.

### Genome-Scale Metabolic Model

All calculations involving the metabolic network of *E. coli* (i.e., distance calculations, thermodynamics of regulated and unregulated reactions, and such) were done using the iJO1366 metabolic model downloaded from the BiGG website (<http://bigg.ucsd.edu/models/iJO1366>) (Orth and Palsson, 2012) and with the help of the cobrapy toolbox (<https://opencobra.github.io/cobrapy/>) (Ebrahim et al., 2013).

### Distance Calculations in the SMRN

First, we removed all the rows corresponding to cofactors from the genome-scale stoichiometric matrix of *E. coli* (the full list of cofactors is provided in the captions of Figure S1). Then, the matrix was converted into an undirected bipartite graph where nodes were either metabolites or reactions. An edge was added between every reaction and all of its substrates and products. Then, the distance between each metabolite and enzyme was calculated by first finding the shortest path between the two on the bipartite graph and counting the number of enzyme nodes along that path (i.e., excluding the metabolite nodes and the target enzyme itself). For example, the distance between an enzyme and one of its substrates is 0.



### Cross-Species Analysis of Small Molecule Regulation

All data regarding activation or inhibition were extracted from the BRENDA. The R package *taxize* (Chamberlain and Szöcs, 2013) was used to recover taxonomic information using the species name provided in the BRENDA. Activating/inhibiting interactions with the same ligand ID (regulating metabolite) and EC number (target enzyme) were aggregated, and the number of such unique entries for each taxonomic group was calculated. Additionally, the number of unique literature references supporting each interaction was recorded, in order to preclude cases where a single interaction reported across multiple species was supported by a small number of independent sources.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and eight tables and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2017.08.066>.

### AUTHOR CONTRIBUTIONS

E.R., D.C., and E.N. designed the analysis, performed the analysis, and wrote the paper. J.E.G. and E.B. performed and contributed to the analysis. U.S. and D.S. supervised the study. All authors read and approved the final manuscript.

### ACKNOWLEDGMENTS

We thank Maren Diether, Hidde de Jong, Wolfram Liebermeister, Hannes Link, and Ruben Mars for fruitful discussions and for helpful suggestions regarding the manuscript. This work was funded by the Swiss Initiative in Systems Biology (SystemsX.ch) TPdF fellowship (2014-230) (to E.N.) and IPdH fellowship (to D.C.). E.B. was supported by the Boston University Undergraduate Research Opportunities Program (UROP). J.E.G. and D.S. were supported by DOE grant DE-SC0012627, MURI grant W911NF-12-1-0390, NIH grant R01GM121950, NSFOCE-BSF grant 1635070, and The Human Frontier Science Program (RGP0020/2016). E.R. was supported by the Sidney Kimmel Center for Prostate and Urologic Cancers, the Marie-Josée and Henry R. Kravis Center for Molecular Oncology, and the MSK Cancer Center Support Grant/Core Grant P30 CA008748.

Received: June 8, 2017

Revised: August 5, 2017

Accepted: August 19, 2017

Published: September 12, 2017

### REFERENCES

Akashi, H., and Gojoberi, T. (2002). Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* *99*, 3695–3700.

Alam, M.T., Olin-Sandoval, V., Stincone, A., Keller, M.A., Zelezniak, A., Luisi, B.F., and Ralser, M. (2017). The self-inhibitory nature of metabolic networks and its alleviation through compartmentalization. *Nat. Commun.* *8*, 16018.

Atkinson, D.E. (1968). The energy charge of the adenylate pool as a regulatory parameter. Interaction with feedback modifiers. *Biochemistry* *7*, 4030–4034.

Bennett, B.D., Kimball, E.H., Gao, M., Osterhout, R., Van Dien, S.J., and Rabinowitz, J.D. (2009). Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* *5*, 593–599.

Buschmeier, B., Hengstenberg, W., and Deutscher, J. (1985). Purification and properties of 1-phosphofructokinase from *Escherichia coli*. *FEMS Microbiol. Lett.* *29*, 231–235.

Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C.A., Keseler, I.M., Kothari, A., Krummenacker, M., Latendresse, M., Mueller, L.A., et al. (2016). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* *44* (D1), D471–D480.

Chamberlain, S.A., and Szöcs, E. (2013). *taxize*: taxonomic search and retrieval in R. *F1000Res.* *2*, 191.

Chandra, F.A., Buzi, G., and Doyle, J.C. (2011). Glycolytic oscillations and limits on robust efficiency. *Science* *333*, 187–192.

Chang, A., Scheer, M., Grote, A., Schomburg, I., and Schomburg, D. (2009). BRENDA, AMENDA and FRENDA the enzyme information system: new content and tools in 2009. *Nucleic Acids Res.* *37*, D588–D592.

Chang, A., Schomburg, I., Placzek, S., Jeske, L., Ulbrich, M., Xiao, M., Sensen, C.W., and Schomburg, D. (2015). BRENDA in 2015: exciting developments in its 25th year of existence. *Nucleic Acids Res.* *43*, D439–D446.

Chen, Z., Rappert, S., and Zeng, A.-P. (2015). Rational design of allosteric regulation of homoserine dehydrogenase by a nonnatural inhibitor L-lysine. *ACS Synth. Biol.* *4*, 126–131.

Christofk, H.R., Vander Heiden, M.G., Harris, M.H., Ramanathan, A., Gerszten, R.E., Wei, R., Fleming, M.D., Schreiber, S.L., and Cantley, L.C. (2008). The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature* *452*, 230–233.

Chubukov, V., Uhr, M., Le Chat, L., Kleijn, R.J., Jules, M., Link, H., Aymerich, S., Stelling, J., and Sauer, U. (2013). Transcriptional regulation is insufficient to explain substrate-induced flux changes in *Bacillus subtilis*. *Mol. Syst. Biol.* *9*, 709.

Chubukov, V., Gerosa, L., Kochanowski, K., and Sauer, U. (2014). Coordination of microbial metabolism. *Nat. Rev. Microbiol.* *12*, 327–340.

Daldal, F., and Fraenkel, D.G. (1983). Assessment of a futile cycle involving reconversion of fructose 6-phosphate to fructose 1,6-bisphosphate during gluconeogenic growth of *Escherichia coli*. *J. Bacteriol.* *153*, 390–394.

Ebrahim, A., Lerman, J.A., Palsson, B.O., and Hyduke, D.R. (2013). COBRApy: COstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.* *7*, 74.

Feng, Y., De Franceschi, G., Kahraman, A., Soste, M., Melnik, A., Boersema, P.J., de Laureto, P.P., Nikolaev, Y., Oliveira, A.P., and Picotti, P. (2014). Global analysis of protein structural changes in complex proteomes. *Nat. Biotechnol.* *32*, 1036–1044.

Fenton, A.W., and Reinhart, G.D. (2009). Disentangling the web of allosteric communication in a homotetramer: heterotropic inhibition in phosphofructokinase from *Escherichia coli*. *Biochemistry* *48*, 12323–12328.

Gerosa, L., and Sauer, U. (2011). Regulation and control of metabolic fluxes in microbes. *Curr. Opin. Biotechnol.* *22*, 566–575.

Gerosa, L., Haverkorn van Rijsewijk, B.R., Christodoulou, D., Kochanowski, K., Schmidt, T.S.B., Noor, E., and Sauer, U. (2015). Pseudo-transition Analysis Identifies the Key Regulators of Dynamic Metabolic Adaptations from Steady-State Data. *Cell Syst.* *1*, 270–282.

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

He, F., Murabito, E., and Westerhoff, H.V. (2016). Synthetic biology and regulatory networks: where metabolic systems biology meets control engineering. *J. R. Soc. Interface* *13*, 20151046.

Henry, C.S., DeJongh, M., Best, A.A., Frybarger, P.M., Lindsay, B., and Stevens, R.L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* *28*, 977–982.

Hofmeyr, J.S., and Cornish-Bowden, A. (2000). Regulating the cellular economy of supply and demand. *FEBS Lett.* *476*, 47–51.

Hofmeyr, J.-H.S., and Rohwer, J.M. (2011). Supply-demand analysis a framework for exploring the regulatory design of metabolism. *Methods Enzymol.* *500*, 533–554.

Khodayari, A., and Maranas, C.D. (2016). A genome-scale *Escherichia coli* kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains. *Nat. Commun.* *7*, 13806.

King, Z.A., Lu, J., Dräger, A., Miller, P., Federowicz, S., Lerman, J.A., Ebrahim, A., Palsson, B.O., and Lewis, N.E. (2016). BiGG Models: A platform for



- integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* *44* (D1), D515–D522.
- Kochanowski, K., Volkmer, B., Gerosa, L., Haverkorn van Rijsewijk, B.R., Schmidt, A., and Heinemann, M. (2013). Functioning of a metabolic flux sensor in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* *110*, 1130–1135.
- Kochanowski, K., Gerosa, L., Brunner, S.F., Christodoulou, D., Nikolaev, Y.V., and Sauer, U. (2017). Few regulatory metabolites coordinate expression of central metabolic genes in *Escherichia coli*. *Mol. Syst. Biol.* *13*, 903.
- Li, X., Wang, X., and Snyder, M. (2013). Systematic investigation of protein-small molecule interactions. *IUBMB Life* *65*, 2–8.
- Lindsley, J.E., and Rutter, J. (2006). Whence cometh the allosterome? *Proc. Natl. Acad. Sci. USA* *103*, 10533–10535.
- Link, H., Kochanowski, K., and Sauer, U. (2013). Systematic identification of allosteric protein-metabolite interactions that control enzyme activity in vivo. *Nat. Biotechnol.* *31*, 357–361.
- Link, H., Christodoulou, D., and Sauer, U. (2014). Advancing metabolic models with kinetic information. *Curr. Opin. Biotechnol.* *29*, 8–14.
- Link, H., Fuhrer, T., Gerosa, L., Zamboni, N., and Sauer, U. (2015). Real-time metabolome profiling of the metabolic switch between starvation and growth. *Nat. Methods* *12*, 1091–1097.
- Millard, P., Smallbone, K., and Mendes, P. (2017). Metabolic regulation is sufficient for global and robust coordination of glucose uptake, catabolism, energy production and growth in *Escherichia coli*. *PLoS Comput. Biol.* *13*, e1005396.
- Morgan, H.P., O'Reilly, F.J., Wear, M.A., O'Neill, J.R., Fothergill-Gilmore, L.A., Hupp, T., and Walkinshaw, M.D. (2013). M2 pyruvate kinase provides a mechanism for nutrient sensing and regulation of cell proliferation. *Proc. Natl. Acad. Sci. USA* *110*, 5881–5886.
- Nikolaev, Y.V., Kochanowski, K., Link, H., Sauer, U., and Allain, F.H.-T. (2016). Systematic Identification of Protein-Metabolite Interactions in Complex Metabolite Mixtures by Ligand-Detected Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* *55*, 2590–2600.
- Noor, E., Bar-Even, A., Flamholz, A., Lubling, Y., Davidi, D., and Milo, R. (2012). An integrated open framework for thermodynamics of reactions that combines accuracy and coverage. *Bioinformatics* *28*, 2037–2044.
- Noor, E., Haraldsdóttir, H.S., Milo, R., and Fleming, R.M.T. (2013). Consistent estimation of Gibbs energy using component contributions. *PLoS Comput. Biol.* *9*, e1003098.
- Orsak, T., Smith, T.L., Eckert, D., Lindsley, J.E., Borges, C.R., and Rutter, J. (2012). Revealing the allosterome: systematic identification of metabolite-protein interactions. *Biochemistry* *51*, 225–232.
- Orth, J.D., and Palsson, B. (2012). Gap-filling analysis of the iJO1366 *Escherichia coli* metabolic network reconstruction for discovery of metabolic functions. *BMC Syst. Biol.* *6*, 30.
- Orth, J.D., Fleming, R.M.T., and Palsson, B.Ø. (2010). Reconstruction and Use of Microbial Metabolic Networks: the Core *Escherichia coli* Metabolic Model as an Educational Guide. *Ecosal Plus* *4*.
- Orth, J.D., Conrad, T.M., Na, J., Lerman, J.A., Nam, H., Feist, A.M., and Palsson, B.Ø. (2011). A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism–2011. *Mol. Syst. Biol.* *7*, 535.
- Page, M.J., and Di Cera, E. (2006). Role of Na<sup>+</sup> and K<sup>+</sup> in enzyme function. *Physiol. Rev.* *86*, 1049–1092.
- Park, J.O., Rubin, S.A., Xu, Y.-F., Amador-Noguez, D., Fan, J., Shlomi, T., and Rabinowitz, J.D. (2016). Metabolite concentrations, fluxes and free energies imply efficient enzyme usage. *Nat. Chem. Biol.* *12*, 482–489.
- Peregrín-Alvarez, J.M., Tsoka, S., and Ouzounis, C.A. (2003). The phylogenetic extent of metabolic enzymes and pathways. *Genome Res.* *13*, 422–427.
- Peregrín-Alvarez, J.M., Sanford, C., and Parkinson, J. (2009). The conservation and evolutionary modularity of metabolism. *Genome Biol.* *10*, R63.
- Reinhard, F.B.M., Eberhard, D., Werner, T., Franken, H., Childs, D., Doce, C., Savitski, M.F., Huber, W., Bantscheff, M., Savitski, M.M., and Drewes, G. (2015). Thermal proteome profiling monitors ligand interactions with cellular membrane proteins. *Nat. Methods* *12*, 1129–1131.
- Savitski, M.M., Reinhard, F.B.M., Franken, H., Werner, T., Savitski, M.F., Eberhard, D., Martínez Molina, D., Jafari, R., Dovega, R.B., Klaeger, S., et al. (2014). Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* *346*, 1255784.
- Schueler-Furman, O., and Wodak, S.J. (2016). Computational approaches to investigating allostery. *Curr. Opin. Struct. Biol.* *41*, 159–171.
- Stryer, L., Berg, J.M., and Tymoczko, J.L. (2002). *Biochemistry*, Fifth Edition (W.H. Freeman & Co Ltd.).
- Thiele, I., Jamshidi, N., Fleming, R.M.T., and Palsson, B.Ø. (2009). Genome-scale reconstruction of *Escherichia coli*'s transcriptional and translational machinery: a knowledge base, its mathematical formulation, and its functional characterization. *PLoS Comput. Biol.* *5*, e1000312.
- Willemoës, M., Hove-Jensen, B., and Larsen, S. (2000). Steady state kinetic model for the binding of substrates and allosteric effectors to *Escherichia coli* phosphoribosyl-diphosphate synthase. *J. Biol. Chem.* *275*, 35408–35412.
- Xu, Y.-F., Amador-Noguez, D., Reaves, M.L., Feng, X.-J., and Rabinowitz, J.D. (2012a). Ultrasensitive regulation of anapleurosis via allosteric activation of PEP carboxylase. *Nat. Chem. Biol.* *8*, 562–568.
- Xu, Y.-F., Zhao, X., Glass, D.S., Absalan, F., Perlman, D.H., Broach, J.R., and Rabinowitz, J.D. (2012b). Regulation of yeast pyruvate kinase by ultrasensitive allostery independent of phosphorylation. *Mol. Cell* *48*, 52–62.