



Modeling the contribution of allosteric regulation for flux control in the central carbon metabolism of *E. coli*

Daniel Machado^{1*}, Markus J. Herrgård² and Isabel Rocha¹

¹ Centre of Biological Engineering, University of Minho, Braga, Portugal, ² The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark

OPEN ACCESS

Edited by:

Firas H. Kobeissy,
University of Florida, USA

Reviewed by:

Osbaldo Resendis-Antonio,
Instituto Nacional de Medicina
Genómica (INMEGEN), Mexico

Jeffrey Varner,
Purdue University, USA

*Correspondence:

Daniel Machado,
Centre of Biological Engineering,
University of Minho,
Braga 4710-057, Portugal
dmachado@deb.uminho.pt

Specialty section:

This article was submitted to Systems
Biology, a section of the journal
Frontiers in Bioengineering and
Biotechnology

Received: 16 July 2015

Accepted: 22 September 2015

Published: 08 October 2015

Citation:

Machado D, Herrgård MJ and Rocha
I (2015) Modeling the contribution
of allosteric regulation for flux
control in the central carbon
metabolism of *E. coli*.
Front. Bioeng. Biotechnol. 3:154.
doi: 10.3389/fbioe.2015.00154

Modeling cellular metabolism is fundamental for many biotechnological applications, including drug discovery and rational cell factory design. Central carbon metabolism (CCM) is particularly important as it provides the energy and precursors for other biological processes. However, the complex regulation of CCM pathways has still not been fully unraveled and recent studies have shown that CCM is mostly regulated at post-transcriptional levels. In order to better understand the role of allosteric regulation in controlling the metabolic phenotype, we expand the reconstruction of CCM in *Escherichia coli* with allosteric interactions obtained from relevant databases. This model is used to integrate multi-omics datasets and analyze the coordinated changes in enzyme, metabolite, and flux levels between multiple experimental conditions. We observe cases where allosteric interactions have a major contribution to the metabolic flux changes. Inspired by these results, we develop a constraint-based method (arFBA) for simulation of metabolic flux distributions that accounts for allosteric interactions. This method can be used for systematic prediction of potential allosteric regulation under the given experimental conditions based on experimental data. We show that arFBA allows predicting coordinated flux changes that would not be predicted without considering allosteric regulation. The results reveal the importance of key regulatory metabolites, such as *fructose-1,6-bisphosphate*, in controlling the metabolic flux. Accounting for allosteric interactions in metabolic reconstructions reveals a hidden topology in metabolic networks, improving our understanding of cellular metabolism and fostering the development of novel simulation methods that account for this type of regulation.

Keywords: metabolism, systems biology, constraint-based modeling, allosteric regulation, *Escherichia coli*

1. Introduction

Mathematical models of metabolism have become a fundamental tool for understanding cellular behavior and for designing genetic or environmental modifications to change that behavior toward a specific purpose (Heinemann and Sauer, 2010). Metabolic models have found applications in both biomedical research and industrial biotechnology. Examples of applications in biomedicine include using metabolic models of human cells to analyze the altered behavior of cancer cells and to suggest potential drug targets (Folger et al., 2011). In the context of industrial biotechnology, models of

microbial metabolism are widely used for rational design of microbial cell factories (Zomorodi et al., 2012).

There are two major approaches for modeling cellular metabolism, namely, kinetic modeling and constraint-based modeling (Machado et al., 2012). The former, based on kinetic rate laws, requires extensive experimental data for determination of the enzymatic mechanisms and respective kinetic parameters. For that reason, these models have been limited to central pathways of well-studied organisms, such as *Escherichia coli* and *Saccharomyces cerevisiae* (Teusink et al., 2000; Chassagnole et al., 2002). Constraint-based modeling, on the other hand, only accounts for the stoichiometry and directionality of biochemical reactions, which can be obtained from genome annotations and limited other information for the organism (Bordbar et al., 2014). With the increasing number of fully sequenced genomes for multiple organisms, the number of genome-scale metabolic reconstructions suitable for constraint-based modeling is also rapidly increasing, with over a hundred reconstructions currently available (Monk et al., 2014).

Constraint-based models can be used to estimate the steady-state flux distribution of a metabolic network, using the so-called Flux Balance Analysis (FBA) approach (Orth et al., 2010). Since the flux solution is not unique with only stoichiometric and directionality constraints, in FBA a single solution is selected based on the assumption of an evolutionary principle of optimality, such as maximization of cellular growth. Methods have been developed to refine metabolic flux predictions by integration of metabolic models with models of other biological processes, such as signaling and transcriptional regulatory networks (Gonçalves et al., 2013). However, some limitations of these methods, such as the reduction of gene expression levels to Boolean states, hamper the predictive ability of the integrated models. More recently, several approaches were developed to directly integrate gene expression data into metabolic models. These methods are based on the assumption that reaction fluxes should be proportional to their respective gene expression levels. However, a recent systematic evaluation of these methods showed little improvement in simulation accuracy when gene or protein expression data are used for flux prediction with a wide range of proposed methods (Machado and Herrgård, 2014). One of the conclusions from this study is that the assumption of proportionality between gene expression levels and reaction rates is not valid for many reactions.

The conclusion that transcriptional or translational regulation does not significantly regulate metabolic fluxes is consistent with recent experimental observations in multiple organisms showing that central carbon metabolism is mostly regulated at post-transcriptional levels (Daran-Lapujade et al., 2007; Chubukov et al., 2013; Kochanowski et al., 2013a). Regulation analysis is a method introduced by ter Kuile and Westerhoff (2001) for quantitatively decomposing flux regulation into *hierarchical* and metabolic coefficients. The former accounts for transcriptional and translational regulation as well as post-translational modifications, whereas the latter accounts for allosteric regulation and thermodynamics. The application of this method to three parasitic protists showed that regulation of glycolytic fluxes is never completely hierarchical, being mostly metabolic in many cases. Similar conclusions were obtained by applying this method

to *S. cerevisiae*, where it was observed that metabolic regulation contributed to 50–80% of the flux change in glycolytic enzymes for the given cultivation conditions (Daran-Lapujade et al., 2007).

The partial contribution of transcriptional regulation for flux control in central carbon metabolism can be explained by the cellular trade-off between lowering the investment of protein synthesis (keeping enzymes saturated), and the need to achieve fast regulatory responses and maintain metabolic homeostasis under environmental changes (Fendt et al., 2010; Wessely et al., 2011). In fact, metabolite measurements in *E. coli* and *S. cerevisiae* have shown that most enzymes in central carbon metabolism are not saturated, with substrate levels being close to their respective K_M values (Bennett et al., 2009; Fendt et al., 2010). A recent study in *B. subtilis* showed that transcriptional regulation is insufficient to explain the observed flux change for growth in different carbon sources (Chubukov et al., 2013). Interestingly, the authors observed that the changes in substrate concentrations were also insufficient to explain the observed flux change, leaving an important contribution for post-translational modifications and allosteric regulation.

Learning how allosteric regulation controls the metabolic flux is fundamental for understanding cellular metabolism. Given the growing scope of the constraint-based modeling approach, we propose to expand this formalism with an explicit representation for allosteric interactions. In this work, we build a constraint-based model of allosteric regulation in the central carbon metabolism of *E. coli* and use it to analyze the role of this type of regulation for controlling the metabolic flux under different perturbations.

Allosteric information data are collected from relevant databases and used to build a constraint-based model expanded with allosteric interactions. We analyze how this new layer of interactions affects the network topology in terms of node connectivity and identify relevant metabolic hubs. The model is used as a scaffold to perform regulation analysis using multiple omics data for *E. coli*. Finally, a new method for constraint-based simulation accounting for allosteric interactions is proposed and used for model-based prediction of regulatory effects on flux control.

2. Results

2.1. Model Reconstruction

In order to analyze the effects of allosteric regulation in the central carbon metabolism, we expanded a constraint-based model of the core metabolism of *E. coli* (Orth et al., 2009) with allosteric interactions obtained from relevant sources (see Figure S3 in Supplementary Material and Methods section for details). The expanded model is presented in **Figure 1**. It can be observed that the integration of regulatory interactions reveals an intricate topology that is not captured by the stoichiometric reconstruction alone. In this case, the connections represent signal flow rather than mass flow. Much like in the case of signaling pathways, it is possible to observe a highly complex *crossstalk* between different subpathways. This includes multiple feedback links between upper and lower glycolysis, upper glycolysis and the oxidative part of the pentose-phosphate (PP) pathway, lower glycolysis and the TCA cycle, and a positive feedback link from citrate to upper glycolysis.

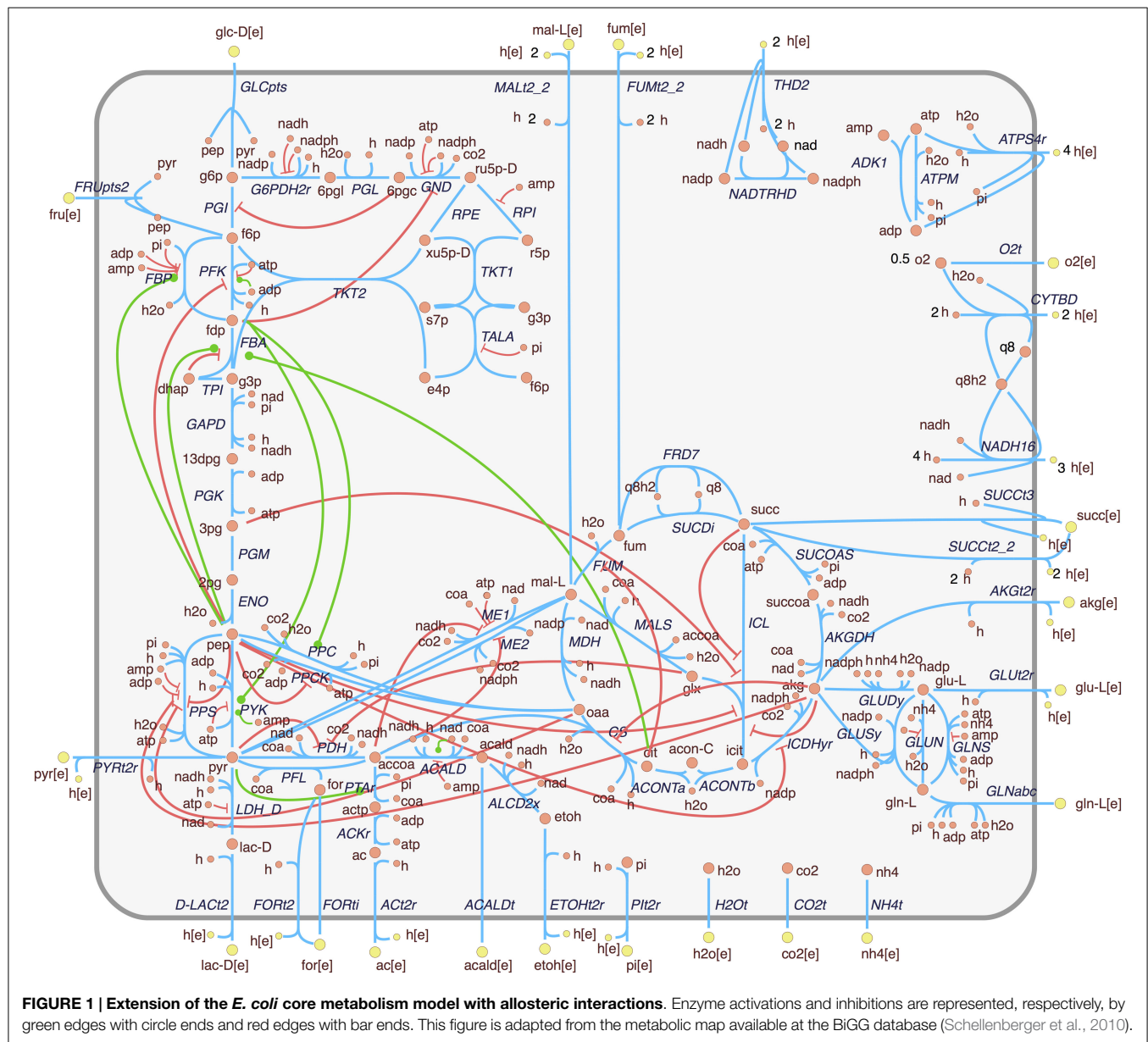


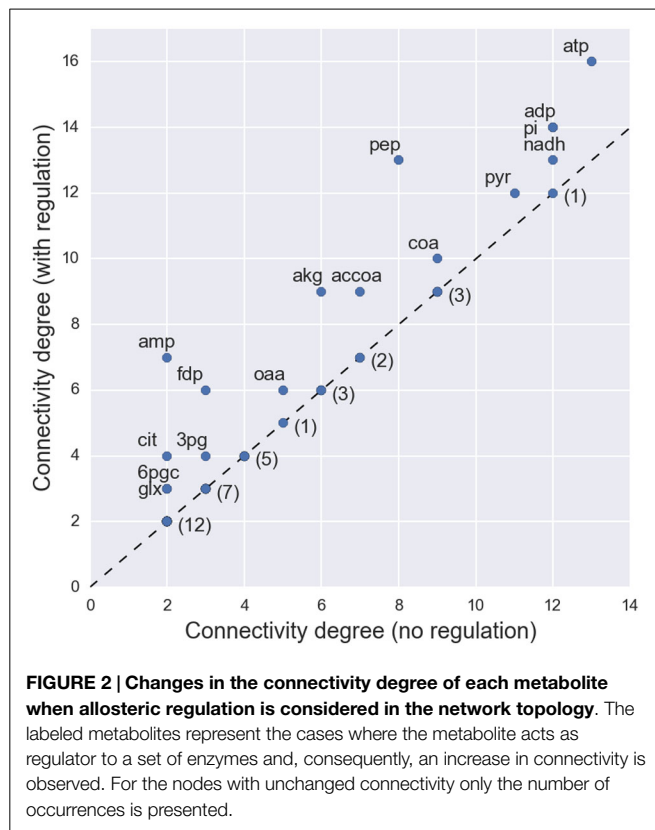
Figure 1 shows that most regulatory interactions are inhibitory. It is possible that some of these inhibitory interactions are competitive rather than allosteric (i.e., the binding site of the effector coincides with the catalytic site). Since the binding mechanisms are not generally reported in the databases, and the regulatory effect is similar, this distinction will be disregarded for the purpose of this work.

Topological analysis in terms of connectivity degree shows an increased connectivity for several metabolites when allosteric regulation is considered (**Figure 2**). However, the median value of connectivity remains the same (4 connections per metabolite). Unsurprisingly, there is an increased connectivity for metabolites that were previously known metabolic hubs. For instance, phosphoenolpyruvate (*pep*) is now connected to a total of 13 reactions (previously 8), reinforcing the importance of this glycolytic compound as a metabolic

hub (Link et al., 2013; Matsuoka and Shimizu, 2015). However, changes are also observed for lowly connected metabolites. A notable case is *fructose-1,6-bisphosphate* (*f6p*), which can now be considered as a hub metabolite (with a total of 6 connections), although its connectivity is below the median if regulation is not considered. This metabolite was recently identified as a key flux-signaling metabolite in the glycolytic flux-sensing mechanism of *E. coli* (Kochanowski et al., 2013b).

2.2. Omics Data-Based Analysis of Allosteric Regulation

In order to understand how the coordination between hierarchical and metabolic regulation drives the metabolic flux, we used the reconstructed model to integrate and analyze a multi-omics dataset for *E. coli* (Ishii et al., 2007). This dataset contains transcript, protein, metabolite, and flux data for *E. coli* strains



growing aerobically in a chemostat. It comprises several experiments, including variations of dilution rate for the wild-type strain ($0.1\text{--}0.7\text{ h}^{-1}$) and 24 single knockout mutants growing at the reference dilution rate (0.2 h^{-1}). Herein, we will refer to the wild-type strain growing at 0.2 h^{-1} as the reference condition, and the remaining as the perturbed conditions (28 in total).

The data were analyzed using the concept of *regulation analysis* introduced by ter Kuile and Westerhoff (2001) to decompose the contribution of hierarchical (ρ_h) and metabolic (ρ_m) control coefficients during flux change between two experimental conditions ($\rho_h + \rho_m = 1$). We applied the generalization proposed in Chubukov et al. (2013) to simultaneously compare multiple conditions (see Methods). This generalization assumes that the coefficients are conserved across conditions. The results are presented in Figure S4 in Supplementary Material. It can be observed that in many cases the slopes are close to zero or even negative, indicating poor evidence of transcriptional control. Only three reactions (*PGI*, *CS*, *FUM*) present an estimated hierarchical control coefficient above 0.5. Hence, only these reactions are likely to be predominantly regulated at the transcriptional level.

Given the lack of evident hierarchical control for most enzymes, one can try to analyze the allosteric control exerted by single effectors in a similar fashion (see Methods). The results are presented in Figure S5 in Supplementary Material. In order to observe active flux control, positive slopes would be expected for enzyme activators and negative slopes for enzyme inhibitors. However, this behavior can only be observed in a few cases. The flux of *FBA* positively correlates with its two activators, citrate and *pep*.

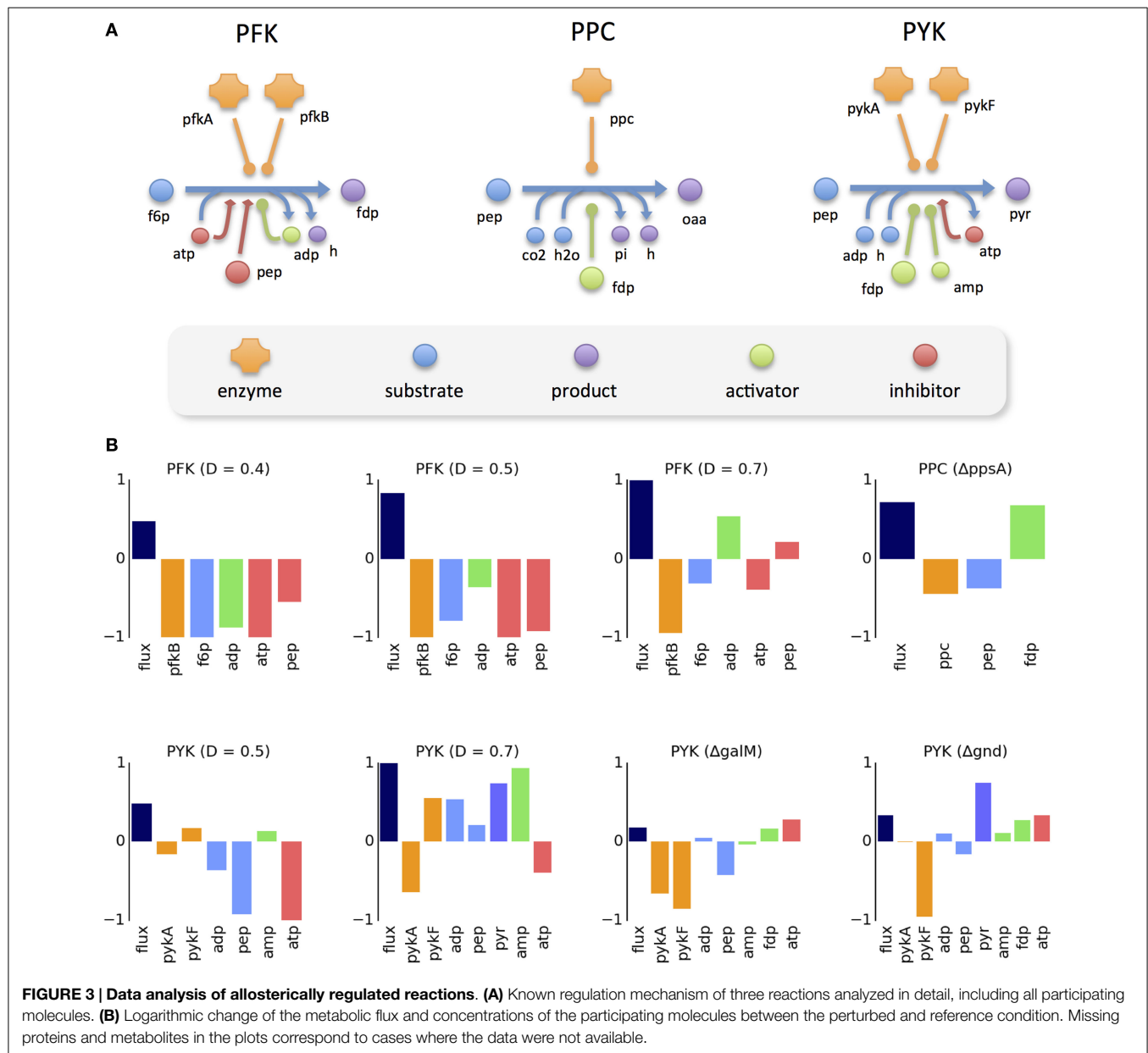
Some correlation is also observed between ATP levels and two of its inhibition targets, *GND* and *PFK*.

Given the large number of reactions without evident transcriptional or allosteric control, we hypothesize that the assumption of constant control coefficients across all conditions does not hold for the given experimental conditions. It is likely that, during different perturbations, different kinds of control are predominant for each reaction. This has also been observed in previous studies in *S. cerevisiae* (Rossell et al., 2006).

We analyzed the flux change for each reaction at each perturbed condition individually, by comparing the logarithmic change of enzyme, flux, and metabolite levels between all 28 perturbed conditions and the reference condition. Although this would result in a total of 672 potential case studies (24 regulated reactions times 28 perturbations), due to the sparsity of the data (especially the metabolome data), this study was restricted to all reaction-condition pairs with sufficient data to perform a meaningful analysis (see Methods). This reduced the number of case studies to 38 (see Figure S6 in Supplementary Material for details). We then analyzed the evidence of allosteric control for these cases (see Methods) and observed a total of 8 cases where allosteric regulation seems to play a role in controlling the reaction flux for the given perturbation (Figure S6 in Supplementary Material). These 8 cases will be analyzed in detail below.

The regulation mechanisms of the three reactions involved (*PFK*, *PPC*, and *PYK*) are depicted in Figure 3A. The intricate regulation of these enzymes is evident, in particular for *PFK* and *PYK*, which are catalyzed by multiple isozymes and regulated by multiple effectors. The logarithmic change of flux and all measured intervening molecules for the selected reaction-condition pairs is presented (Figure 3B). It can be observed that, in most cases, the change in enzyme concentration is in the opposite direction of the flux change. For *PFK*, only one of the isozymes is measured. In the case of *PYK*, where both isozymes are measured, it can be observed that the level of one isozyme increases while the other decreases. In the few cases where the flux change follows the direction of the enzyme level, the magnitude of enzyme change is still insufficient to explain the flux change (since the reaction rate would be directly proportional to the enzyme concentration). Regarding the change in substrate levels, it can be observed that, in most cases, it is also opposite to the direction of the flux change.

The effect of allosteric control is evident in some scenarios. For instance, in the $\Delta ppsA$ mutant, the flux of *PPC* largely increases, despite the decrease of its only enzyme (*ppc*) and its main substrate (*pep*). This increase can be explained by the increased concentration of its allosteric activator (*fdp*). There are cases where the different allosteric regulators have a cooperative effect in flux control (e.g., *PYK* at 0.7 h^{-1}) and cases where there is a competing effect (e.g., *PFK* at 0.4 h^{-1}). One can observe that flux change is not always controlled by the same combination of effectors. For instance, at high dilution rates ($0.4\text{--}0.5\text{ h}^{-1}$) the flux of *PFK* increases with the decrease of its inhibitors (ATP and *pep*), despite the decrease of its activator (ADP). However, at an even higher dilution rate (0.7 h^{-1}), the flux increase coincides with higher levels of the activator, whereas the two inhibitors change in opposite directions.



The interpretation of the results is hampered by the lack of protein and metabolite measurements for many experimental conditions. One cannot exclude the possibility that some flux changes are also driven by changes in unmeasured isozymes, cofactors, or reaction products.

2.3. Model-Based Prediction of Allosteric Regulation

Given the scarcity of multi-omics datasets with all the data required to perform a quantitative analysis of allosteric regulation, we developed a constraint-based approach for model-based predictions. This method is based on the assumption that, if a reaction is activated (respectively, inhibited) by a compound present in a pathway, then its flux change should be positively (respectively, negatively) correlated with the flux change in that

pathway (see Supplementary Material for details). It has been proposed that allosteric intermediates function as flux-signaling metabolites that directly translate flux information to metabolite concentration (Kotte et al., 2010; Matsuoka and Shimizu, 2015). The method, named allosteric regulation FBA (arFBA), is a variation of parsimonious FBA (pFBA) (Lewis et al., 2010) where the objective function is extended as follows:

$$\min_{\nu} \sum_i |v_i| + \sum_{R_{ij}>0} w_{ij} \left| \frac{v_j}{v_j^0} - \frac{t_i}{t_i^0} \right| + \sum_{R_{ij}<0} w_{ij} \left| \frac{v_j}{v_j^0} + \frac{t_i}{t_i^0} - 2 \right|.$$

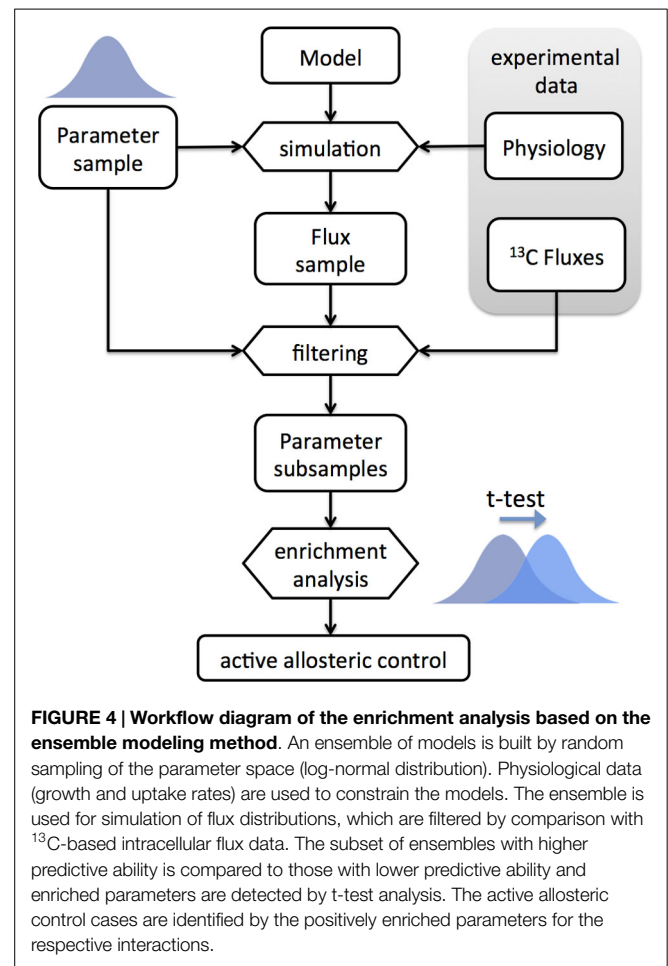
Here, ν is the flux distribution to be estimated, v^0 is the flux distribution for a given reference condition, t_i is the turnover rate of metabolite i . The allosteric interactions are represented in a new matrix R , which has a structure similar to the stoichiometric

matrix, with $R_{ij} = 1$ (respectively, -1) if metabolite i activates (respectively, inhibits) reaction j , and 0 otherwise (note that the stoichiometric matrix S is not changed). The w_{ij} parameters are arbitrary weights that represent the strength of the interaction between effector i and reaction j . If all w_{ij} are close to zero, then the method defaults to a simple pFBA simulation. The minimization of the extra terms in the objective function affects the respective fluxes when regulation is active. For an activation, the subtraction forces the flux and turnover ratios to be the same. For an inhibition, the term forces that a change in the turnover is compensated by an opposite change in the flux. A detailed justification for these terms is given in the Supplementary Material. The full implementation of the method is slightly more complex due to the presence of reversible reactions and reactions without flux in the reference condition (see Supplementary Material for a complete description).

In general, it is not possible to know the strength of the allosteric interactions beforehand. Therefore, we implemented an ensemble modeling approach in order to find the most plausible models (Figure 4). The approach is similar, albeit different, to the ensemble modeling approach used for kinetic modeling (Tran et al., 2008). A model ensemble was built by randomly sampling the w_{ij} parameters (see Methods). The simulated flux distributions are then compared with the intracellular flux data from Ishii et al. (2007). The accuracy of each model is given by the (L_1 -norm) distance between the predicted and measured flux distributions. The original ensemble is split into two groups containing the models with prediction accuracy above and below the median. We then perform enrichment analysis by comparing the distributions of each parameter between the two ensembles. For a particular experimental condition, if a parameter w_{ij} has systematically higher values in the ensemble with higher predictive accuracy, then the assumption of allosteric control between effector i and reaction j results in improved flux predictions for that condition.

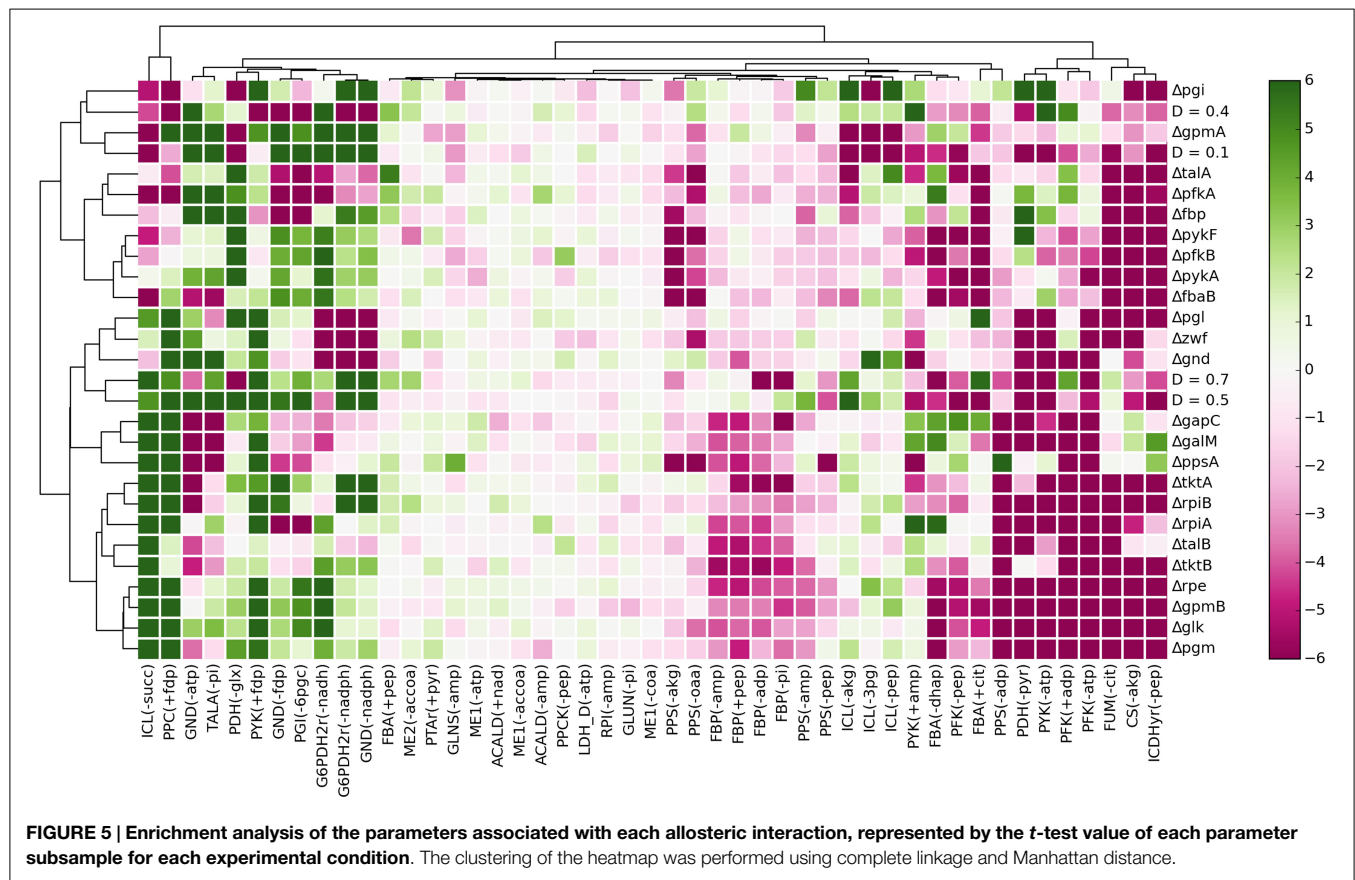
Figure 5 shows t -test values for all parameters across all experimental conditions. Although there are not clearly defined clusters in the clustered heatmap, some general patterns can be observed. About one-quarter of the interactions are positively enriched for most experimental conditions, representing probable cases of active allosteric control for those conditions. On the other hand, almost half of the parameters are negatively enriched for a majority of conditions. These represent allosteric constraints that, in most cases, hamper the predictive ability of the models. Finally, there is a subset of allosteric interactions which are neither positively nor negatively enriched. Accounting for these interactions has very little effect in the prediction of flux distributions for the given experimental conditions.

The most frequent positively enriched interactions include inhibition of the oxidative phase of the pentose-phosphate pathway (PPP) by reducing agents NADH and NADPH; mutual inhibition between PPP and upper glycolysis; feedforward activation of PPC and PYK by *fdp*; and inhibition of the glyoxylate shunt by multiple effectors. Interestingly, several parameters that are positively enriched for a subset of conditions are also negatively enriched for some of the remaining conditions. Hence, although the respective interactions improve the flux predictions in some conditions, in other conditions they make predictions worse.



In order to test the predictive ability of our *in silico* approach, we analyzed the enrichment results for the potential cases of allosteric control previously detected by data-driven analysis (Figure 3B). Some of the allosteric interactions were significantly enriched, namely the activation of *PFK* by *ADP* at the highest dilution rate ($t = 4.28$, $p = 1.88\text{e-}5$), activation of *PPC* by *fdp* in the $\Delta ppsA$ mutant ($t = 19.0$, $p = 8.17\text{e-}79$), and activation of *PYK* by *fdp* in the Δgnd mutant ($t = 4.70$, $p = 2.63\text{e-}6$) and the $\Delta galM$ mutant ($t = 7.09$, $p = 1.44\text{e-}12$).

It should be noted that we are using our simulation method (arFBA) in the reverse direction, i.e., a model ensemble is compared with experimental data to find which parameters (weighting factors) result in improved predictions. Although, in theory, one could use the method in the forward direction, i.e., to perform simulations with improved flux predictions, this would require finding a “universal” parameter configuration that fits all conditions. The previous results show that such universal configuration cannot be found due to the condition-specific nature of allosteric regulation. Nonetheless, we tested the accuracy of arFBA by measuring the distance between simulated and experimental flux distributions. Figure 6 shows the frequency distribution of the distances obtained by random sampling of the weighting factors for each experimental condition. The distance obtained with FBA is shown for comparison. It can be observed that,



for most experimental conditions, the average distance obtained with arFBA is lower than that obtained with FBA, indicating a higher accuracy of the former. Finally, we tested the accuracy of arFBA with *a posteriori* calibration of the weighting factors (see methods). It can be observed that, after calibration, the accuracy of arFBA is higher than FBA for 26 of the 28 conditions.

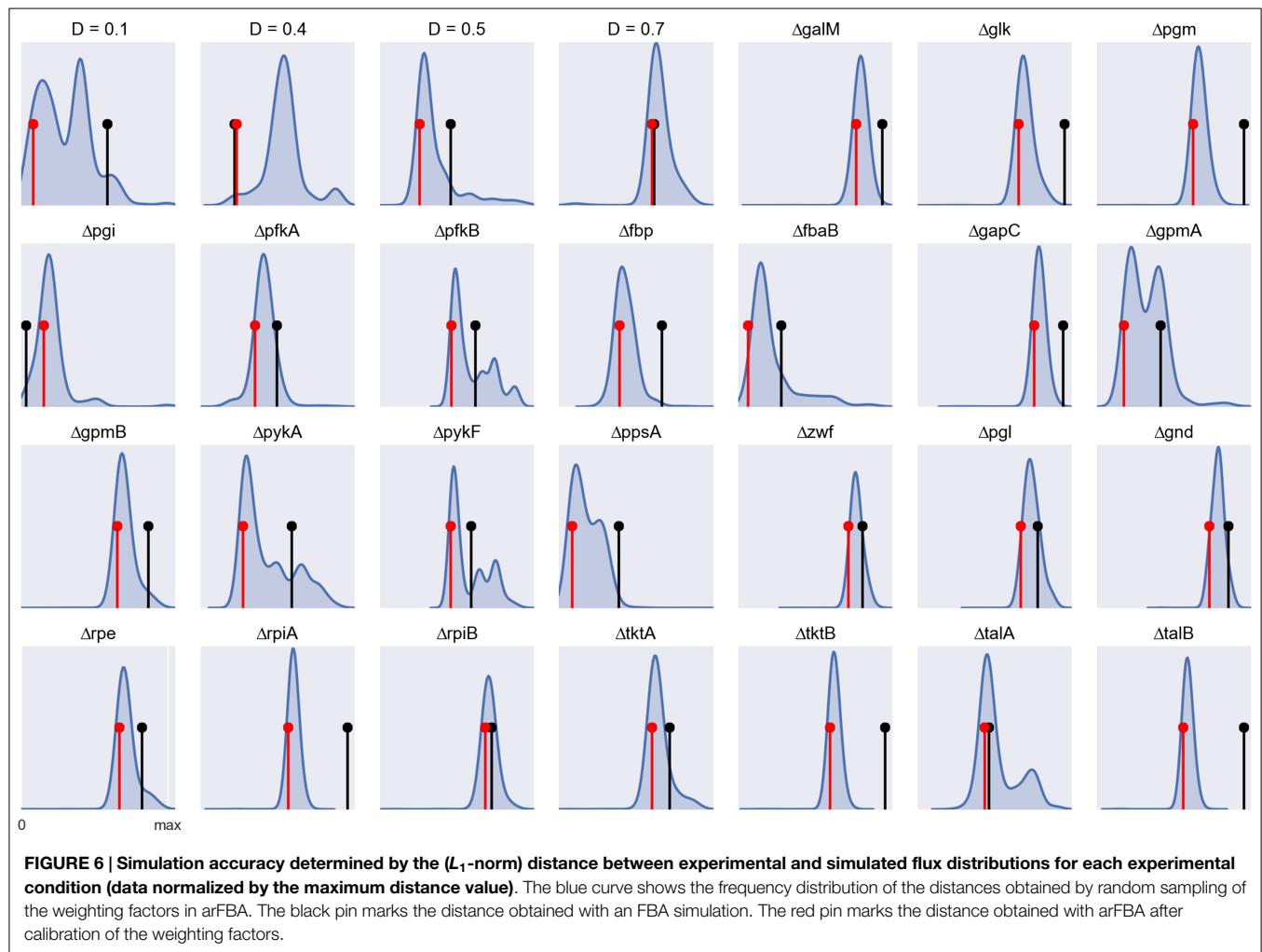
3. Discussion

In this work, we analyzed the role of allosteric regulation for flux control in the central carbon metabolism of *E. coli*. For this, we extended a constraint-based metabolic model of *E. coli* with allosteric regulation. The application of such a model is twofold. First, it can be used as an integrative scaffold for multi-omics dataset analysis, revealing the coordination between enzyme, metabolite, and flux levels. Second, it can be used for *in silico*-based predictions that account for allosteric regulation in the simulation of the metabolic phenotype. For that purpose, we implemented an FBA variant, named arFBA, that accounts for allosteric interactions in the determination of the flux distribution.

Using the expanded model and a multi-omics dataset for *E. coli* (Ishii et al., 2007), we analyzed the impact of allosteric regulation in controlling the metabolic flux under multiple environmental and genetic perturbations. We implemented a generalized form of regulation analysis (ter Kuile and Westerhoff, 2001) in order to find which reactions are predominantly under transcriptional or allosteric control. The results reveal that most reactions are

generally not controlled by the same mechanism across all conditions. This led us to analyze the effects of perturbations in single reactions for each experimental condition. This analysis is hampered by missing protein and metabolite measurements, which does not allow accounting for all participating compounds in the reactions analyzed. Although we neglected the effect of missing isozyme and cofactor measurements, as well as product concentrations for irreversible reactions, only 38 out of 672 possible case studies (24 reactions \times 28 perturbations) could be analyzed in a meaningful way (Figure S6 in Supplementary Material). Nonetheless, it was possible to identify 8 (out of 38) cases where the reaction flux is predominantly controlled by allosteric mechanisms.

Considering that the dataset published by Ishii et al. (2007) is one of the most comprehensive multi-omics dataset for a model organism published so far, we can conclude that purely data-driven analysis is very limited for studying metabolic regulation. Therefore, we applied our simulation method using an ensemble modeling approach to identify which allosteric interactions result in improved flux predictions. Enrichment analysis of the weighting factors in our model revealed that several allosteric interactions were significantly enriched when the models were filtered by their agreement with experimental flux data. A comparison between the *in silico* results and the data-driven analysis showed that 4 of the 8 cases of allosteric control previously identified were also detected by the computational approach.



Given the very limited scope of the cases analyzed in detail, the cross-comparison between the data driven and *in silico* results can hardly be considered a validation of the latter. In order to determine the accuracy of the simulation method it would be necessary to estimate the number of false positive and false negative results for the whole dataset. Instead, the two approaches should be seen as complementary methods to guide the analysis of allosteric regulation. Furthermore, the data analysis revealed that the predominant mode of regulation for each reaction is condition dependent. This was also observed in the *in silico* analysis, hampering the determination of a universal set of weighting factors for arFBA. Given the interplay between different regulation mechanisms, the approach developed herein could be suitable for integration with other methods for identification of regulation mechanisms (Bordel et al., 2010).

An ensemble modeling approach was also employed by Link et al. (2013) for systematic identification of allosteric interactions in *E. coli*. The authors measured metabolite concentrations using rapid sampling and ^{13}C -labeled substrates (glucose and fructose) to determine the transient profile of glycolytic intermediates in dynamic cultures switching between glycolysis and gluconeogenesis. A kinetic ensemble model for glycolysis was used to test 126

putative interactions. The results not only confirmed previously known interactions but also predicted new interactions that had not been previously reported. Although the model used in this study differs from ours, the results regarding interactions common to both models are consistent. In particular, both studies revealed the importance of PFK as an active regulation target for controlling the glycolytic flux, and the role of *fdp* as key regulator of PPC and PYK to control *pep* consumption.

At the end of our data-driven analysis, some flux changes remain unexplained by hierarchical or metabolic control. One main reason for this is the lack of coverage of the metabolomics data, which only accounts for approximately half of the metabolites in the model. Another possibility is that the regulatory mechanisms for the respective enzymes are not fully known or the relevant allosteric interactions were not included in the model. It is also possible that the enzyme concentrations do not correlate with the respective enzymatic activity due to post-translational modifications (PTMs). It has been shown that PTMs, such as acetylation, have important regulatory functions in *E. coli* (Castaño-Cerezo et al., 2014).

The generation of high-quality multi-omics datasets will be necessary for a deeper understanding of metabolic regulation.

Herein, we used a previously published dataset for chemostat cultures. However, steady-state data may be insufficient to analyze regulatory responses. It has been observed that fast metabolic responses precede the slower transcriptional response during metabolic adaptation (Ralser et al., 2009). Since allosteric regulation operates on a faster time-scale compared to transcriptional regulation, transient profiles on short time scales should be particularly informative (Link et al., 2013).

4. Conclusion

In this work, we focused on the role of allosteric regulation in central carbon metabolism. The reconstruction of an allosteric model revealed that allosteric information is inconsistent among different data sources even for these highly studied pathways. The allosteric interactions added a new layer to the network topology, changing the overall network connectivity and revealing metabolic hubs that would otherwise be ignored (e.g., *fdp*). Hierarchical and allosteric regulation analysis using a multi-omics dataset revealed that there is no predominant mechanism of regulation across all experimental conditions. Nonetheless, situations of predominant allosteric control could be identified for some reactions at particular conditions. Our new method for model-based prediction of allosteric control was able to capture at least a few of these situations. However, the assessment of the predictive ability of this method is hampered by the lack of more comprehensive data.

For central carbon metabolism, it would have been feasible to perform this analysis using a kinetic modeling approach [similarly to Link et al. (2013)]. However, as we move toward regulatory analysis at the genome-scale, the constraint-based approach should become especially useful. Building a genome-scale model of allosteric regulation is a daunting task that will require literature mining, extensive manual curation, and prediction of putative interactions. Our knowledge of the *allosterome* is currently limited by the lack of high-throughput screening methods for detecting metabolite–enzyme interactions. It is likely that the vast majority of allosteric interactions are yet to be discovered (Lindsley and Rutter, 2006). Recent experimental methods have been developed toward systematic identification of metabolite–protein interactions (Gallego et al., 2010; Li et al., 2010; Orsak et al., 2011; Feng et al., 2014). However, we are still far from a genome-scale screening of the hundreds of thousands of potential interactions between all metabolites and enzymes in an organism.

Notebaart et al. (2014) have recently unraveled the *underground* metabolism of *E. coli* by expanding a genome-scale metabolic model with reactions resulting from promiscuous enzyme activity. With the *allosterome*, we can unravel yet another hidden layer in the network topology of cellular metabolism. New expanded models of metabolism will be certainly useful for applications, such as drug discovery and rational strain design, as we slowly move toward what has been called the “second secret of life” (Fenton, 2008).

A python implementation of arFBA as well as the allosteric model in SBML format are available on GitHub: <https://github.com/cdanielmachado/arfb>.

5. Materials and Methods

5.1. Model Reconstruction

The original model of the core metabolism of *E. coli* (Orth et al., 2009) was extended with allosteric interactions obtained from BRENDA (Schomburg et al., 2002), EcoCyc (Keseler et al., 2011), and two previously published kinetic models (Chassagnole et al., 2002; Kotte et al., 2010). We searched for evidence of regulatory interactions for each possible combination of enzymes and metabolites in the model. A total of 148 regulatory interactions were found (Figure S3 in Supplementary Material). Since the majority of these interactions can only be found in one data source, for the sake of curation we only included in the model the interactions that are reported in at least two different sources. In a few cases the same metabolite is reported as activator and inhibitor of an enzyme (e.g., *phosphoenolpyruvate* binding to *fructose-bisphosphatase*). In these cases, we used the most frequently reported effect.

5.2. Regulation Analysis

5.2.1. Cross-Condition Analysis

The metabolic flux of a reaction (J_i) can be generically described in terms of the concentrations of the respective enzyme(s) (E_i) and all the intervening metabolites (substrates, products, effectors):

$$J_i = k_{\text{cat}} E_i f(M)$$

where k_{cat} is the turnover rate of the enzyme, and $f(M)$ represents a non-linear function of the metabolite concentrations. *Regulation analysis* introduced by ter Kuile and Westerhoff (2001) decomposes the contribution from hierarchical and metabolic control by considering the logarithmic change between two experimental conditions:

$$\Delta \log(J_i) = \Delta \log(E_i) + \Delta \log(f(M))$$

and estimating the respective contribution coefficients:

$$1 = \frac{\Delta \log(E_i)}{\Delta \log(J_i)} + \frac{\Delta \log(f(M))}{\Delta \log(J_i)} = \rho_h + \rho_m.$$

Since $f(M)$ is generally unknown, one can estimate ρ_h (and consequently ρ_m) by measuring the enzyme and flux levels across different conditions. Chubukov et al. (2013) generalized this comparison from two to multiple conditions in order to decrease the effects of experimental error. The estimation is performed by linear regression between $\log(E_i)$ and $\log(J_i)$ across all experimental conditions using a robust linear regression method (Theil–Sen estimator).

We further generalized this concept to the study of allosteric regulation, by decoupling the effect of allosteric regulators in the reaction flux from the non-linear $f(M)$ component, using a power-law approximation:

$$f(M) \approx g(S, P) \prod_j A_j^{\gamma_{ij}} \prod_j I_j^{-\gamma_{ij}}$$

where S, P, A, I represent, respectively, the set of substrates, products, activators and inhibitors of reaction i , and γ_{ij} is the apparent

kinetic order of effector j in reaction i , as defined in Biochemical-Systems Theory (Voit, 2013). This allows us to estimate individual allosteric regulation coefficients (ρ_a) for each effector as:

$$\rho_{a(j)} = \begin{cases} \gamma_{ij} \frac{\Delta \log(A_j)}{\Delta \log(I_i)} & \text{if } j \text{ is an activator of } i \\ -\gamma_{ij} \frac{\Delta \log(I_j)}{\Delta \log(I_i)} & \text{if } j \text{ is an inhibitor of } i \end{cases}$$

With the exception of effectors exhibiting cooperative binding, we can assume that the kinetic orders are close to or below unity ($\gamma_{ij} \leq 1$). Hence, the allosteric control coefficient is bound by the slope of the linear regression.

Regulation analysis was performed for all allosterically regulated reactions with available fluxomics and proteomics data. A total of 18 (out of 24) regulated reactions were experimentally measured. Due to gaps in the proteomics dataset, we restricted the analysis to enzymes with available data for at least 10 (out of 29) experimental conditions.

5.2.2. Single-Condition Analysis

Allosteric effects were analyzed for each perturbation individually by comparing the logarithmic change of enzyme, flux, and metabolite levels between all 28 perturbed conditions and the reference condition. Due to the sparsity of the data (especially the metabolome data), this analysis was restricted to all reaction-condition combinations where the following criteria were satisfied: (1) at least one associated enzyme was measured; (2) all main substrates (excluding cofactors) were measured; (3) at least one effector was measured. Furthermore, we excluded flux changes that were not significant (i.e., the perturbed flux falls within a 95% confidence interval of the reference flux).

Evidence of allosteric control was detected by selecting conditions where the flux change is not fully explained by changes in enzyme concentration ($\Delta \log(E)/\Delta \log(J) < 0.5$) or substrate abundance ($\Delta \log(S)/\Delta \log(J) < 0.5$), and is at least partly related with changes in one allosteric activator ($\Delta \log(A)/\Delta \log(J) > 0.25$) or inhibitor ($-\Delta \log(I)/\Delta \log(J) > 0.25$). For reversible reactions,

the effect of flux changes arising from changes in the thermodynamic driving force cannot be excluded. Therefore, for these reactions we only considered reactions where the products were experimentally measured (excluding cofactors) and the flux change cannot be fully explained by the change in product abundance ($-\Delta \log(P)/\Delta \log(J) < 0.5$).

5.2.3. Ensemble Modeling with arFBA

For each experimental condition, an ensemble of 10^4 models was built by sampling the weighting factors (w_{ij} parameters) from a log-normal distribution. Each model is constrained with the experimentally measured glucose and oxygen uptake rates, and the growth rate, which is given by the dilution rate.

5.2.4. Calibration of Weighting Factors in arFBA

Condition-specific weighting factors were calibrated for each experimental condition as follows: an ensemble of 10^4 arFBA models was built as described above; the accuracy of each model was determined by the L_1 -norm distance between the experimental and simulated flux distributions; the calibrated weighting factors were calculated as the average of the 10% most accurate models.

Acknowledgments

DM and IR would like to thank the FCT Strategic Project PEst-OE/EQB/LA0023/2013 and the Project “BioInd – Biotechnology and Bioengineering for improved Industrial and Agro-Food processes”, REF. NORTE-07-0124-FEDER-000028, co-funded by the Programa Operacional Regional do Norte (ON.2 – O Novo Norte), QREN, FEDER. MJH would like to thank the Novo Nordisk Foundation for support.

Supplementary Material

The Supplementary Material for this article can be found online at <http://journal.frontiersin.org/article/10.3389/fbioe.2015.00154>

References

- 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. doi:10.1038/nchembio.186
- Bordbar, A., Monk, J. M., King, Z. A., and Palsson, B. O. (2014). Constraint-based models predict metabolic and associated cellular functions. *Nat. Rev. Genet.* 15, 107–120. doi:10.1038/nrg3643
- Bordel, S., Agren, R., and Nielsen, J. (2010). Sampling the solution space in genome-scale metabolic networks reveals transcriptional regulation in key enzymes. *PLoS Comput. Biol.* 6:e1000859. doi:10.1371/journal.pcbi.1000859
- Castaño-Cerezo, S., Bernal, V., Post, H., Fuhrer, T., Cappadona, S., Sánchez-Díaz, N. C., et al. (2014). Protein acetylation affects acetate metabolism, motility and acid stress response in *Escherichia coli*. *Mol. Syst. Biol.* 10, 762. doi:10.15252/msb.20145227
- Chassagnole, C., Noisommit-Rizzi, N., Schmid, J. W., Mauch, K., and Reuss, M. (2002). Dynamic modeling of the central carbon metabolism of *Escherichia coli*. *Biotechnol. Bioeng.* 79, 53–73. doi:10.1002/bit.10288
- Chubukov, V., Uhr, M., Le Chat, L., Kleijn, R. J., Jules, M., Link, H., et al. (2013). Transcriptional regulation is insufficient to explain substrate-induced flux changes in *Bacillus subtilis*. *Mol. Syst. Biol.* 9, 709. doi:10.1038/msb.2013.66
- Daran-Lapujade, P., Rossell, S., van Gulik, W. M., Luttkik, M. A., de Groot, M. J., Slijper, M., et al. (2007). The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15753–15758. doi:10.1073/pnas.0707476104
- Fendt, S.-M., Buescher, J. M., Rudroff, F., Picotti, P., Zamboni, N., and Sauer, U. (2010). Tradeoff between enzyme and metabolite efficiency maintains metabolic homeostasis upon perturbations in enzyme capacity. *Mol. Syst. Biol.* 6, 356. doi:10.1038/msb.2010.11
- Feng, Y., De Franceschi, G., Kahraman, A., Soste, M., Melnik, A., Boersema, P. J., et al. (2014). Global analysis of protein structural changes in complex proteomes. *Nat. Biotechnol.* 32, 1036–1044. doi:10.1038/nbt.2999
- Fenton, A. W. (2008). Allosteric: an illustrated definition for the ‘second secret of life’. *Trends Biochem. Sci.* 33, 420–425. doi:10.1016/j.tibs.2008.05.009
- Folger, O., Jerby, L., Frezza, C., Gottlieb, E., Rupp, E., and Shlomi, T. (2011). Predicting selective drug targets in cancer through metabolic networks. *Mol. Syst. Biol.* 7, 501. doi:10.1038/msb.2011.35
- Gallego, O., Betts, M. J., Gvozdenovic-Jeremic, J., Maeda, K., Matetzki, C., Aguilar-Gurrieri, C., et al. (2010). A systematic screen for protein–lipid interactions in *Saccharomyces cerevisiae*. *Mol. Syst. Biol.* 6, 430. doi:10.1038/msb.2010.87
- Gonçalves, E., Bucher, J., Ryll, A., Niklas, J., Mauch, K., Klamt, S., et al. (2013). Bridging the layers: towards integration of signal transduction, regulation and

- metabolism into mathematical models. *Mol. Biosyst.* 9, 1576–1583. doi:10.1039/c3mb25489e
- Heinemann, M., and Sauer, U. (2010). Systems biology of microbial metabolism. *Curr. Opin. Microbiol.* 13, 337–343. doi:10.1016/j.mib.2010.02.005
- Ishii, N., Nakahigashi, K., Baba, T., Robert, M., Soga, T., Kanai, A., et al. (2007). Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. *Science* 316, 593–597. doi:10.1126/science.1132067
- Keseler, I. M., Collado-Vides, J., Santos-Zavaleta, A., Peralta-Gil, M., Gama-Castro, S., Muñiz-Rascado, L., et al. (2011). EcoCyc: a comprehensive database of *Escherichia coli* biology. *Nucleic Acids Res.* 39(Suppl. 1), D583–D590. doi:10.1093/nar/gkq1143
- Kochanowski, K., Sauer, U., and Chubukov, V. (2013a). Somewhat in control – the role of transcription in regulating microbial metabolic fluxes. *Curr. Opin. Biotechnol.* 24, 987–993. doi:10.1016/j.copbio.2013.03.014
- Kochanowski, K., Volkmer, B., Gerosa, L., van Rijsewijk, B. R. H., Schmidt, A., and Heinemann, M. (2013b). Functioning of a metabolic flux sensor in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1130–1135. doi:10.1073/pnas.1202582110
- Kotte, O., Zaugg, J. B., and Heinemann, M. (2010). Bacterial adaptation through distributed sensing of metabolic fluxes. *Mol. Syst. Biol.* 6, 355. doi:10.1038/msb.2010.10
- Lewis, N. E., Hixson, K. K., Conrad, T. M., Lerman, J. A., Charusanti, P., Polpitiya, A. D., et al. (2010). Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Mol. Syst. Biol.* 6, 1–13. doi:10.1038/msb.2010.47
- Li, X., Gianoulis, T. A., Yip, K. Y., Gerstein, M., and Snyder, M. (2010). Extensive in vivo metabolite-protein interactions revealed by large-scale systematic analyses. *Cell* 143, 639–650. doi:10.1016/j.cell.2010.09.048
- Lindsley, J. E., and Rutter, J. (2006). Whence cometh the allosterome? *Proc. Natl. Acad. Sci. U.S.A.* 103, 10533–10535. doi:10.1073/pnas.0604452103
- 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. doi:10.1038/nbt.2489
- Machado, D., Costa, R. S., Ferreira, E. C., Rocha, I., and Tidor, B. (2012). Exploring the gap between dynamic and constraint-based models of metabolism. *Metab. Eng.* 14, 112–119. doi:10.1016/j.ymben.2012.01.003
- Machado, D., and Herrgård, M. (2014). Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS Comput. Biol.* 10:e1003580. doi:10.1371/journal.pcbi.1003580
- Matsuoka, Y., and Shimizu, K. (2015). Current status and future perspectives of kinetic modeling for the cell metabolism with incorporation of the metabolic regulation mechanism. *Bioresour. Bioprocess.* 2, 1–19. doi:10.1186/s40643-014-0031-7
- Monk, J., Nogales, J., and Palsson, B. O. (2014). Optimizing genome-scale network reconstructions. *Nat. Biotechnol.* 32, 447–452. doi:10.1038/nbt.2870
- Notebaart, R. A., Szappanos, B., Kintsjes, B., Pál, F., Györkei, Á., Bogos, B., et al. (2014). Network-level architecture and the evolutionary potential of underground metabolism. *Proc. Natl. Acad. Sci. U.S.A.* 111, 11762–11767. doi:10.1073/pnas.1406102111
- Orsak, T., Smith, T. L., Eckert, D., Lindsley, J. E., Borges, C. R., and Rutter, J. (2011). Revealing the allosterome: systematic identification of metabolite-protein interactions. *Biochemistry* 51, 225–232. doi:10.1021/bi201313s
- Orth, J., Fleming, R., and Palsson, B. (2009). “Reconstruction and use of microbial metabolic networks: the core *Escherichia coli* metabolic model as an educational guide,” in *EcoSal – Escherichia coli and Salmonella: Cellular and Molecular Biology*, eds A. Bock, I. R. Curtiss, J. Kaper, P. Karp, F. Neidhardt, T. Nyström, et al. (Washington, DC: ASM Press), 56–99.
- Orth, J. D., Thiele, I., and Palsson, B. Ø (2010). What is flux balance analysis? *Nat. Biotechnol.* 28, 245–248. doi:10.1038/nbt.1614
- Ralsler, M., Wamelink, M. M., Latkolik, S., Jansen, E. E., Lehrach, H., and Jakobs, C. (2009). Metabolic reconfiguration precedes transcriptional regulation in the antioxidant response. *Nat. Biotechnol.* 27, 604–605. doi:10.1038/nbt0709-604
- Rossell, S., van der Weijden, C. C., Lindenberg, A., van Tuijl, A., Francke, C., Bakker, B. M., et al. (2006). Unraveling the complexity of flux regulation: a new method demonstrated for nutrient starvation in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2166–2171. doi:10.1073/pnas.0509831103
- Schellenberger, J., Park, J. O., Conrad, T. M., and Palsson, B. Ø (2010). BiGG: a biochemical genetic and genomic knowledgebase of large scale metabolic reconstructions. *BMC Bioinformatics* 11:213. doi:10.1186/1471-2105-11-213
- Schomburg, I., Chang, A., and Schomburg, D. (2002). BRENDA, enzyme data and metabolic information. *Nucleic Acids Res.* 30, 47–49. doi:10.1093/nar/30.1.47
- ter Kuile, B. H., and Westerhoff, H. V. (2001). Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. *FEBS Lett.* 500, 169–171. doi:10.1016/S0014-5793(01)02613-8
- Teusink, B., Passarge, J., Reijenga, C. A., Esgalhado, E., van der Weijden, C. C., Schepper, M., et al. (2000). Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. *Eur. J. Biochem.* 267, 5313–5329. doi:10.1046/j.1432-1327.2000.01527.x
- Tran, L. M., Rizk, M. L., and Liao, J. C. (2008). Ensemble modeling of metabolic networks. *Biophys. J.* 95, 5606–5617. doi:10.1529/biophysj.108.135442
- Voit, E. O. (2013). Biochemical systems theory: a review. *ISRN Biomath.* 2013, 897658. doi:10.1155/2013/897658
- Wessely, F., Bartl, M., Guthke, R., Li, P., Schuster, S., and Kaleta, C. (2011). Optimal regulatory strategies for metabolic pathways in *Escherichia coli* depending on protein costs. *Mol. Syst. Biol.* 7, 515. doi:10.1038/msb.2011.46
- Zomorodi, A. R., Suthers, P. F., Ranganathan, S., and Maranas, C. D. (2012). Mathematical optimization applications in metabolic networks. *Metab. Eng.* 14, 672–686. doi:10.1016/j.ymben.2012.09.005

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Machado, Herrgård and Rocha. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.