Out of (transcriptional) control? Design principles of the regulatory network controlling metabolic pathways in Escherichia coli

Frank Wessely[1,2,†], Martin Bartl[3], Reinhard Guthke[4], Pu Li[3], and Christoph Kaleta[1,2*]

[1] *Research Group Theoretical Systems Biology, Friedrich Schiller University Jena, Jena, Germany,* [2] *Department of Bioinformatics, Friedrich Schiller University Jena, Jena, Germany,* [3] *Institute for Automation and Systems Engineering, Ilmenau University of Technology, Ilmenau, Germany,* [4] *Systems Biology/Bioinformatics Group, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute, Jena, Germany,* [†] *present address: School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington Campus, Loughborough, UK*

*christoph.kaleta@uni-jena.de

Abstract: While a large number of previous studies has explored the link between the structure of metabolism and its regulation, the extent to which transcriptional regulation controls metabolism has not yet been fully elucidated. We address this problem by integrating a large number of experimental data sets with a genome-scale metabolic model of *Escherichia coli* metabolism. We find that there is a strong connection between the extent of transcriptional regulation in a metabolic pathway and the protein investment into this pathway. While pathways associated to a low protein cost tend to be controlled only in key steps, pathways associated to a high protein cost are controlled by fine-tuned transcriptional regulatory programs. These different strategies for the control of metabolic pathways can be explained by a trade-off between the conflicting requirements to minimize protein investment and to maintain the ability to quickly respond to changes in environmental conditions.

1 Introduction

The increasing availability and decreasing prices of experimental techniques have led to an explosion in the number of available experimental data sets [LVW+07, BKG+09]. These data sets provide an increasingly comprehensive view on the principles that influence the evolution of the

regulatory network controlling metabolism [NTSP08]. In this submission we discuss the results of a previous work [WBG$^+$11], in which we have used different types of OMICs data sets in order to identify these global principles of regulatory network evolution in the model organism *Escherichia coli*.

2 Results

In order to understand to which extent transcriptional regulation controls metabolism, we investigated the coexpression of enzymes within the pathways of all biochemically annotated subsystems of *E. coli* metabolism. This analysis was based on the concept of elementary flux patterns [KdFS09], which allowed us to identify pathways in all subsystems of metabolism. By mapping gene expression data to the corresponding pathways, we found that pathways in many subsystems of metabolism show a large degree of coexpression. However, pathways in the subsystems cofactor and prosthetic group biosynthesis, glycerophospholipid metabolism, murein recycling, nucleotide salvage pathway and pentose phosphate pathway show only weak coexpression of pathways. We call these subsystems with a low coexpression of pathways "transcriptionally sparsely regulated subsystems".

To provide an explanation for these distinct patterns of transcriptional regulation, we constructed a simplified model of a linear metabolic pathway that converts a substrate s via four intermediates into a product p. Dynamic optimization was used to identify specific regulatory programs (representing time-courses of enzyme concentrations) that allow the cell to precisely adjust the concentration of the product in a changing environment while obeying a set of physiological constraints. As objective function we used the minimization of the change of enzyme concentrations from initial concentrations and protein costs.

The results of this optimization procedure showed that for a full control of flux through a pathway, transcriptional regulation of initial and terminal positions of a pathway is sufficient (sparse transcriptional regulation). The role of the control of the first enzyme of a pathway is to regulate the flux into the pathway and avoid the accumulation of intermediates. In contrast, the control of the terminal reaction of a pathway allows the cell to precisely adjust the rate of synthesis of the product. Performing the same optimization for a large number of pathways with randomized kinetic parameters, we found that these principles hold true regardless

of kinetic parameters. Moreover, we found that with increasing cost of enzymes of a pathway (i.e. increasing enzyme concentrations) there is a shift from the sparse transcriptional regulation of a metabolic pathway to the coordinated transcriptional control of all enzymes in a pathway (pervasive transcriptional regulation).

We validated these predictions by an analysis of the position-specific frequency of regulatory events in the pathways of transcriptionally sparsely regulated subsystems. We confirmed that there is a significant increase in the frequency of transcriptional regulation at the beginning and end of pathways. Moreover, we found a significant increase of the frequency of post-translational regulation at the beginning of pathways. Thus, the control at the initial positions of pathways is achieved through a combination of transcriptional as well as post-translational regulation, while control at the end of pathways is achieved through transcriptional regulation. In other subsystems that were not identified as being transcriptionally sparsely regulated by the expression analysis, we did not find this pattern of transcriptional regulation, while the pattern of post-translational regulation prevailed. Investigating data of protein costs (defined as the total mass of a particular protein in the cell) for different subsystems we found that in particular subsystems with a small cost of proteins show a pattern of transcriptional sparse regulation.

3 Discussion

Since we were able to confirm the predictions of the optimization, there appears to be an evolutionary mechanism favoring sparse transcriptional regulation in pathways with low-cost enzymes. We propose an evolutionary trade-off between the two conflicting objectives of the minimization of protein investment and the minimization of response time. The optimal strategy to reduce protein investment is to transcriptionally control proteins and express them only if they are needed. However, response times on a transcriptional level are usually very slow. Optimal response times can be achieved through a constitutive expression of most enzymes in a pathway and a transcriptional control of key steps. The interplay between both objectives results in a pervasive transcriptional control of all enzymes within a pathway if they are associated to a high cost. In pathways with low-cost enzymes, transcriptionally sparse regulation prevails. In support of these results, we found that even costly pathways such as the pentose phosphate pathway, for which rapid response times are required, are

sparsely regulated due to a strong advantage of a rapid response time. Finally, if there is only a small fitness advantage of both cellular objectives, sparse transcriptional regulation is a minimum requirement to precisely control the flux through a pathway.

These results demonstrate that, in contrast to the classical picture of regulation, the control of key positions of metabolic pathways is sufficient to achieve a full control over the flux through a pathway. Such a pattern of sparse transcriptional regulation is useful if a higher fitness advantage can be achieved through rapid response times in comparison to the fitness advantage of a reduced protein cost.

References

[BKG+09] Bryson D Bennett, Elizabeth H Kimball, Melissa Gao, Robin Osterhout, Stephen J Van Dien, and Joshua D Rabinowitz. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol*, 5(8):593–599, Aug 2009.

[KdFS09] Christoph Kaleta, Luís Filipe de Figueiredo, and Stefan Schuster. Can the whole be less than the sum of its parts? Pathway Analysis in genome-scale metabolic networks using Elementary Flux Patterns. *Genome Res*, 19(10):1872–1883, Oct 2009.

[LVW+07] Peng Lu, Christine Vogel, Rong Wang, Xin Yao, and Edward M Marcotte. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat Biotechnol*, 25(1):117–124, Jan 2007.

[NTSP08] Richard A Notebaart, Bas Teusink, Roland J Siezen, and Balázs Papp. Co-regulation of metabolic genes is better explained by flux coupling than by network distance. *PLoS Comput Biol*, 4(1):e26, Jan 2008.

[WBG+11] Frank Wessely, Martin Bartl, Reinhard Guthke, Pu Li, Stefan Schuster, and Christoph Kaleta. Optimal regulatory strategies for metabolic pathways in *Escherichia coli* depending on protein costs. *Mol Syst Biol*, 7:515, 2011.