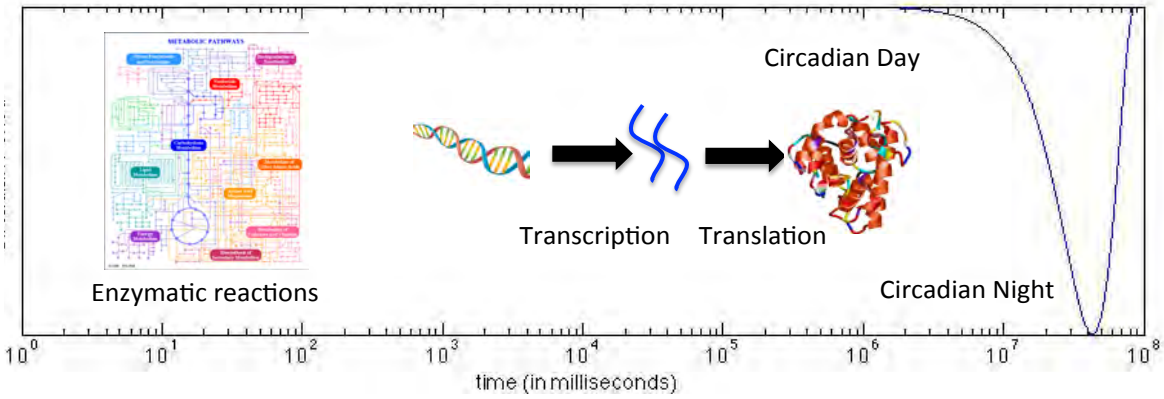
LEARNING REGULATION AND OPTIMAL CONTROL OF ENZYME ACTIVITIES AND APPLICATION TO SYSTEMS BIOLOGY DATA OF NEUROSPORA CRASSA.

William R. Cannon[1], Samuel Britton[2], Mark Alber[2], Jennifer M. Hurley[3], Meaghan S. Jankowski[3], Tina Keliher[4], Jeremy D. Zucker[1], Scott E. Baker[5], Jay C. Dunlap[4]

[1] Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, 99352;
[2] Department of Mathematics, University of California, Riverside, Riverside, CA 92521
[3] Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY, 12180
[4] Department of Molecular and Systems Biology, Geisel School of Medicine at Dartmouth, Hanover, NH 03755
[5] Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, 99352

Project Goals: The goal of this research is to develop and implement a new computational and theoretical method for modeling biological systems that fills a gap in modeling mass action dynamics. Based on statistical thermodynamics, the method bridges data-poor scales (parameters for mass action kinetics) and data-rich scales (chemical potentials of metabolites, and metabolite, protein & transcript data) to enable predictive modeling from enzymatic reactions (10^{-3} to 1 s^{-1}) to gene and protein regulation (~20 minutes) to circadian rhythms (24 hours).

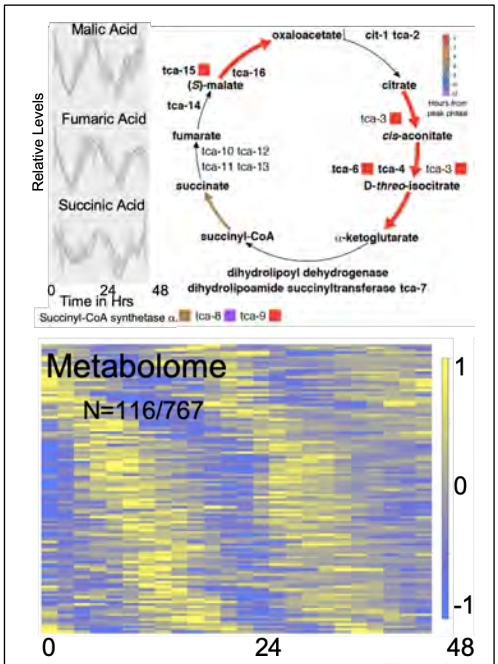


Timescales that the simulations using statistical thermodynamics will cover. Enzymatic reactions occur on the millisecond to second timescale while gene and protein expression occur on the minute to ~30-minute scale and the circadian rhythm occurs over a period of 24 hours.

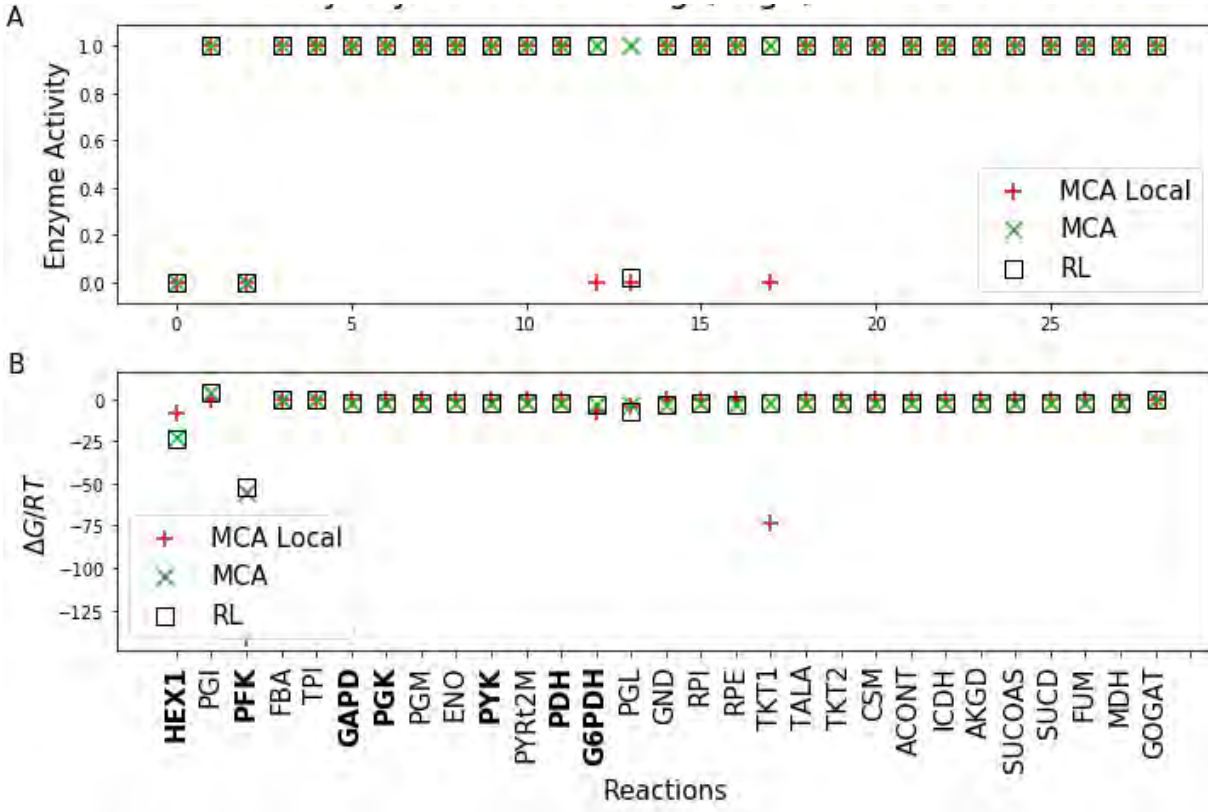
Abstract. Experimental measurement or computational inference/prediction of the enzyme regulation needed in a metabolic pathway is hard problem. Consequently, regulation is known only for well-studied reactions of central metabolism in model organisms. In the last year, we use statistical thermodynamics and metabolic control theory as a theoretical framework to determine the enzyme activities that are needed to control metabolite concentrations such that they are consistent with experimentally measured values. A reinforcement learning approach is utilized to learn optimal regulation policies that match physiological levels of metabolites while maximizing the entropy production rate and minimizing the work needed to maintain a steady state. The learning takes a minimal amount of time, and efficient regulation schemes were learned that agree surprisingly well with known regulation. The learning is facilitated by a new approach in which steady state solutions are obtained by optimization rather than ODE solvers, making the time to solution seconds rather than days. The optimization is based on the Marcelin-De Donder formulation of mass action kinetics [1]. Consequently, a full ODE-based, mass action simulation with rate parameters and post-translational regulation is obtained.

We demonstrate the process on the central metabolism of *Neurospora crassa* which requires different regulation schemes under different nutrient conditions. While many reasons have been proposed as to why enzyme activities are regulated, very few of the proposals lead to specific hypotheses that can be tested. We hypothesize that the post-transcriptional regulation of enzymes is at least in part driven by the need to maintain the solvent capacity in the cell, a directly testable hypothesis.

We investigate this hypothesis computationally by combining experimental metabolomics data with steady state concentrations predicted computationally from equations for reformulated mass action kinetics, which can be solved either by simulation or by optimization. Using quantitative metabolomics data, including both absolute quantitation and relative quantitation over circadian time (right) as well as physical and biological principles, we can predict the control of activity required to bring metabolite levels down to observed values using reinforcement learning (RL) or metabolic control analysis (MCA, local MCA). The predicted activities (below) agree with known regulation of central metabolism in model organisms (highlighted in bold in figure axis). Moreover, the results show that regulated enzymes have higher free energies of reaction precisely because of the regulation, turning common wisdom about enzyme regulation upside-down. Instead of highly non-equilibrium reactions being the targets for regulation in metabolic pathways [2, 3], regulation results in reactions being much further from equilibrium than non-regulated reactions. Being further away from equilibrium than other reactions is an effect, not a cause, of regulation.



Key metabolite levels in the glycolysis pathway in *Neurospora* oscillate in a circadian manner in time with the circadian oscillation of the enzymes that are responsible for producing them.



(Top) Plot of predicted enzyme activity vs enzyme for enzymes of glycolysis, the pentose phosphate pathway and the TCA cycle. Enzymes in bold on x-axis are known to be post-translationally regulated. Activities with values of 1.0 are not regulated while values of 0.0 indicate full regulation. (Bottom) Free energy of reactions for the reactions catalyzed by the respective enzymes.

References

1. Cannon, W.R., et al., *Prediction of metabolite concentrations, rate constants and post-translational regulation using maximum entropy-based simulations with application to central metabolism of Neurospora crassa.* Processes, 2018. **6**(6).
2. Newsholme, E.A. and C. Start, *Regulation in Metabolism*. 1973, London: John Wiley & Sons. 349.
3. Lehninger, A.L., et al., *Lehninger principles of biochemistry*. 5th ed. 2008, New York: W.H. Freeman.

This project is supported by the U.S. Department of Energy's Office of Biological and Environmental Research and the National Institute of Biomedical Imaging and Bioengineering.