PI(s) of MSM U01: Cannon/Dunlap
Institution(s): Pacific Northwest National Laboratory/ Geisel School of Medicine at Dartmouth
MSM U01 Grant Number: 1U01EB022546
Title of Grant: Multiscale Modeling of Circadian Rhythms

Abstract Authors

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

[1] Department of Mathematics, University of California, Riverside, CA, 92521
[2] Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY, 12180
[3] Physical and Computational Sciences Division, Pacific Northwest National Laboratory, Richland, WA, 99352
[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

Abstract Text

Experimental determination of which enzymes need to be regulated in a metabolic pathway is a hard problem. Consequently, regulation is known only for well-studied reactions of central metabolism in various model organisms. In this study, we use control theory and statistical thermodynamics as a theoretical framework to learn optimal enzyme regulation policies from experimental data. The reinforcement learning method learns to match physiological levels of metabolites while maximizing the entropy production rate, or analogously, minimizing the heat loss while maximizing the power generated by the pathway. The learning takes a minimal amount of time when metabolic control analysis is used as a guide, and a efficient greedy method takes only a few seconds to give similar results. We demonstrate the process on four pathways in the central metabolism of *Neurospora crassa* (gluconeogenesis, glycolysis-TCA, Pentose Phosphate-TCA, and cell wall synthesis) that each require different regulation schemes.

The greedy method is applied in four steps: (1) a new convex optimization approach based on Marcelin's 1910 mass action equation is used to obtain the maximum entropy distribution, (2) the predicted metabolite concentrations are compared to experimental observations using a loss function from which post-translational regulation of enzymes is inferred, (3) enzyme regulation is selected using metabolic control analysis, (4) the system is re-optimized until loss function values are acceptable. After convergence, rate constants are determined from the metabolite concentrations and reaction fluxes, and a full ODE-based, mass action simulation with rate parameters and post-translational regulation is obtained.

10 Simple Rules: The **general context** of the model is defined in all publications and in the Jupyter notebooks accompanying the model. Currently, the model only addresses the rate equations of metabolic processes, and only parameters for this **context** are available. In each publication and in the Jupyter notebooks, care is taken to describe how **metabolomics data** can **appropriately** be incorporated into the model, and the limits of the data itself. The model is **currently limited** to a well-mixed assumption regarding distribution of metabolites in the cellular compartments. **Version control** is maintained through Github repositories. The simulation method is **documented** not only in the methods section of each publication, but each modeling publication includes a Jupyter notebook demonstrating the implementation and use of the model. Each model is **disseminated** as a Jupyter notebook that includes not only the model, but the implementation of the model in code. The model and approach were **evaluated with respect to uncertainty quantitation**. We have now **implemented** various versions of general model in Matlab, Python and C. The results are implementation independent. Our current metabolic model is available in **SBML**, but not yet is the reinforcement learning model available in SBML.