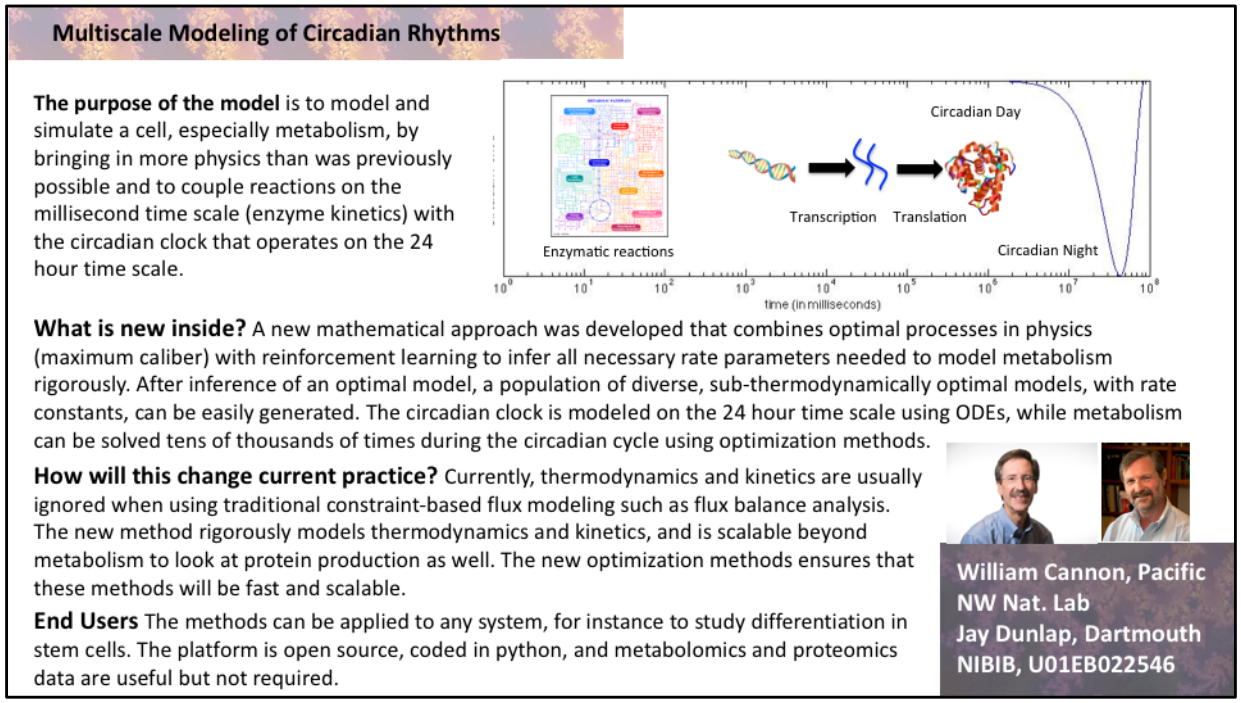

Public statement: A new mathematical approach has been developed that combines our ability to define optimal processes in physics with machine learning to infer all necessary rate parameters needed to rigorously model metabolism. After inference of the optimal model, diverse models representing the metabolisms that would be present in a population can be easily generated. The circadian operation of the cell is modeled on the 24 hour time scale along with metabolism which occurs on the millisecond time scale.

Machine learning, specifically reinforcement learning, is used to learn how regulation should be applied to metabolism. The reinforcement learning algorithm uses metabolomics data, but highly accurate data is not needed. Further advances that combine predictive simulations with reinforcement learning will enable us to go beyond metabolism and model protein and enzyme production rigorously, as well.

Future challenges are (1) to map these activities and dynamics to mesoscale structures of the cell obtained from imaging studies, and (2) to scale these physics-based methods up from metabolism to protein synthesis and then gene expression. Going from metabolism to proteins to genes – a bottom up approach – allows us to propagate the physics from one scale up to another by a series in which we model

the lower scale, average over and parameterize the dynamics, and then model the next higher scale.

Multiscale modeling: The multiscale challenge is being addressed in a manner analogous to operator splitting, in which two time scales can be separated if one time scale reaches a quasi-steady state rapidly compared to the slower time scale. In our approach, ODEs are used to simulate the circadian clock regulation and proteins, while very fast non-linear optimization methods are used to solve the time dependent equations for the kinetics of metabolism.

Medical simulation: The math challenge being overcome is how to model metabolism rigorously when both *in vitro* and *in vivo* rate parameters are extremely difficult to determine experimentally. We overcome this problem by (1) using a new constrained optimization approach to obtain the non-equilibrium, steady state, maximum entropy distribution of metabolism, (2) the predicted metabolite concentrations are compared to those generally expected from experiment using a loss function from which post-translational regulation of enzymes is inferred, (3) the system is re-optimized with the inferred regulation from which rate constants are determined from the metabolite concentrations and reaction fluxes, and finally (4) a full ODE-based, mass action simulation with rate parameters and allosteric regulation is obtained.

Analysis: What makes it possible to infer mechanism using this approach is that the same physical principles are used to formulate the math used in both the simulations and the data analysis methods present in the reinforcement learning implementation. The statistical method is statistical thermodynamics, which is also a multinomial likelihood model used to analyze the data.

Interestingly, the metaphor for self-organizing systems being used in this project – dissipative structures – is also being used in Europe to study and elucidate how the **brain** works.