122. Development of Genome-Scale Dynamic Modeling Framework for Simulating Photoautotrophic Growth of Cyanobacteria

Hyun-Seob Song*(email: hyunseob.song@pnnl.gov), Jim K. Fredrickson, **Allan E. Konopka, and Alexander S. Beliaev**

Pacific Northwest National Laboratory, Biological Sciences Division, Richland, Washington, 99352

Project Goals: The PNNL BSFA conducts fundamental research of microbial photoautotrophs with specific emphasis on photosynthetic energy conversion, reductant partitioning, and central carbon metabolism. As a model system, we utilize unicellular prokaryotic organism *Synechococcus* sp. PCC 7002, which exhibits one of the fastest growth rates known among cyanobacteria and which is also remarkably tolerant to high light intensities. Understanding the genetic and physiological bases of these properties could provide fundamental new insights that are broadly applicable to the optimization of other biological systems for biofuels development. To that end, we are interrogating fluxes through central metabolic pathways to define the major constraining factors (*i.e.* metabolic and regulatory controls) governing carbon partitioning through the metabolic subsystems of cyanobacteria that can be manipulated to increase productivity of specific molecules that are either precursors or fuel molecules themselves.

Systems-level analysis of cellular metabolism has a level of complexity that requires a   simulation model to integrate critical information on key genetic and metabolic mechanisms governing fluxes and predict the most likely outcome of metabolic engineering. Flux balance analysis (FBA) has contributed to a number of important successes in metabolic engineering and has been used for genome-scale metabolic reconstructions. The aim of our work is to develop an   accurate dynamic model as a tool for investigating the fundamental mechanism driving metabolic shift in prokaryotic photoautotrophs which will explicitly incorporate regulation of metabolism (and thus can use global transcriptomic or proteomic datasets to refine the model in a sophisticated fashion). From a practical perspective, industrial-scale processes have rarely operated at steady-state; the capacity to capture culture dynamics in time or after experimentally imposed perturbations will also give these simulation models practical benefits.

We have focused our efforts on the development of a framework that can facilitate the use of dynamic metabolic models for different organisms and applications. This framework features the capability to account for dynamic regulation thus allowing the accurate prediction of growth rate changes and flux distributions in response to environmental variations. Dynamic   modeling approach is particularly useful not only for providing fundamental understanding of metabolic properties but also for exploring the rational strategies of improving the productivity of biofuels and chemicals potentially produced therefrom. However, development of a large- scale dynamic modeling framework faces several critical barriers, including: (*i*) the difficulty in   estimating a large number of parameters, and (*ii*) the lack of efficient algorithms for the   identification of relevant metabolic pathways from genome-scale networks. With regard to the   first issue, our framework avoids the over-parameterization problem by accounting for dynamics   of slowly reacting metabolites (such as extracellular metabolites) only, while neglecting fast metabolites (mostly intracellular). This quasi steady-state approximation results in the model with only a few parameters, which are readily identified from collected experimental data. In contrast to constraint-based approaches that consider only a single optimal pathway, our approach analyzes metabolism in terms of diverse pathways options (termed elementary modes). As addressed in (*ii*), extraction of elementary modes from genome-scale networks poses a challenge due to their combinatorial explosion in a complex, large-scale network. Therefore, we developed a novel algorithm based on a new optimization concept that enables selective sequential computation, i.e., one pathway at a time. The developed

algorithm performs alternate   implementation of integer and linear programming, which led to fast and numerically stable   computation.

As the proof-of-principle, our team has been developing a genome-scale network-based model of *Synechococcus* sp. PCC 7002 (hereafter, *Synechococcus* 7002 with the goal of determining the metabolic characteristics of this organism under carbon, nitrogen and light- limiting conditions, and carrying out dynamic simulation of its metabolic behaviors. Using our   previously developed framework, we incorporated gene and protein expression data into the genome-scale network of *Synechococcus* 7002 (*i*Syp708) available in the literature, and obtained   an initial estimation of flux distributions under each of three growth conditions. This framework is based on the flux minimization principle; flux distribution is estimated by suppressing one flux over another according to the associated gene (or protein) expression level. Then, we applied the   alternate integer and linear programming to *i*Syp708 to selectively extract elementary modes that are close to initially estimated flux distributions. The analysis of elementary modes obtained as such provides condition-specific metabolic characteristics of *Synechococcus* 7002 under   different environmental stresses. Identification of kinetic parameters is in progress for the dynamic simulation of temporal metabolic shift of *Synechococcus* 7002 subject to variations of environmental conditions.