Development of an accelerated procedure for parameterizing kinetic metabolic models for *C. thermocellum*

Satyakam Dash[1]* (sud25@psu.edu), Saratram Gopalakrishnan[1]* (sxg375@psu.edu), Charles Foster[1] (cjf33@psu.edu), and **Costas D. Maranas**[1]

[1]Department of Chemical Engineering, The Pennsylvania State University of State, University Park, PA 16802

http://www.maranasgroup.com/

Project Goals: The goal of the project is to systematically construct dynamic models of *Clostridium thermocellum* by making use of Ensemble Modeling (EM) paradigm through integration of multiple omic information (metabolomic & fluxomic). These models will be instrumental in exploring genetic interventions for overproduction of biofuel products.

In this study, we describe the development of a metabolic model for a cellulolytic microbe, *C. thermocellum* which shows improved predictive capabilities upon incorporation of kinetic information. We further describe the limitations of existing parametrization procedures for kinetic model development and methods to overcome them using *E. coli* as a model system. For *C. thermocellum*, we constructed a core kinetic model k-ctherm118 to capture the regulatory impact of changes in metabolite pools on reaction fluxes using the Ensemble Modeling paradigm. k-ctherm118 was parameterized by using fermentation yield data in lactate, malate, acetate and hydrogen production pathways for 19 measured metabolites, 19 distinct single and multiple gene knockout mutants along with 18 intracellular metabolite concentration data for a *Δgldh* mutant and ten experimentally measured Michaelis-Menten kinetic parameters. k-ctherm118 captures metabolic perturbations caused by (i) nitrogen limitation leading to increased yields for lactate, pyruvate and amino acids, and (ii) ethanol stress triggering an increase in intracellular ammonia and sugar phosphate concentrations due to upregulation of cofactor pools. A secondary activity of ketol-acid reductoisomerase and possible regulation by valine and/or leucine pool levels were revealed by robustness analysis of k-ctherm118. k-ctherm118 also captures the growth inhibitory effect of pentose sugars by upregulating the non-oxidative pp-pathway genes. Overall, the *C. thermocellum* case study demonstrates that the developed kinetic model (k-ctherm118) provides greater insight into metabolic pathways and regulations than the stoichiometric model.

Kinetic models simulate genetic perturbations by modifying specific enzyme levels *a priori* based on the mutant genotype, which in turn modulates the concentration of all intracellular metabolites. We expand the scope of the developed kinetic models by incorporating a transcriptional regulatory layer which refines enzyme levels based on a linear combination of log-normalized changes in growth rate (global) and select intracellular metabolite pool (specific) levels. A major computational bottleneck in the kinetic parameterization process is the lack of a fast and efficient algorithm to identify the optimal set of kinetic parameters that minimizes deviations between predicted and MFA-derived steady-state flux distributions in response to genetic perturbations. The implementation of a gradient-based procedure is limited by failures in steady-state flux evaluations due to slow numerical integration. To this end, we have implemented an algorithm to compute steady-state flux distributions under multiple genetic perturbations for a given set of kinetic parameters that overcome the limitations of numerical integration. In conjunction with an efficient gradient-based scheme for updating kinetic parameters, we have developed a fast and automated algorithm for parametrization of a kinetic model that includes allosteric and transcript

regulations. We demonstrate the strength of this procedure using a recent study which provides a comprehensive metabolic flux characterization of wild-type *E. coli* and 22 knockouts of enzymes in the upper part of central carbon metabolism, including glycolysis, pentose phosphate pathway and ED pathway. We apply these flux datasets and the parametrization procedure to construct a core kinetic model of *E. coli* containing 138 model reactions, 93 metabolites, and 60 substrate-level regulatory interactions.

The work was supported by the genomic science grant from Department of Energy, USA (grant # DE-SC0012377).