

159. Ensemble cell-wide kinetic modeling of anaerobic organisms to support fuels and chemicals production

Satyakam Dash^{1*}(sud25@psu.edu), Ali Khodayari^{1*}(auk241@psu.edu), M. Ahsanul Islam², Yuting Zheng², Paul Lin³, James C. Liao³, Gregory Stephanopoulos², and Costas D. Maranas¹

¹The Pennsylvania State University, University Park; ²Massachusetts Institute of Technology;
³University of California, Los Angeles.

<http://maranas.che.psu.edu/> <http://bamel.scripts.mit.edu/gns/> <http://www.seas.ucla.edu/~liao/>

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

Thermophilic microorganisms have garnered the interest of the bioprocess industry due to their high temperature optimal growth conditions. In particular, two phylogenetically close organisms *Clostridium thermocellum* and *Moorella thermoacetica* have been focused on in the recent years. While *C. thermocellum* can metabolize cellulose into biofuels such as ethanol, *M. thermoacetica* can metabolize syn gas using the unique Wood-Ljungdahl pathway. Despite their increasing role as bio-production platforms, they remain poorly characterized with significant uncertainty in their metabolic repertoire. To this end, we develop cell-wide dynamic models of transcription and metabolism of these two organisms using the concept of Ensemble Modeling (EM) which requires curated genome-scale metabolic (GSM) models of the organisms as its foundation.

The second generation GSM for *C. thermocellum* (iCth446) has been developed, which contains 446 genes and includes 598 metabolites and 637 reactions, along with gene-protein-reaction associations. The GSM is devoid of thermodynamically infeasible cycles and contains elementally and charge-balanced reactions. The GSM was simulated with down-regulation of phosphoenolpyruvate carboxykinase, or down-regulation of malic enzyme and malate dehydrogenase knocked out along with exogenous pyruvate kinase knocked in and lactate dehydrogenase knocked out. The simulations showed higher yield of ethanol production compared to wild-type conditions as observed experimentally [1]. Likewise, the GSM results showed that only lactate hydrogenase knock out did not have any effect on growth rate as observed experimentally [2].

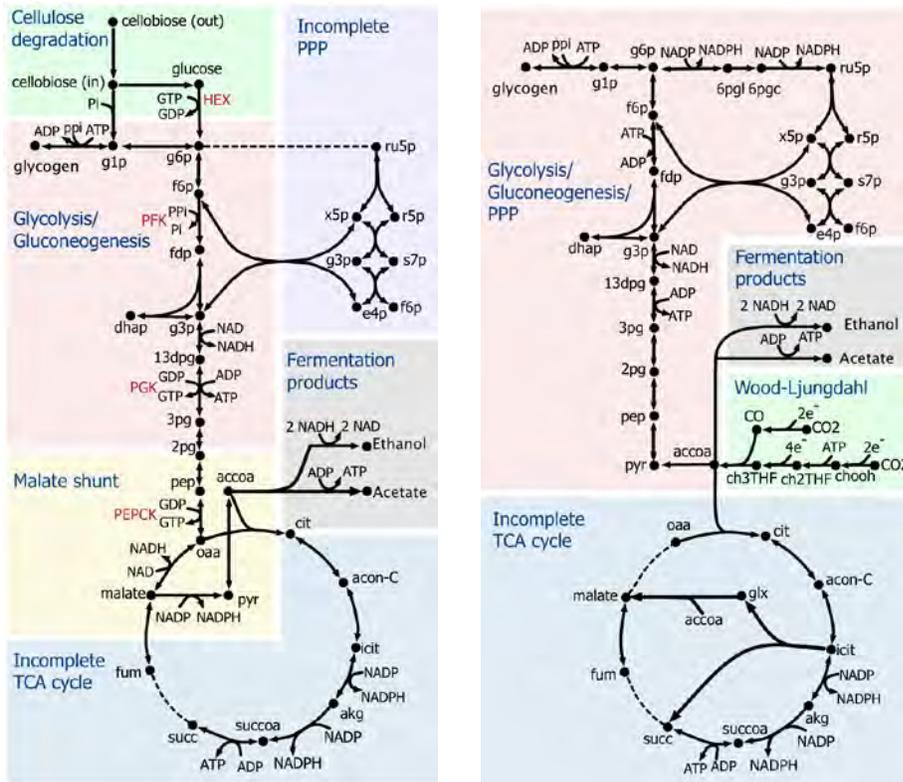
The GSM model for *M. thermoacetica* has been constructed using a semi-automatic pipeline in the SEED database [3]. This model contains 517 genes and 812 reactions, supplemented with 57 exchange reactions including biomass demand reaction. The GSM will be curated, gap-filled, and incorporated with the estimated growth and non-growth associated ATP (GAM & NGAM) requirements. Steady-state ¹³C-labelling experiments will be used to validate flux distribution predicted by the GSM, followed by instationary ¹³C-labeling experiments to identify the robustness of the GSM under various growth conditions with different substrates in distinct growth phases. Ultimately, these experiments will resolve the long-standing debate surrounding the hypothetical incomplete TCA cycle of *M. thermoacetica*.

The constructed stoichiometry representations will subsequently serve as the scaffold for building kinetic model using the EM approach for *C. thermocellum* and *M. thermoacetica*. The EM procedure will allow us to integrate substrate-level as well as transcriptional level regulatory interactions into the framework. The constructed kinetic models will be ultimately used to identify the effect of transcription factors as

well as enzyme level manipulations on metabolic fluxes leading to explore key metabolic drivers that underpin various biofuels production.

The work was supported by the genomic science grant from Department of Energy, USA (grant # DE-FOA-0001060).

1. Deng Y, Olson DG, Zhou JL, Herring CD, Shaw AJ, Lynd LR: Redirecting carbon flux through exogenous pyruvate kinase to achieve high ethanol yields in *Clostridium thermocellum*. *Metab Eng* 2013, 15:151-158.
2. van der Veen D, Lo J, Brown SD, Johnson CM, Tschaplinski TJ, Martin M, Engle NL, van den Berg RA, Argyros AD, Caiazza NC et al: Characterization of *Clostridium thermocellum* strains with disrupted fermentation end-product pathways. *J Ind Microbiol Biotechnol* 2013, 40(7):725-734.
3. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, et al: The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* 2005, 33:5691- 5702.



Central metabolism of *Clostridium thermocellum* (left) and *Moorella thermoacetica* (right)