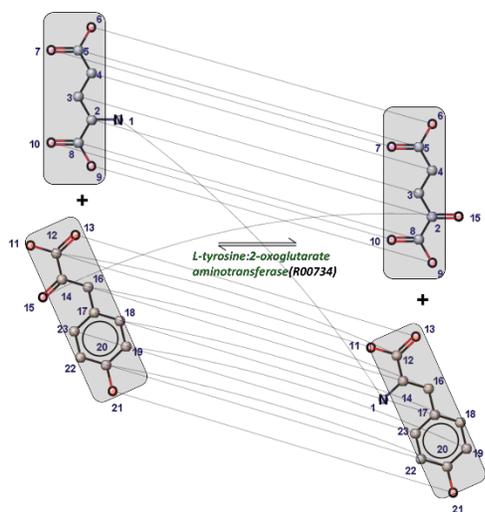


162. MetRxn2.0: Integrating atom mapping information for pathway comparisons and metabolic flux elucidation through MFA

Akhil Kumar^{2*} (azk172@psu.edu), Saratram Gopalakrishnan^{1*} (sxg375@psu.edu) and Costas D. Maranas¹

¹Department of Chemical Engineering, Pennsylvania State University, University Park, PA; ² The Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, PA;

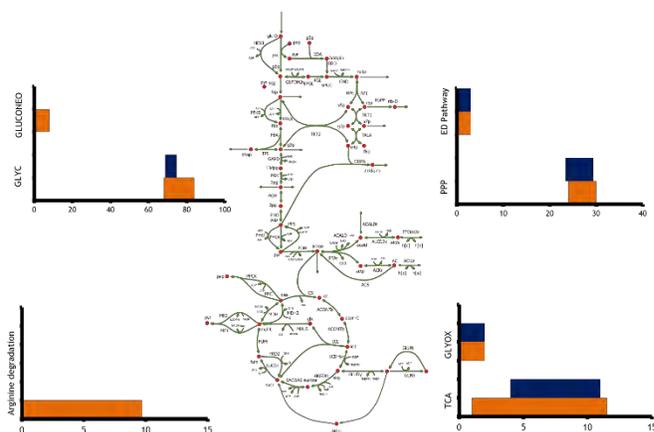
Project Goals: This project aims to organize and disseminate standardized metabolite and reaction information to improve metabolic modeling by accurately describing reaction stoichiometry, directionality, atom mapping from reactants to products, and gene to protein to reaction relations. The generated mapping model will then be used to elucidate metabolic fluxes using ¹³C Metabolic Flux Analysis (MFA) at the genome-scale to provide insights into the role of model scale-up and biomass composition on metabolic flux resolution.



MetRxn (<http://www.metrxn.che.psu.edu/>) is a freely available searchable web-resource that integrates heterogeneous information from 8 metabolic databases and 112 metabolic models. The MetRxn project aims to organize and disseminate standardized metabolite and reaction information to improve metabolic modeling by accurately describing reaction stoichiometry, directionality, atom mapping from reactants to products, and gene to protein to reaction relations. A host of standardization algorithms on the integrated dataset is applied to automatically curate information by removing incompatibilities in content representation, fixing stoichiometric errors such as elemental or charge imbalances and resolving incomplete atomistic details. For each reaction, metabolite stoichiometry, atom transition and metabolite compartment information is stored. The reaction

and metabolite information is downloadable in SBML 3.0 and in a tabular format. The number of distinct reactions that have been mapped is greater than 27,000 and MetRxn contains tools that allow users to download atom mapping data for each reaction. In addition, all charge and mass balanced reactions within the database are processed by our novel algorithm; Canonical Labelling for Clique Approximation (CLCA). CLCA leverages prime factorization to quickly generate unique molecular graphs, detect symmetries for all metabolites and resolve atom/bond transition maps of reactions. The atom transition information is then utilized for the construction of genome-scale mapping models to address the impact of model scale-up on prediction fidelity of metabolic fluxes using ¹³C Metabolic Flux Analysis (MFA).

Metabolic models used in ¹³C metabolic flux analysis generally include a limited number of reactions from the central metabolic network while omitting degradation pathways, complete cofactor balances, or atom transition contributions for reactions outside central metabolism. The base mapping model employed in this study accounts for (75 reactions and 65 metabolites) primarily from central metabolism. The genome-scale mapping model (GSMM) (697 reaction and 595 metabolites) is constructed using as a basis the iAF1260 model upon eliminating reactions guaranteed not to carry flux based on growth and fermentation data for a minimal glucose growth medium. Metabolic fluxes and confidence intervals are estimated, for both base and genome-scale mapping models, by minimizing the sum of square of differences between predicted and experimentally measured labeling patterns using the EMU decomposition algorithm.



Overall, we find that both the topology and estimated values of the metabolic fluxes remain largely consistent between the base and GSM. Stepping up to a genome-scale mapping model leads to wider flux inference ranges for 20 key reactions in the base model. The glycolysis flux range doubled due to the possibility of active gluconeogenesis, the TCA flux range expanded by 80% due to the availability of a bypass through arginine consistent with labeling data, and the transhydrogenase reaction flux was essentially unresolved due to the presence of

as many as 5 routes for the inter-conversion of NADPH to NADH afforded by the genome-scale model. Owing to higher growth-associated ATP demands, the available free ATP decreased drastically with the lower bound corresponding to maintenance ATP requirement. A non-zero flux for the arginine degradation pathway was identified for meeting biomass precursor demands as detailed in the iAF1260 model. Inferred ranges for 81% of the reactions in the GSM model varied less than one-tenth of the basis glucose uptake rate (95% confidence test). This is because as many as 521 reactions in the GSM are growth coupled meaning that the single measurement of biomass formation rate locks the reaction flux values. This implies that accurate biomass formation rate and composition are critical for resolving metabolic fluxes away from central metabolism and suggests the importance of biomass composition (re)assessment under different genetic and environmental backgrounds.

Supported by funding from the U.S. Department of Energy to Dr. Costas D. Maranas grant DE-FG02-05ER25684.