**High-throughput Screening for Carboligase Activity in ThDP-dependent Enzymes**

Tracey Dinh[1]*(tracey.dinh@u.northwestern.edu), Rohin Devanathan[1], Bradley W. Biggs[1], Lindsay Caesar[2], Neil L. Kelleher[2,3,4], Paul M. Thomas[4], Linda J. Broadbelt[1], Keith E.J. Tyo[1]

[1]Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL; [2]Department of Molecular Biosciences, Northwestern University, Evanston, IL; [3]Department of Chemistry, Northwestern University, Evanston, IL; and [4]Proteomics Center of Excellence, Northwestern University, Evanston, IL

https://pamspublic.science.energy.gov/CCBond

**Project Goals:** The goal of this project is to prospect novel carbon-carbon bond ligation reactions in thiamine diphosphate-dependent enzymes using predictive activity models. Carboligase enzymes will be screened for α-keto acid condensation and machine learning models will be developed to elucidate their catalytic landscape. Activity models will be also be used to predict potential effects of selected enzymes on the *E. coli* metabolome.

**Abstract:** Promiscuous enzymatic reactions create 'biological noise' that can be either advantageous or disruptive to global metabolism. Redundancy among promiscuous reactions can act as resistance to deleterious genetic mutations while others can inadvertently siphon essential metabolites or create toxic products. The potential metabolomic burden resulting from promiscuous, heterologous enzyme expression remain challenging to characterize and mitigate. Toward this end, our team seeks to develop a platform to systematically predict enzyme substrate promiscuity and the resulting metabolomic consequences. Merging both experimental and computational approaches, this work aims to comprehensively characterize the catalytic landscape of thiamine diphosphate (ThDP)-dependent enzymes and uncover novel promiscuous transformations for carbon-carbon bond formation. Predictive cheminformatics-based tools will be developed to gain insight into the chemical properties contributing to altered substrate selectivity and the resulting impact of promiscuous activity on cell metabolism.
A high-throughput activity assay was developed for precise functional analysis of ThDP-dependent enzymes capable of catalyzing carbon-carbon ligation (carboligases). Diverse sets of α-keto acid substrates were screened in multiplexed reactions, and the resulting products were detected by liquid chromatography mass spectrometry (LC-MS) to generate high-quality enzyme activity data. The developed carboligase activity assay will be combined with machine learning classification to rapidly characterize the enzyme-specific activity landscapes. Establishing a method for rapid enzyme screening and characterization will also facilitate subsequent rational mutagenesis of select enzymes and assessing the engineered enzymes for metabolic burden. Predictive activity models will be applied to E. coli metabolites to identify potential cross-reactivity and toxic byproducts. This work demonstrates a platform for rapid biocatalyst development based on substrate promiscuity. We present novel reactions discovered for multiple carboligase enzymes; however the analytical methods and cheminformatics tools developed for reaction screening can be widely applied to other chemistries.

This work is supported by DOE grant DE-SC0019339.