

# Computational approaches to understand the atomistic drivers of enzyme catalysis

by  
Natasha Seelam

Submitted to the Department of Chemical Engineering on January 20, 2021, in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy in Chemical Engineering

Enzymes readily perform chemical reactions several orders of magnitude faster than their uncatalyzed versions in ambient conditions with high specificity, making them attractive design targets for industrial purposes. Traditionally, enzyme reactivity has been contextualized through transition-state theory (TST), in which catalytic strategies are described by their ability to minimize the activation energy to cross the reaction barrier through a combination of ground-state destabilization (GSD) and transition-state stabilization (TSS). While excellent progress has been made to rationally design enzymes, the complexity of the design space and the highly optimized nature of enzymes make general application of these approaches difficult. This thesis presents a set of computational methods and applications in order to investigate the larger perspective of enzyme-assisted kinetic processes.

For the first part of the thesis, we analyzed the energetics and dynamics of proficient catalyst orotidine 5'-monophosphate decarboxylase (OMPDC), an enzyme that catalyzes decarboxylation nearly 17 orders of magnitude more proficiently than the uncatalyzed reaction in aqueous solvent. Potential-of-mean-force (PMF) calculations on wild type (WT) and two catalytically hindered mutants, S127A and V155D (representing TSS and GSD, respectively), characterized the energy barriers associated with decarboxylation as a function of two parameters: the distance between the breaking C–C bond and a proton-transfer coordinate from the nearby side chain of K72, a conserved lysine in the active site. Coupling PMF analyses with transition path sampling (TPS) approaches revealed two distinct decarboxylation strategies: a simultaneous, K72-assisted pathway and a stepwise, relatively K72-independent pathway. Both PMF and TPS rate calculations reasonably reproduced the empirical differences in relative rates between WT and mutant systems, suggesting these approaches can enable *in silico* inquiry into both pathway and mechanism identification in enzyme kinetics.

For the second study, we investigated the electronic determinants of reactivity, using the enzyme ketol-acid reductoisomerase (KARI). KARI catalyzes first a methyl isomerization and then reduction with an active site comprised of several polar residues, two magnesium divalent cations, and NADPH. This study focused on isomerization, which is rate limiting, with two objectives: characterization of chemical mechanism in successful catalytic events (“reactive”) versus failed attempts to cross the barrier (“non-reactive”), and the interplay between atomic positions, electronic descriptors, and reactivity. Natural bonding orbital (NBO) analyses provided detailed electronic description of the dynamics through the reaction and revealed that successful

catalytic events crossed the reaction barrier through a 3-center-2-electron (3C) bond, concurrent to isomerization of hydroxyl/carbonyls on the substrate. Interestingly, the non-reactive ensemble adopted a similar electronic pathway as the reactive ensemble, but its members were generally unable to form and sustain the 3C bond. Supervised machine learning classifiers then identified small subsets of geometric and electronic descriptors, “features”, that predicted reactivity; our results indicated that fewer electronic features were able to predict reactivity as effectively as a larger set of geometric features. Of these electronic features, the models selected diverse descriptors representing several facets of the chemical mechanism (charge, breaking–bond order, atomic orbital hybridization states, etc.). We then inquired how geometric features reported on electronic features with classifiers that leveraged pairs of geometric features to predict the relative magnitude of each electronic feature. Our findings indicated that the geometric, pair-feature models predicted electronic structure with comparable performance as cumulative geometric models, suggesting small subsets of features were capable of reporting on electronic descriptors, and that different subsets could be leveraged to describe various aspects of a chemical mechanism.

Lastly, we revisited OMPDC in order to learn the key geometric features that distinguished between the simultaneous and stepwise pathways of decarboxylation, aggregating and labeling pathways drawn from WT and mutant systems ensembles. We leveraged classifiers that predicted between reactive pathways by selecting small subsets of structural features from 620 geometric features comprised of atoms from the active site. The classifiers performed comparably, with greater than 80% testing accuracy and AUC, between times starting from in the reactant basin to 30 fs into crossing the reaction barrier. Remarkably, model-selected features reported on chemically meaningful interactions despite no explicit prior knowledge of the mechanism in training. To illustrate this, we focused analyses on two particular features shown to be predictive while in the reactant basin, prior to crossing the barrier: a potential hydrogen-bond between D75\*, an aspartate in the active site, and the 2'-hydroxyl of OMP, and electrostatic repulsion through the proximity of a different aspartate, D70, to the leaving group carboxylate of OMP. Analysis between the simultaneous and stepwise ensembles demonstrated that the simultaneous ensemble adopted shorter distances for both features, generally suggesting stronger interactions. Both features were additionally shown to be associated with the ability to distort the planarity of the orotidyl ring, where shorter distances for either feature were correlated with larger degrees of distortion. Taken together, this suggested the simultaneous ensemble was more effective at distorting the ground state structure prior to crossing the reaction barrier.

**Thesis Supervisor:** Bruce Tidor

**Title:** Professor of Biological Engineering and Computer Science