

## Putative Benzoylformate Decarboxylases and the Annotation Problem

**Ahmed M. Malik**<sup>1</sup>, Matthew G. Logsdon<sup>1</sup>, Forest H. Andrews<sup>1</sup>, and Michael J. McLeish<sup>1</sup>

<sup>1</sup>Department of Chemistry and Chemical Biology, School of Science, IUPUI

Benzoylformate decarboxylases (BFDCs) are a relatively uncommon class of thiamin diphosphate-dependent enzymes of commercial interest that catalyze the decarboxylation of benzoylformate, with BFDC from *Pseudomonas putida* being the most extensively studied among them. Based upon sequence homology, the recently established Thiamin Enzyme Engineering Database (TEED) has identified dozens of sequences in a variety of other microorganisms and annotated them as BFDCs. Interestingly, the majority of these putative BFDCs share >40% sequence identity with *Pp*BFDC but many lack certain amino acids thought to be important in its function. Further, most of the annotated sequences are from organisms with no known pathway involving benzoylformate. To determine the integrity of these sequence annotations, these previously unstudied enzymes must be functionally characterized to determine if they are, in fact, true BFDCs. Currently, we are studying putative BFDCs from *Polynucleobacter necessaries* and *Mycobacterium smegmatis*, both of which share most of the same catalytic residues as *Pp*BFDC but have alterations in residues involved in substrate specificity. We have successfully expressed, purified these supposed BFDCs, and characterized them by assaying with a variety of both metabolic and unnatural 2-ketoacid substrates. Comparison of their activity with that of *Pp*BFDC suggests that these two sequences were both incorrectly annotated. We are currently in the process of trying to identify their natural substrate.

Mentor: Michael J. McLeish, Department of Chemistry and Chemical Biology, School of Science, IUPUI