Prediction of enzymatic functions – Identification of a galactarate dehydratase from *Oceanobacillus iheyensis*

John F. Rakus and John A. Gerlt

In the post-genomic era, reliable annotation of the function of new protein sequences presents a significant challenge. To date, over 6.5 million unique sequences have been deposited in protein databases but a majority of these are either annotated vaguely, incorrectly or not at all. With the frequency at which new genome sequencing projects are being completed, the disparity between protein discovery and proper functional annotation is only widening. The mechanistically-diverse enolase superfamily is a model for functional evolution of homologous enzymes. Members of this superfamily share common mechanistic and structural features. With functions known for fewer than half its members, we are focused on identifying new functions in this superfamily. Herein I describe work to identify a new function in the mandelate racemase (MR) subgroup of the enolase superfamily. An *apo* crystal structure of a highly divergent MR homologue from the thermophile *Oceanobacillus iheyensis* was solved and *in silico* ligand docking in this active site predicted a dianionic carbohydrate-derived substrate. The protein was screened for dehydration with a library of probably substrates which revealed the function as a *meso*-galactarate dehydratase with k_{cat} = 8.0 s^{-1}, k_{cat}/K_M = 1.1 x 10^4 $M^{-1}s^{-1}$. Curiously, this is the second example of the galactarate dehydratase function within the MR subgroup. However, these functions are not strictly orthologous because they catalyze dehydrations at regiochemically different positions on the *meso*-galactarate substrate.