

## **A Cellulase Family Reunion: Observing and Predicting the Structural Changes Accompanying the Evolution of GH5 Enzyme Specificity**

Kirk Vander Meulen<sup>1\*</sup> (kirk.vandermeulen@wisc.edu), Evan Glasgow<sup>1\*</sup> (eglasgow@wisc.edu), Taichi Takasuka<sup>2</sup>, Christopher Bianchetti<sup>3</sup>, Sam Deutsch<sup>4</sup>, and **Brian G. Fox**<sup>1</sup>

<sup>1</sup>Great Lakes Bioenergy Research Center, University of Wisconsin – Madison; <sup>2</sup>Hokkaido University, Sapporo, Hokkaido, Japan; <sup>3</sup>University of Wisconsin – Oshkosh; <sup>4</sup>Joint Genome Institute, Walnut Creek, CA.

**Project Goals: Synthesize and screen a subfamily of polysaccharide-degrading enzymes, determine the sequence and structural determinants of substrate specificity, predict functions of unknown enzymes, and model the evolution of enzymatic activities.**

We have performed a large-scale analysis of the diverse activities exhibited by members of subfamily 4 in glycoside hydrolase family 5 (GH5\_4), a subfamily notable for the range of substrate specificities among its members. Using high-throughput cell-free expression, we tested for the specificities exhibited by 237 GH5\_4 enzymes. Of the activities tested, lichenase activity was the most prevalent (present in 85% of enzymes), followed by xylanase (70%) and mannanase (44%). Two loop regions, each connecting a TIM-barrel core beta-strand with its respective C-terminal alpha-helix, contribute to the geometry of the GH5 hallmark binding channel, and each plays a critical role in determining activity and specificity. One loop is primed for contacts with the positive subsites of a substrate oligosaccharide, and correspondingly an absence of all activities is highly correlated with the absence of a channel-exposed tryptophan. The second critical loop is located both across and upstream from this channel position, and is correspondingly known to form contacts with substrate negative subsites. We have found that in one entire clade of enzymes, both loops are truncated considerably, affecting a striking change in binding channel geometry.

*DOE Great Lakes Bioenergy Research Center and Joint Genome Institute are supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research, through contracts DE-FC02-07ER64494 (GLBRC), and DE-AC02-05CH11231 (JBEI/JGI), respectively.*