Glycosyltransferases (GTs) from the Arabidopsis CAZy Family: High-throughput Cloning of a Library of GT and GT-related Genes

Peter McInerney[1], Stephanie Morrison[1], Hou Cheng Chu[1], Ai Oikawa[2], Henrik Scheller[2], **Pam Ronald**[2], **Paul Adams**[1] and Masood Hadi[1] Presenting author: [1]Masood Hadi – MZHadi@lbl.gov Technology1 and Feedstocks2 Division, Joint BioEnergy Institute, 5885 Hollis Avenue, Emeryville, CA 94608

The CAZy (Carbohydrate Active EnZyme) family of enzymes includes glycosyltransferases (GTs), glycosylhydrolases (GHs), polysaccharide lyases (PLs), carbohydrate esterase (CEs) and carbohydrate binding modules (CBMs). Many enzymes in this family are involved in various aspects of plant cell wall metabolism. The GTs represent one of the most diverse CAZy groups, with 91 separate protein families (not including non-classified sequences) that are assigned based on 3D protein structure, catalytic mechanism and donor/acceptor substrate requirements. In the simplest terms, GTs catalyze the transfer of sugar molecules from a donor molecule to an acceptor. However, the seemingly limitless combination of specific sugar, donor and acceptor molecules underscores the necessity for a large number of enzymes of this family. Our group is undertaking an effort to clone all 455 GTs in the CAZy database from Arabidopsis thaliana as well as the 90 GT-like proteins identified from other bioinformatic analyses. This library of GT clones will be a valuable resource at JBEI for a wide range of applications. At the level of biofuels research, results applicable to cell wall engineering are expected, based on the observation that a significant proportion of GT genes play roles (or are proposed to play roles) in cell wall metabolism. Furthermore, from a basic science standpoint, a great deal of new information should result from the study of these genes, since many of these genes are hypothetical or have unknown functions. Our approach relies heavily on automation, for informatics steps such as PCR primer generation and DNA sequence analysis, in addition to laboratory robotics, for assembly of enzymatic reactions and purification steps. To date clones for 80% of the targets have gone through the pipeline and are being sequence verified. Current efforts are centered on maximizing the number of clones that perfectly match the target DNA sequence by optimizing our cloning workflow. Once production of sequence-validated clones is complete, we will transfer genes to vectors suited to specific needs (e.g. expression for biochemical analyses or crystallography trials, fluorescence localization studies, etc.).

This work was part of the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.