NREL
NATIONAL RENEWABLE ENERGY LABORATORY

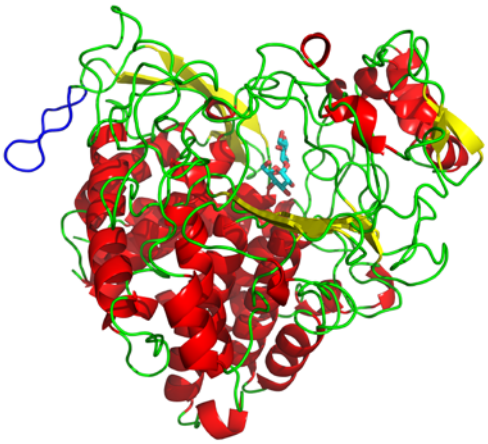
New Cellulase Identification Method Holds Promise for Lower-Cost Biofuels

Highlights in Science

A new computational approach to genomic data effectively distinguishes cellulases and non-cellulases within the protein family GH48, a key component for degrading lignocellulose for biofuels.

Cellulase enzymes, particularly from the glycoside hydrolase family 48 (GH48), are a critical component of many lignocellulose-degrading systems that produce biomass to use for biofuels. Typically it has been difficult and costly—but necessary—to distinguish required cellulase enzymes from unwanted non-cellulases in GH48. The expense has been a major economic impediment to commercialization of biofuels. Collaborating researchers from the National Renewable Energy Laboratory (NREL), the University of Tennessee, and Oak Ridge National Laboratory within the BioEnergy Science Center show that by using a robust computational approach supported by experimental studies, cellulases and non-cellulases within a given protein family can be effectively identified by genome sequence and structural features, potentially reducing the cost of biofuels production.

Although computational mining of large genomic data sets is a promising new approach for identifying cellulase activities, these computational methods are unable to distinguish between cellulases and enzymes with different substrate specificities that belong to the same protein family. The researchers show that cellulases from GH48 have distinct, evolutionarily conserved sequence and structural features that can be used to differentiate cellulases from non-cellulases in genomic data sets.



***Structure of GH48 from* H. chejunsis.** *The additional ω-loop identified in all cellulases is labeled in blue. The α-helices are shown in red, β-strands in yellow, and loops in green. The cellobiose molecule is shown with carbon atoms in cyan and oxygen atoms in red.*
Image by Markus Alahuta, NREL

The research team proposes that the structural element that can be used for in silico discrimination between cellulases and non-cellulases belonging to GH48 is an ω-loop located on the surface of the molecule and characterized by rare, highly conserved amino acids. These markers were used to screen metagenomics data for "true" cellulases. Such unambiguous identification of cellulases in genomic data is critical in searching for novel cellulolytic activities needed for bioenergy research.

Technical Contact: Vladimir Lunin, *vladimir.lunin@nrel.gov*

Reference: Sukharnikov, S.O.; Alahuhta, M.; Brunecky, R.; Upadhyay, A.; Himmel, M.E.; Lunin, V.V.; Zhulin, I.B. (2012). "Sequence, Structure, and Evolution of Cellulases in Glycoside Hydrolase Family 48." *The Journal of Biological Chemistry*, Vol. 287, No. 49, pp. 41068–41077.

Key Research Results

Achievement

Researchers show that by using a robust computational approach supported by experimental studies, cellulases and non-cellulases can be effectively identified within a given protein family, specifically GH48.

Key Result

Computational and structural studies of GH48 enzymes identified conserved sequence and structure features that define cellulases and that can be used to computationally distinguish them from non-cellulases.

Potential Impact

The ability to reliably identify cellulase enzymes from non-cellulases in GH48 may reduce the cost and time otherwise required for the production process, potentially accelerating the commercialization of biofuels.

NREL is a national laboratory of the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, operated by the Alliance for Sustainable Energy, LLC.

15013 Denver West Parkway
Golden, CO 80401
303-275-3000 | www.nrel.gov

NREL/FS-2700-58228 | April 2013

Printed with a renewable-source ink on paper containing at least 50% wastepaper, including 10% post consumer waste.