www.emsl.pnl.gov

EMSL

Science Made Possible

De Novo Sequencing Characterizes Protein Modifications

Research featured on cover of top 10 analytical chemistry journal

To improve proteomics studies of heart disease, cancer, neurodegenerative diseases, and diabetes, scientists at the Department of Energy's EMSL developed a novel strategy that expands protein identifications beyond those predicted from genomic sequences. Their approach combines *de novo* sequencing and the UStag approach to discover post-translational modifications in proteins. *De novo* sequencing derives a peptide sequence without the help of a database of known protein sequences. The UStag approach uses a peptide sequence to unambiguously and uniquely identify a single protein.

A new approach that characterizes previously unknown protein modifications was featured on the cover of Analytical Chemistry.

Scientific impact: The combined *de novo*—UStag approach, funded in part by EMSL's 2007 Intramural Research Program, allows scientists to discover protein post-translational modifications, including complex multiple unknown/unexpected modifications on single protein sequences, and discover sequence mutations and genome-predicted database sequence errors. This research and other discoveries are advancing EMSL's goal of being able to predict biological functions from molecular and chemical data.

Societal impact: Unambiguous and highly confident protein identifications can help better identify changes that signal disease, potentially improving techniques for early disease detection and monitoring changes in environmentally important organisms.

For more information, contact EMSL Communications Manager Mary Ann Showalter (509-371-6017).

References: Shen Y, N Tolic, KK Hixson, SO Purvine, GA Anderson, and RD Smith. 2008. "*De novo* sequencing of unique sequence tags for discovery of post-translational modifications of proteins." *Analytical Chemistry* 80(6):7742-7754.

Shen Y, N Tolic, KK Hixson, SO Purvine, L Paša-Tolic, WJ Qian, JN Adkins, RJ Moore, and RD Smith. 2008. "Proteome-wide identification of proteins and their modifications with decreased ambiguities and improved false discovery rates using unique sequence tags." *Analytical Chemistry* 80(6):1871-1882.

Acknowledgments: This research was funded by EMSL's 2007 Intramural Program, the DOE Office of Biological and Environmental Research, and the NIH National Center for Research Resources.