EMSL Research and Capability Development Proposals

Developing calculation methods to precisely measure proteome proteins and their posttranslational modifications

Project timeline: Spring 2007–Fall 2009

EMSL Lead Investigator:

Nikola Tolić

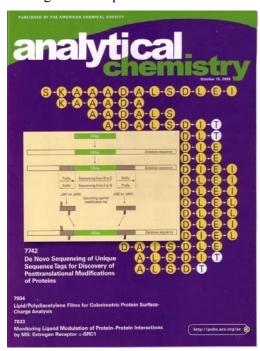
Instrument Development Laboratory, EMSL, PNNL

Co-investigators:

Yufeng Shen, FCSD, PNNL Kim Hixson, Mass Spectrometry Facility, EMSL, PNNL Samuel Purvine, Instrument Development Laboratory, EMSL, PNNL

Unique sequence tags—UStags—are a new method for protein identifications from tandem mass spectrometry (MS/MS) experiments. It assumes high resolution on fragment ion species and relies on

precise mass measurements. Thus, it is targeted toward data analysis of datasets produced on hybrid linear trap quadrupole-Fourier transform (LTQ-FT), LTQ-Orbitrap, and Fourier transform ion cyclotron resonance (FTICR) class of instruments. We have developed the software— Tagitus—capable of *de novo* sequencing of unique sequence tags and comparison of results with the peptide candidates pool derived from the proteome database. Although not new per se, the *de novo* sequencing idea went through a recent renaissance brought on via advancements in mass spectrometry instrumentation and the realization of the importance of mass measurement precision for reduction of incorrect peptide assignments and inaccurate error rate estimates through poorly defined analysis of false discovery rates. The difference of this method is that accurate mass determination eliminates the vast majority of random de novo hits, and sequencing of consecutive peptide fragments allows for unrestricted search of post-



translational modifications and protein changes due to genomic mutations and transcription errors. To the best of our knowledge, this is the first sequencing software designed and modeled exclusively for high-resolution tandem mass spectra with the capability of generating identifications with near-zero false discovery rates.

Products and Output

New Capability for EMSL Users

EMSL users are able to employ the Tagitus software for custom data analysis in the EMSL mass spectrometry facility.

Publications¹

Shen, Y., N. Tolić, K.K. Hixson, S.O. Purvine, G.A. Anderson, and R.D. Smith. 2008. "De Novo Sequencing of Unique Sequence Tags for Discovery of Post-Translational Modifications of Proteins." *Analytical Chemistry* 80(20):7742-7754. DOI: 10.1021/ac801123p.

Shen, Y., K.K. Hixson, N. Tolić, D.G. Camp, S.O. Purvine, R.J. Moore, and R.D. Smith. 2008. "Mass Spectrometry Analysis of Proteome-Wide Proteolytic Post-Translational Degradation of Proteins." *Analytical Chemistry* 80(15):5819-5828. DOI: 10.1021/ac800077w.

Shen Y., N. Tolic, K.K. Hixson, S.O. Purvine, L. Paša-Tolic, W-J Qian, J.N. Adkins, R.J. Moore, and R.D. 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. DOI: 10.1021/ac702328x. (Cover feature)

Currents: Journal of Proteome Research. April 4, 2008. "Unambiguous protein identification by unique sequence tags." 7(4):1369. DOI: 10.1021/pr083725r.

Highlight

http://www.emsl.pnl.gov/news/inbriefs/docs/shen20080521.pdf

http://www.emsl.pnl.gov/news/highlights/shen20081208.pdf

¹ After project closeout, work using the UStags method continued through the method development and applications of projects studying human blood plasma peptidome and degradome. At the moment, two additional manuscripts accepted for publication in the *Journal of Proteome Research* and one submitted to *PLoS ONE* are direct products of the work on and developments stemming from the UStags project.