

Biological Sciences Division

Simulation of PNNL protein identification process featured in journal special edition

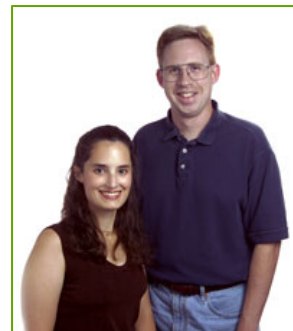
The results of a computational study done by Pacific Northwest National Laboratory scientists will be published in the August (Special Edition honoring John Yates III) issue of the *Journal of the American Society for Mass Spectrometry (JASMS)*. A figure from that *JASMS* article will be prominently featured on the cover of the issue (Figure 1).

The paper, entitled “The utility of accurate mass and LC elution time information in the analysis of complex proteomes,” presents an analysis of a PNNL-developed protein/peptide characterization method. Authors are Angela Norbeck, Matthew Monroe, Kevin Anderson, Don Daly, Joshua Adkins, and Richard Smith from PNNL. The work is sponsored by the Department of Energy’s Office of Biological and Environmental Research, the National Institute of Allergy and Infectious Diseases, and PNNL’s Biomolecular Systems Initiative.

Mass spectrometry is a technique that can be used to observe the masses of proteins or peptides, thereby providing accurate identifications from biological samples. For very complex samples, such as found in human cells, peptides generally cannot be differentiated by mass alone; mass measurements must be combined with additional information to provide a unique characterization of each peptide.

The PNNL article explores the applicability of a method in which a combination of an accurate mass measurement with a normalized elution time (NET) of a peptide, as observed by liquid chromatography coupled to mass spectrometry, is used to create a unique peptide signature in the context of complex biological systems.

Two-dimensional reference databases, defined by mass and elution time, can be used to locate such peptide signature points for protein identification. The paper addresses the applicability of this signature method to a variety of peptide reference databases and examines an approach to addressing ambiguities when two peptide signatures have similar masses and normalize elution



Angela Norbeck
(angela.norbeck@pnl.gov)
and Matthew Monroe
(matthew.monroe@pnl.gov)

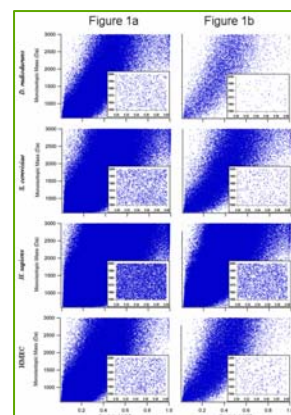


Figure 1. a) Global representation of tryptic digests of all proteins for all four systems studied. b) Cysteine-containing peptides from tryptic digests for all four systems studied. Predicted Normalized Elution Time (NET) is plotted along the x-axis, and monoisotopic mass in Daltons is plotted along the y-axis. Inset views are representative of the region contained within 1950–2000 Da and 0.3–0.4 NET.

times. The creation of a signature from NET values, in addition to clarifying ambiguity, represents an improved approach to identifying proteins using data acquired on high-resolution, high mass accuracy LC-MS instrumentation.

Mass-NET values of a range of peptides were predicted using Protein Digestion Simulator, a PNNL-written computer program that is freely available to the public. The values were then placed into a two-dimensional reference space, and the probability of a match between the actual measurement and the signature reference point was calculated. Despite inherent uncertainties, it was found that accurate mass and elution time values provide the basis for an acceptable level of certainty in peptide identification. Further, the researchers show that the value of these approaches can be augmented by using specific chemistries in the sample preparation techniques for isolating select peptides (see Figures 2 and 3).

This unique mass-NET signature method of characterization has broad application in proteomics, benefiting both the throughput and sensitivity of measurements.

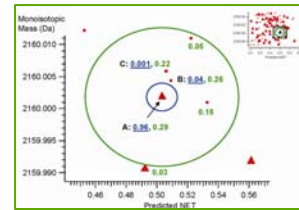


Figure 2. Granular view of dense region from *H. sapiens*. Cysteine-containing peptides are represented as triangles. The SLiC scores for ± 1 ppm/0.01 NET are underlined, while those for ± 5 ppm/0.05 NET are present for each data point. The peptide sequences for points A-C are LVWEEAMSRFCEA-EFSVK, FGLLMVENLEE-HSEASNIE, and DDLDE-QIRHMLFSWAER, respectively.

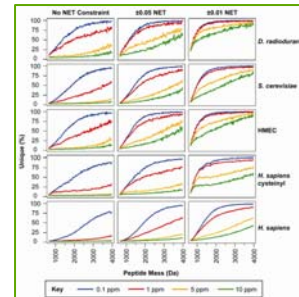


Figure 3. Percent of peptides that are unique (i.e., distinctive) versus peptide mass for the four systems for no NET constraint and for ± 0.05 and 0.01 NET, as well as for different levels of mass accuracy. Cysteinylyl-only peptides are shown for *H. sapiens* in addition to all peptides for *H. sapiens*.