

Open Source Tools for the Accurate Mass and Time (AMT) Tag Proteomics Pipeline

Navdeep Jaityl, Matthew E. Monroe, Nikola Tolić, Kyle Littlefield, Don S. Daly, Joshua N. Adkins, Gordon A. Anderson, and Richard D. Smith
Pacific Northwest National Laboratory, Richland, WA 99352

Overview

- The accurate mass and time (AMT) tag pipeline combines liquid chromatography (LC)-mass spectrometry (MS) and tandem mass spectrometry (MS/MS) analyses to provide high confidence identifications and quantitative information¹.

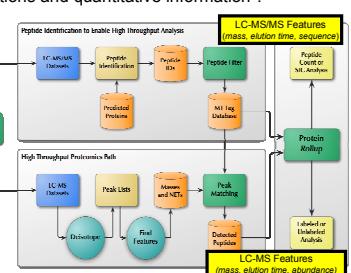


Figure 1. The AMT tag pipeline for quantitative, high throughput proteomics

- Implementation of the pipeline depends on software for:
 - Feature discovery in LC-MS and LC-MS/MS experiments
 - Alignment of LC-MS datasets
 - Management of datasets
 - Execution of a workflow of processing steps
- Tools required for implementation of the AMT tag pipeline are being made available at: <http://ncrr.pnl.gov>
- We describe here our tools for deisotoping of high mass accuracy data (Decon2LS) and chromatographic peak picking and peak matching (Viper and MASIC)

Introduction

Accurate Mass and Time (AMT) tag strategy

- Analyze peptide mixtures
 - Gradient LC separations coupled to ion trap MS/MS analyses to obtain large numbers of potential mass and time (PMT) tags with calculated accurate mass and observed normalized elution time (NET)
 - High throughput LC-FTICR-MS detection to measure intact peptide masses and elution times for various conditions
 - Find features, i.e., unique mass classes (UMCs)
- Align and match features to PMT tags

Methods

LC-MS Feature discovery: Features are found by deisotoping each scan using a modified implementation of the THRASH algorithm² and by finding elution time profiles of deisotoped masses over multiple scans.

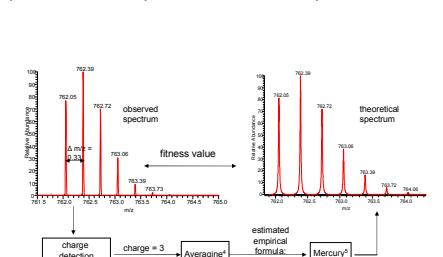


Figure 2. THRASH² algorithm for deisotoping a high resolution mass spectrum



Figure 3. The same monoisotopic mass detected in multiple scans is grouped together into one chemical species with an elution profile

LC-MS/MS Feature discovery: Selected ion chromatograms (SICs) are constructed for each parent ion fragmented. These can be tied with identification from, for example, X!Tandem or SEQUEST. Each identified peptide is converted into an LC-MS/MS feature with monoisotopic mass, elution time and sequence information. SICs from multiple LC-MS/MS experiments can be used for quantitative comparisons.

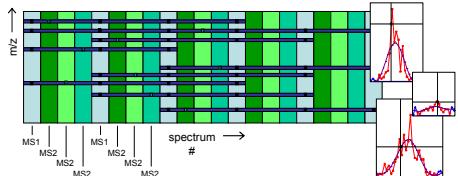


Figure 4. SICs for parent ions are constructed from parent MS1 scans for each ion chosen for fragmentation in an LC-MS/MS analysis.

Results

Decon2LS Functionalities

- Calculate expected composition using Averagine
- Create theoretical isotopic profile using Mercury
- Score observed vs. theoretical distribution
- Deisotope normal and isotopically labelled samples
- Read various formats: mzXML, mzData, Micromass TOF, Finnigan .Raw, Bruker, and Agilent TOF.

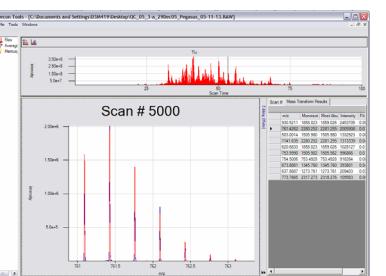


Figure 5. Decon2LS: Software for deisotoping high resolution LC-MS experiments. The top part of the figure shows the TIC for the spectrum. The bottom shows an observed isotopic profile (blue) and the best fit theoretical distribution (red) overlaid on top.

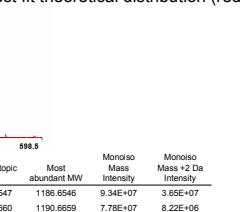


Figure 7. Example of the isotopic profile of a peptide and an isotopically labelled pair where the isotopically labelled peptide is heavier than the unlabelled peptide by 4 Da because of the presence of two O¹⁸ atoms instead of two O¹⁶ atoms.

LC-MS FEATURE DISCOVERY

LC-MS/MS FEATURE DISCOVERY

MASIC: MS/MS Automated Selected Ion Chromatogram

- Process LC-MS/MS data to:
 - Generate scan-based statistics
 - Generate dataset-wide chromatograms
 - Summarize chromatographic statistics for m/z values chosen for fragmentation
- Several statistics computed for each SIC peak
- Available as GUI, command line and DLL
- Read various formats: mzXML, mzData, Finnigan .Raw, and .cdf/.mgf combo files.

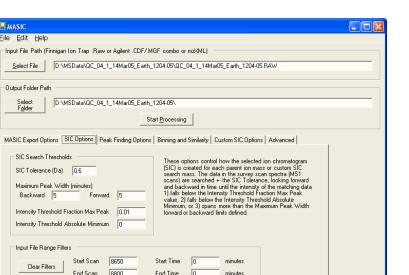


Figure 9. MASIC: Software for constructing selected ion chromatograms.

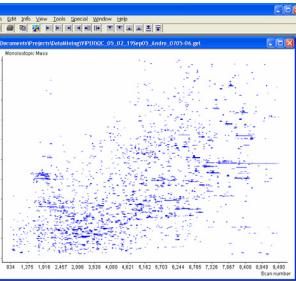


Figure 6. VIPER: Software for clustering deisotoped masses from high resolution LC-MS experiments into features. The figure shows deisotoped masses vs the scan number they were found in. Features are seen over a range of spectra in a chromatographic profile.

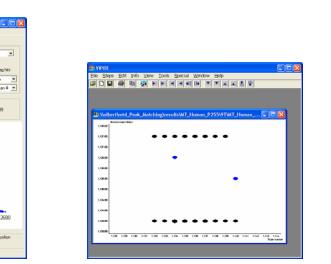


Figure 8. Different utilities in Viper. (a) Feature browser. The chromatographic profile of an LC-MS feature can be viewed in this browser. (b), (c) Isotopically labeled pairs can be viewed in mass and time dimensions and as chromatograms for relative quantitation.

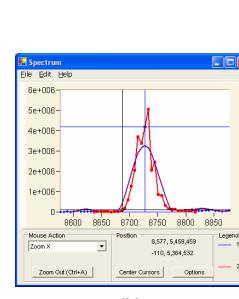
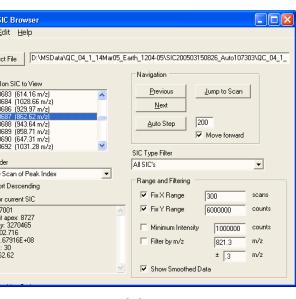


Figure 10. (a) MASIC Browser can be used to browse results from analysis of an LC-MS/MS dataset with MASIC. (b) The selected ion chromatogram of one ion chosen for MS/MS. The scan at which the ion was chosen for fragmentation is shown in black and the peak top of the chromatographic peak is shown in blue.

Conclusions

- We have applied and leveraged our experience in informatics and mass spectrometry as well as our knowledge of important metrics to develop improved open source versions of our analytical tools for the proteomics community.
- Source code and binary versions of our software are available at <http://ncrr.pnl.gov/software>

Acknowledgements

Portions of this research were supported by the National Institute of Allergy and Infectious Diseases (NIH/DHHS interagency agreement Y1-AI-4894-01), the NIH National Center for Research Resources (RR18522), and by the U.S. Department of Energy (DOE) Office of Biological and Environmental Research. Experimental portions of this research were performed in the Environmental and Molecular Sciences Laboratory, a DOE national scientific user facility located at Pacific Northwest National Laboratory (PNNL) in Richland, Washington. PNNL is a multiprogram national laboratory operated by Battelle for the DOE under Contract No. DE-AC05-76RLO 1830.

References

- Smith, R. D.; Pasar-Tolic, L.; Lipton, M. S.; Jensen, P. K.; Anderson, G. A.; Shen, Y.; Conrads, T. P.; Udseth, H. R.; Harkewicz, R.; Belov, M. E.; Masselon, C.; Veenstra, T. D. Electrophoresis 2001, 22, 1652-1668.
- Horn, D.M., Zubarev, R.A., McLafferty, F.W. Automated Reduction and Interpretation of High Resolution Electrospray Mass Spectra of Large Molecules. J. Am. Soc. Mass Spectrom. 2000, 11, 320-332
- Senko, M. W.; Beu, S. C.; McLafferty, F. W. Automated assignment of charge states from resolved isotopic peaks for multiplycharged ions. J. Am. Soc. Mass Spectrom. 1995, 6, 52-56.
- Senko, M. W.; Beu, S. C.; McLafferty, F. W. Determination of monoisotopic masses and ion populations for large biomolecules from resolved isotopic distributions. J. Am. Soc. Mass Spectrom. 1995, 6, 229-233.
- Rockwood, A. L.; Van Orden, S. L.; Smith, R. D. Rapid Calculation of Isotope Distributions. Anal. Chem. 1995, 67, 2699-2704.
- Zimmer, J. D.; Monroe, M. E.; Qian, W. J.; Smith, R.D. Advances in Proteomics Data Analysis and Display Using an Accurate Mass and Time Tag Approach. Mass Spec. Rev. 2006, 25, 450-482.

Contact Information

Navdeep Jaityl
Pacific Northwest National Laboratory, K9-91
P.O. Box 999, Richland, WA 99352
E-mail: navdeep.jaityl@pnl.gov

Matthew Monroe
Pacific Northwest National Laboratory, K8-98
P.O. Box 999, Richland, WA 99352
E-mail: matthew.monroe@pnl.gov