

Integration of HPLC-FTICR MS and HPLC-QIT MS² to Achieve Enhanced Proteome Characterization

Chongle Pan^{1,2,3} Nathan VerBerkmoes^{1,3} Praveen Chandramohan² Nagiza Samatova^{2,3} Robert Hettich^{1,3}

¹Chemical Sciences Division; ²Computer Science and Mathematics Division, Oak Ridge National Lab; ³Genome Science and Technology Graduate School, ORNL-University of Tennessee

OVERVIEW

Objectives

Proteome characterization generally consists of peptide separation by high performance liquid chromatography (HPLC) and peptide identification by tandem mass spectrometry (MS/MS). We aim to enhance the specificity and sensitivity of peptide identification by *in silico* integrating nanoLC-FTICR-MS and nanoLC-QIT-MS/MS.

Methods

- nanoLC-FTICR-MS: 1D reverse phase LC coupled online to IonSpec 9.4T FTICR with MS scans
- nanoLC-QIT-MS/MS: 1D reverse phase LC coupled online to LQCY with data dependent MS/MS scans
- *in silico* integration: Retention time normalization; peptide correlation by retention time and mass

Results

- Development of a robust nanoLC-FTICR-MS system
- Development of retention time normalization and peptide correlation
- Demonstration of enhanced peptide identification from simple protein mixture digest

INTRODUCTION

Proteomics

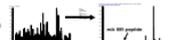
- Systematic analysis of protein complement in a given cell, tissue or organism for their identity, quantity and function.
- Major challenges: proteome coverage, low abundance proteins, membrane proteins, post-translationally modified proteins,
- General procedures: Gel-based or LC-based separation, MS-based characterization, informatics-based identification.

Current approaches to proteomics

- MudPit technology (1)
 1. Biphasic column separation of tryptic digested proteome, integrating SAX resin and RP resin in one column
 2. Introduction of eluent directly to QIT through ESI interface; data-dependent MS/MS scan of eluting peptides
 3. Identification of peptides by SEQUEST database searching (2)
 4. Assemble of peptide identification to protein identification by DTAselect (3)

Switching LC technology (4)

1. Autosampler loading of tryptic digested proteome onto a trapping cartridge
2. 1D HPLC separation or 2D switching (SICSW)/LC separation
3. QIT data-dependent MSⁿ scan of eluting peptides, possible multiple mass range scan
4. SEQUEST/MSMOT database searching and protein identification



- Advantages:
 - HPLC separation of peptides are relatively unbiased.
 - MS/MS scans are informative of peptide sequence
- Disadvantages:
 - The scores from MSⁿ database searching algorithm is the only base for identification calls
 - The selection of cutoff is a compromise between identification specificity and sensitivity
 - Common cutoff for SEQUEST (Xcorr (+1) > 1.8, (+2) > 2.5, (+3) > 3.5)

AMT technology (5)

"An AMT tag is a peptide with a sufficiently distinctive mass and LC elution time to act as a biomarker for a given protein."

1. Construction of potential mass tag (PMT) library
 - 1.1 Capillary HPLC / QIT MSⁿ characterization of a given organism's proteome under a variety of conditions
 - 1.2 Establishment of PMT for a peptide using a very liberal cutoff (SEQUEST Xcorr (+1, +2, +3) > 2.0)
2. Construction of accurate mass and time tag (AMT) library
 - 2.1 Capillary HPLC / FTICR MS characterization of those proteome samples
 - 2.2 Establishment of AMT for a peptide matched with its PMT's retention time and calculated mass
3. Proteome examination by capillary HPLC / FTICR MS
 - 3.1 Capillary HPLC / FTICR MS characterization of the proteome samples of interest
 - 3.2 Identification of protein by searching AMT library with measured mass and retention time

- Advantages:
 - High throughput proteome examination once AMT library is constructed
 - High proteome coverage given a confident AMT library
- Disadvantages:
 - Large initial effort investment in constructing AMT library
 - Multiple dimensional separation not straight-forward
 - False positive protein identification due to insufficiently distinctive mass and LC elution time of an AMT tag

NANOLC-FTICR-MS METHODS AND RESULTS

Methods:

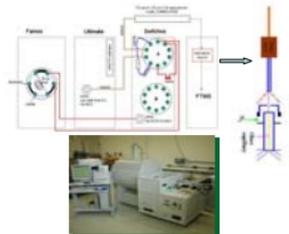
- 1D-reverse phase Liquid Chromatography
 - Autosampler injection, LC packings FMMOS 50 ul
 - Preconcentration: 300 um x 5 mm C18 PepMap
 - RP-LC: Vydac 75 um id x 25 cm C18 nanocolumn

Nanospray

- Tip: 10 um ID New Objective Picospip
- V_{capillary}: -1800V V_{0V} V_{nebulizer}: -1800V
- Distance: 3mm

FT ICR MS

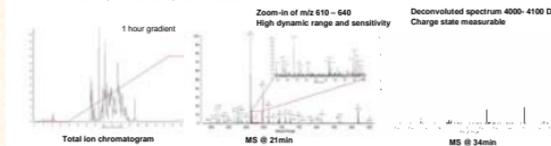
- 9.4 Tesla IonSpec
- 2 sec: hexapole ion accumulation
- 256K data points @ 1MHz ADC
- 2-scan signal averaging, ~ 9 sec: per spectrum



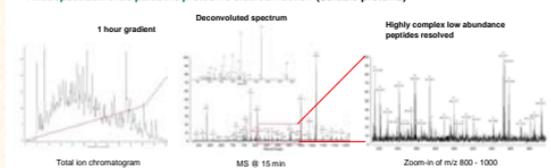
Results

Standard protein mixture tryptic digest

- Constituents: ALCOHOL DEHYDROGENASE I and II, CARBONIC ANHYDRASE, HEMOGLOBIN A and B chains, MYOGLOBIN, LYSOZYME C, SERUM ALBUMIN



Rhodopseudomonas palustris proteome cleared fraction (Soluble proteins)



3 consecutive nanoLC- FTICR MS runs with proteome samples



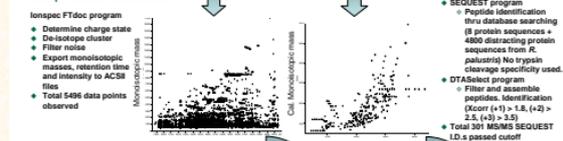
- Automatic 3 repetitive sample injections
- Highly reproducible results (mass spectra)
- A robust LC/MS system

INTEGRATION METHODS AND RESULTS

1. Parallel LC/MS experiments



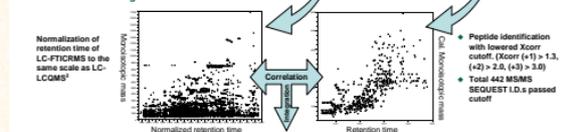
2. Data process and extraction



3. Normalization of retention time

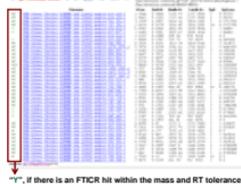
1. Selection of highly confident peptide hits from LQO data
Xcorr (+1) > 2.5, (+2) > 3.0, (+3) > 3.5 Total 208 I.D.
2. Search against FT ICR data with calculated monoisotopic masses
Mass error < 0.05 Da
3. Data representation and visual inspection to remove spurious hits
Deletion of the obvious outliers
4. Linear regression: RT_{LQO} = 0.9669*RT_{FT} - 11.5872
correlation coefficient = 0.94

4. Correlation and Integration



Mass tolerance < 0.05 Da
RT tolerance < 3 min
Xcorr (+1) > 1.3, (+2) > 2.0, (+3) > 3.0

Export of results to DTASelect readable format. Use of manual validation flag to flag the presence of FTICR data hits



"", if there is an FTICR hit within the mass and RT tolerance

LC QIT MS² METHODS

Sample preparation and protein digestion

- The protein sample is denatured with 6-M Guanidine or Urea, and reduced with DTT or some other reducing agent at 60°C for 10-60 minutes.
- The denaturant concentration is lowered by dilution with Tris; sequencing grade trypsin is added to the sample and incubated overnight
- The sample is desalted by solid phase extraction and organic solvent is removed by SpeedVac

1D LC / QIT MS²

- One-dimensional LC-MS/MS experiments were performed with an Ultimate HPLC LC Packings, a division of Dionex, San Francisco, CA) equipped with an LQO-DECA ion trap mass spectrometer (Thermo Finnigan, San Jose, CA) coupled with an electrospray source. Injections were made with a Famos LC Packings autosampler onto a 50ul loop. Flow rate was ~4ul/min with a 240min gradient for each run.
- A VYDAC 218MS5.325 (Grace-Vydac, Hesperia, CA) C18 column (300um id x 15cm, 300A with 5um particles) or a VYDAC 238IEV5.325 monomeric C18 (300um id x 15cm, 300A with 5um particles) was directly connected to the Finnigan electrospray source with 100um id fused silica.
- For all 1D LC/MS/MS data acquisition, the LQO was operated in the data dependent mode with dynamic exclusion enabled, where the top four peaks in every full MS scan were subjected to MS/MS analysis.

SEQUEST and DTASelect

- The resultant MS/MS spectra from the sample were searched with SEQUEST against the six constituent protein sequence and all predicted ORFs from R. palustris. The ORFs from R. palustris serves as indication of false position rate.
- The raw output files were filtered and sorted with DTASelect. The filter criteria include the minimal Xcorr, the validation flag for FT ICR data hits and the FTICR mass measurement error

CONCLUSIONS

Development of a robust nanoLC-FTICR-MS system

- Stable electrospray and good sensitivity in highly aqueous solution
- Low sample consumption (~ 1 ug of total peptide loaded)
- Good mass accuracy, mass resolution, dynamic range, and reproducibility

Development of retention time normalization and peptide correlation

- Use of linear regression to offset gradient start time and normalize gradient slope
- High correlation coefficient in retention times between LC-FTICR-MS and LQO
- Demonstration of comparability of FTICR data and QIT data

Demonstration of enhanced peptide identification from simple protein mixture digest

- FTICR data used as a validation method for SEQUEST identification
- Improved specificity and sensitivity in peptide identification for simple protein mixtures; should be much more enhanced for proteomes
- Improvements in the mass measurement accuracy of FTICR and retention time reproducibility of LC system should provide more narrow windows
- Improvements in peptide identification scoring should provide a more rigorous method to integrate FTICR accurate mass measurements and QIT MS/MS measurements.

REFERENCES

1. Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR 3rd *Nat Biotechnol*. 1999 17, 676-82.
2. Eng JK, McCormack AL, Yates JR 3rd *J. Am. Soc. Mass Spectrom*. 1995 6, 1426-1436.
3. Tabb DL, McDonald WH, Yates JR 3rd *J. Proteome Res*. 2002 1, 21-6.
4. VerBerkmoes NC, Bundy AL, Hasseril L, Asano KG, Razumovskiy J, Lammer F, Hettich RL, Stephenson JL Jr *J. Proteome Res*. 2002 1, 239-52.
5. Smith RD, Anderson GA, Lipton MS, Masselon C, Pasa-Tolic L, Shen Y, Udelski RH. *OMICS*. 2002 6, 61-90

Acknowledgement:

- Dr. David Tabb and Dr. Hayes McDonald are acknowledged for technical input and discussions.
- C.P. and N.V. thank Genome Sciences and Technology graduate program for financial support.
- Research was sponsored by the U.S. Department of Energy, Office of Biological and Environmental Research. Oak Ridge National Laboratory is operated and managed by the University of Tennessee-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725