Functional characterization of microbial proteins using top-down proteomics

Si Wu[1], Charles Ansong[2], Da Meng[3], Rui Zhao[1], Nikola Tolić[1], Li Cao[2], Yi Qu[2], Samuel Payne[2], Ronald J. Moore[2], Mary Lipton[2], Joshua N. Adkins[2], and Ljiljana Paša-Tolić[1]

[1]Environmental Molecular Sciences Laboratory, [2]Biological Sciences Division, [3]Computational Science and Mathematics Division, Pacific Northwest National Laboratory, Richland, WA, 99352

EMSL at Pacific Northwest National Laboratory

Overview

- Development of a high throughput LC-MS platform for comprehensive and sensitive intact protein identification.
- Application of the optimized platform for top-down microbial proteomics.
  - *Escherichia coli*
  - *Salmonella* Typhimurium
  - *Novosphingobium aromaticivorans*

Introduction

Functional characterization of the entire microbial proteome is often complicated by various protein processing events and posttranslational modifications (PTMs) that are frequently untraceable using traditional bottom-up approach.

While top-down analyses provide this information, throughput and sensitivity have been limiting factors due to extensive pre-fractionation efforts required for adequate proteome coverage.

To tackle this challenge, we have developed a high throughput LC-MS platform for comprehensive top-down characterization of microbial proteome.

Specifically, we have optimized commercially available LC-MS platforms for top-down analysis using an *E. coli* lysate, and applied this workflow to other microbial systems, including *S.* Typhimurium *and N. aromaticivorans.*

CONTACT: Si Wu, Ph.D.
Environmental Molecular Sciences Laboratory
Pacific Northwest National Laboratory
P.O. Box 999, Richland, WA 99352
E-mail: si.wu@pnnl.gov

Methods

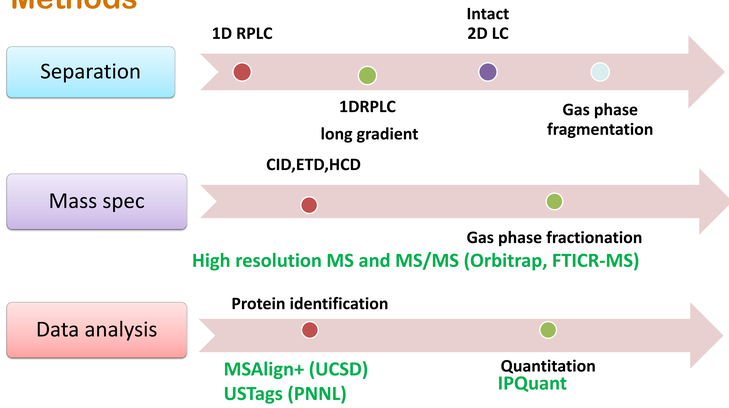

Figure 1. Overall experimental design. Intact proteins were extracted from microbial cells using a bead-beating method. The separation method, MS analysis, and data analysis were optimized for intact microbial protein analysis.

Intact *E coli* proteome analysis

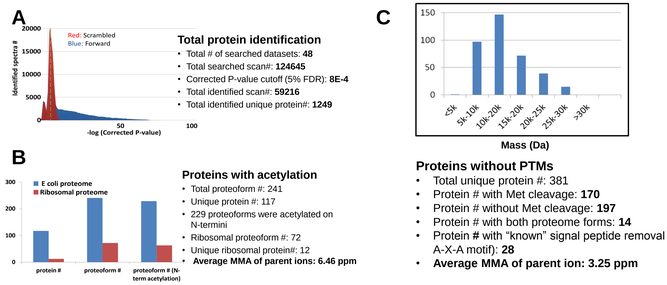

Figure 2. Intact *E. coli* proteome analysis. **A)** Protein identification summary. Proteins were identified using MSAlign+[2] (FDR<5%); fragment ion error tolerance was 15 ppm, and up to two mass shifts were allowed. **B)** Proteins identified with acetylation. Our results indicated that most of the acetylation events occurred on protein N-termini. **C)** Proteins identified in their intact form and without PTMs.

Top-down proteomic analysis of *S.* Typhimurium response to infection-like conditions[3]

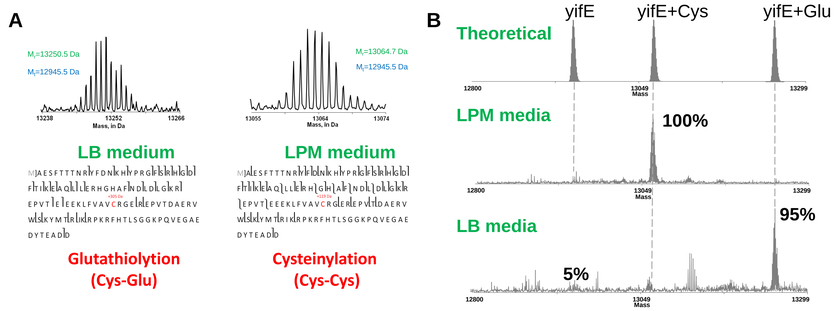

Figure 3. Switch in modification from glutathionylation to cysteinylation in response to environmental conditions. **A)** Representative zero charge state spectra from top-down LC-MS data showing switch in S-thiolation forms for the hypothetical protein YifE (STM14_4694). Fragmentation map shows specific cysteine residue on which switch occurred. **B)** Proteoform stoichiometry can be estimated from intact protein mass spectra summed across the corresponding LC peaks.

Analysis of intact periplasmic proteome from *N. aromaticivorans*[4]

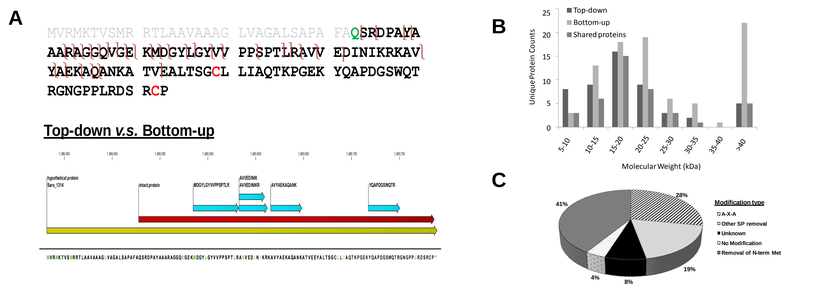

Figure 4. A) Top-down and bottom-up analysis of the hypothetical protein Saro_1314. Upper panel shows the fragmentation ion map illustrating highly confident identification ("Q" highlighted in green was modified as pyroglutamic acid, and two "C" highlighted in red formed a disulfide bond). Lower panel shows sequence coverage achieved using top-down approach (red arrow) vs. bottom-up approach (blue arrows). **B)** Molecular mass distribution of proteins identified using top-down and bottom-up analysis. Theoretical molecular masses were calculated using amino acid sequence. **C)** Pie chart representation of protein modifications observed by top-down LC-MS.

Conclusions

- 1249 unique *E. coli* proteins from a total of 59216 tandem mass spectra were confidently identified (5% FDR), which corresponded to >3000 proteoforms.
- Top-down proteomic analysis of *S.* Typhimurium identified 563 unique proteins corresponding to1665 proteoforms.
- Top-down proteomics enabled discovery of the differential utilization of the protein S-thiolation forms, S-glutathionylation and S-cysteinylation, in response to infection-like conditions (vs. basal conditions).
- First report of S-cysteinylation in Gram-negative bacteria.
- Top-down analysis of intact periplasmic proteome from *N. aromaticivorans* confidently identified 55 proteins with various PTMs. This study provides the first experimental evidence for the expression and periplasmic localization of many hypothetical and uncharacterized proteins and the first unrestrictive, large-scale data on PTMs in the bacterial periplasm.

Acknowledgements

This work were supported by EMSL intramural funding, the National Institute of Allergy and Infectious Diseases NIH/DHHS through Interagency agreement Y1-AI-8401, and the U. S. Department of Energy Office of Biological and Environmental Research (DOE/BER). Work was performed in the Environmental Molecular Sciences Laboratory (EMSL), a DOE/BER national scientific user facility located on the campus of Pacific Northwest National Laboratory in Richland, Washington.

References

1. Tran JC, *et al.* (2011) Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature* 480:254-258.
2. Liu X, *et al.* (2012) Protein identification using top-down. *Mol Cell Proteomics* 11:M111 008524.
3. Ansong C, *et al.* (2013) Top-down proteomics reveals a unique protein S-thiolation switch in *Salmonella* Typhimurium in response to infection-like conditions. *Proc Natl Acad Sci U S A* 2013 May 29 [Epub ahead of print].
4. Wu S, *et al.* (2013) Top-Down Characterization of the Post-Translationally Modified Intact Periplasmic Proteome from the Bacterium *Novosphingobium aromaticivorans*. *Int J Proteomics* 2013:279590.