Rapid Profiling of Proteomes and Sub-Proteomes on the Orbitrap Fusion Tribrid Mass Spectrometer

Steven R. Danielson,[1] Todd Markowski,[2] LeeAnn Higgins,[2] Pratik Jagtap,[2] Tim Griffiin,[2,3] Yue Chen[3]
[1]Thermo Fisher Scientific, San Jose, CA; [2]Center for Mass Spectrometry and Proteomics, University of Minnesota, Minneapolis MN; [3]Dept. of Biochemistry, University of Minnesota, Minneapolis, MN

Thermo SCIENTIFIC

Overview

Purpose: High throughput characterization of proteomic samples by combining rapid immunoaffinity enrichment sample preparation, fast liquid chromatography and detection utilizing very sensitive, rapid, high resolution, accurate mass (HRAM) mass spectrometry.

Methods: Tryptic digests of complex lysates and immunoaffinity enriched acetyl-lysine modified samples were separated using fast chromatography and analyzed by an Thermo Scientific™ Orbitrap Fusion™ Tribrid™ Mass Spectrometer.

Results: We identified 1,666 lysine-acetylated peptides from an enriched DH10B *E. coli* strain with a 45 minute analytical LC gradient. Additionally 3,452 unique proteins were identified from a *C. elegans* tryptic lysate from a single 55 minute analytical LC gradient run, which constitutes ~ 17% of the *C. elegans* proteome.

Introduction

Comprehensive identification of modified and unmodified proteins from complex samples is a central component of proteomic mass spectrometry experiments. In particular, the identification of lysine acetylation has been demonstrated to be critical in the regulation of transcription factors and cytoskeletal proteins. The regulated activity of protein acetylases and de-acetylases leads to a very dynamic spectrum of acetylation within the cell. Thus, it is essential to develop a rapid and sensitive method to identify site specific alterations in lysine acetylation in a global/proteomic scale. We combined rapid immunoaffinity enrichment with a very sensitive high resolution/high mass accuracy mass spectrometer utilizing a very fast liquid chromatography gradient to fully characterize proteomes and sub-proteomes in a high throughput method.

Methods

Sample Preparation

Tryptic digests from *E. coli* and *C. elegans* cell lysates were prepared as previously described[1, 2]. The acetyl enrichment of *E. coli* samples were performed using anti-acetyllysine agarose beads (PTM Biolabs Inc., Chicago IL)[3]. Samples were reduced with DTT and then alkylated with iodoacetamide (*E. coli*) and methyl methanethiosulfante (*C. elegans*)[1,2].

Liquid Chromatography

LC separation was conducted using a Thermo Scientific™ EASY-nLC™ 1000 HPLC system. Mobile phases were composed of [A] water (0.1% formic acid) and [B] acetonitrile (0.1% formic acid). Peptides were directly loaded onto a Thermo Scientific™ EASY-Spray™ PepMap™ C18 column, 75 µm ID x 25 cm, 2 µm, 100 Å. The analytical gradient for the *C. elegans* sample was from 10-35% B in 55 min. The gradient for the *E. coli* sample was from 5-22% B in 40 min followed by 22-30% B in 5 min, with a total analytical gradient of 55 and 45 minutes for the *C. elegans* and *E. coli* samples respectively. Flow rate for analytical gradients was 300 nL/min. Analytical column was heated to 40 ºC.

Mass Spectrometry

All samples were analyzed using a Orbitrap Fusion mass spectrometer. Data was acquired on the Orbitrap Fusion MS using a resolution of 120,000 (@ 200 *m/z*) for full MS scans followed by HCD fragmentation and detection of the fragment ions in the ion trap. Detailed MS parameters are provided in Table 1.

TABLE 1. Mass Spectrometry Instrument Parameters.

	C. elegans	*E. coli*
Full Scan		
Resolution (@ *m/z* 200)	120,000	120,000
Scan Range (*m/z*)	400-2000	400-2000
Max Injection (ms)	50	50
AGC Target	2.00E+05	2.00E+05
Data-dependent MS/MS		
Fragmentation	HCD	HCD
NCE (%)	35	35
Detector Type	Ion Trap	Ion Trap
AGC Target	1.00E+04	1.00E+04
Max Injection (ms)	35	100
Dynamic Exclusion (sec)	35	35

Data Analysis

Proteins/Peptides were confidently identified using Thermo Scientific™ Proteome Discoverer™ ver. 1.4 software with SEQUEST® HT. Databases for *E. coli* and *C. elegans* used were obtained from Uniprot containing 138,865 and 26,061 sequences respectively. Precursor and fragment mass tolerances were set to 10 ppm and 0.8 Da respectively. Static modifications used were carbamidomethyl (*E. coli*) and methylthio (*C. elegans*) for cysteine. Dynamic modifications for *E. coli* included oxidation (M), acetyl (K) and deamidation (NQ). Dynamic modification used for *C. elegans* was oxidation (M). Peptides were confidently matched to sequences with a false discovery rate of 1% as determined by Percolator®.

Results

Data-Dependent Top Speed Method Analysis of *C. elegans* Lysate

1.5 µg of *C. elegans* cell lysate was directly loaded onto a reverse phase EASY-Spray column using a EASY-nLC 1000 HPLC system. Peptides were separated over a 55 minute gradient. MS data was acquired on a Orbitrap Fusion mass spectrometer using a data-dependent Top Speed method with a maximum cycle time of 3 seconds. Precursors dynamicaly chosen from the full MS scan acquired in the Orbitrap were isolated by the quadrupole followed by HCD fragmentation and a MS/MS scan of the fragment ions acquired in the ion trap.

The parallelization of the Orbitrap and ion trap scans increases efficiency and maximizes the identification of eluting peptides. The enhanced speed of the Orbitrap Fusion MS allows for shorter gradients to be utilized while still characterizing a significant portion of a proteome. In this case utilizing a short 55 minute gradient yielded 20,973 unique peptides and 3,452 unique proteins (5,288 total proteins) representing 17-26% of the *C. elegans* proteome (C. elelgans proteome = 20,470 proteins[4]) at a 1% FDR.

Figure 3 demonstrates that very high sequence coverage was obtained for many *C. elegans* proteins identified , while Figure 4 and 5 indicates a normal distribution of missed cleavages and charge states for a tryptic digest.

FIGURE 1. Basepeak chromatogram of 1.5 µg of *C. elegans* lysate analyzed by the Orbitrap Fusion

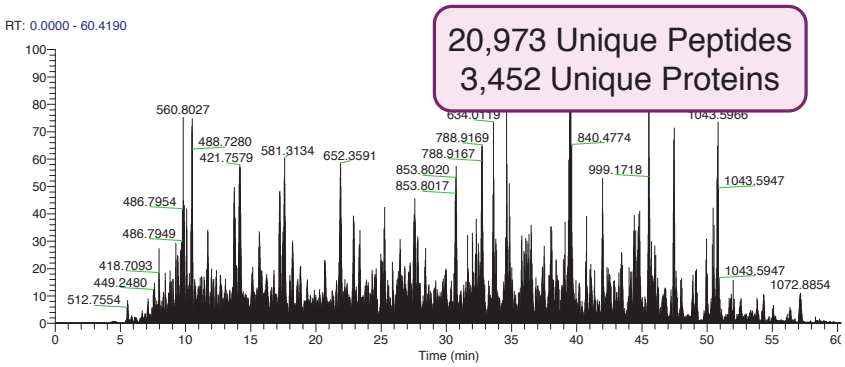


FIGURE 2. Representative duty cycle. Each stick represents one scan, the first and last are full MS scans with MS/MS scans in between. The entire duty cycle took only 1.19 seconds, during which 1 full MS scan and 24 MS/MS scans were acquired.

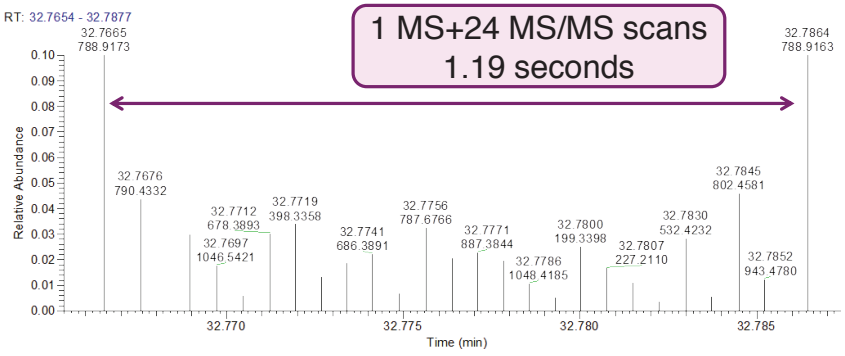


FIGURE 3. Distribution of sequence coverage for all *C. elegans* proteins identified. Color represents number of amino acids in protein; red<4000 AA, yellow>4000 AA, blue/purple>10,000 AA.

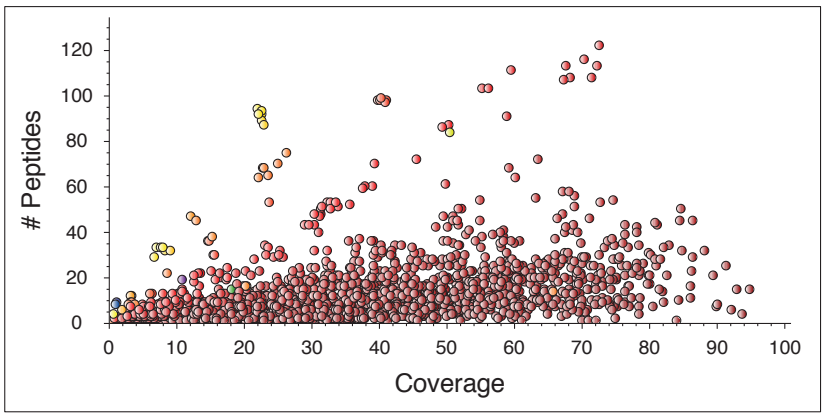


FIGURE 4. Distribution of missed cleavages detected in all C. elegans peptides detected by Proteome Discoverer/Sequest software

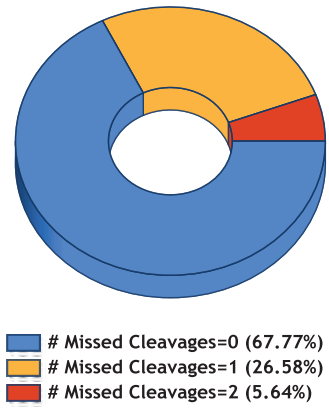
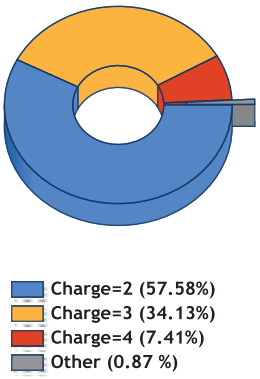


FIGURE 5. Charge state distribution of all C. elegans peptides detected by Proteome Discoverer/Sequest software.



Data-Dependent Top Speed Method Analysis of Acetyl-Lysine Enriched *E. coli*

Through the combination of rapid immunoaffinity acetyl-lysine enrichment with a Orbitrap Tribrid Mass Spectrometer we identified 1,666 acetyl-lysine modified peptides from a DH10B *E. coli* strain with a 45 min analytical LC gradient (Figure 6). Proteins/Peptides were confidently identified using Proteome Discoverer/SEQUEST software. Peptides were confidently matched to sequences with a false discovery rate of 1% as determined by Percolator. Analysis of this sub-proteome (acetylated proteins) will allow for a more complete understanding of key cellular regulatory pathways. The immunoaffinity enrichment resulted in 50% acetyl-lysine modified peptides (Figure 7).

To explore the distribution of acetyl-lysine peptides in the E. coli proteome in terms of molecular function and biological process. The identified proteins were grouped using GO ontology classifications by Thermo Scientific™ Proteome Discoverer ™ software and Thermo Scientific™ ProteinCenter™ software. The results of the acetyl-lysine enriched samples were compared to non-enriched *E. coli* tryptic lysate. For this comparison only the acetyl-lysine modified peptides were used from the immunoaffinity enriched sample, the non-acetyl lysine peptides were excluded. The comparison of molecular function is shown in Figure 8 and the results of the biological process are shown in Figure 9. The similar distribution of identified proteins for the acetyl-lysine enriched sample and the non-enriched sample indicate lysine acetylation is present throughout the *E. coli* proteome and involved in many molecular functions and biological processes.

FIGURE 6. Basepeak chromatogram of Acetyl-Lysine enriched *E. coli* sample analyzed by the Orbitrap Fusion MS.

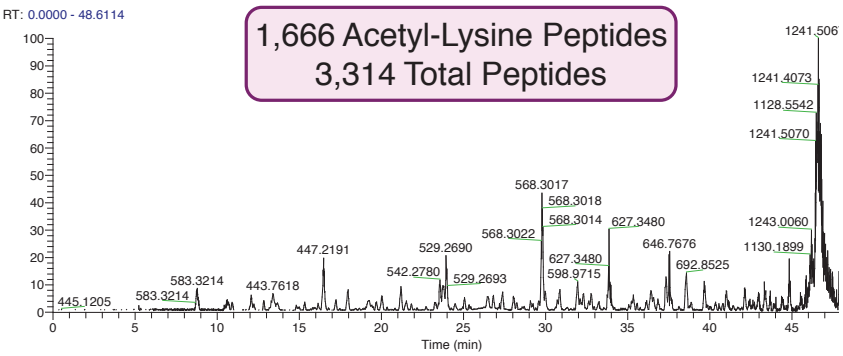


FIGURE 7. Immunoaffinity enrichment results in 50% aceytl-lysine modified peptides identified by Orbitrap Fusion analysis. Yellow circle represents all peptides identified (3,314) while the red circle represents only acetyl-lysine modified peptides (1,666). LCMS analysis of an non-enriched E. coli sample resulted in only 3 acetyl-lysine peptides identified (data not shown).

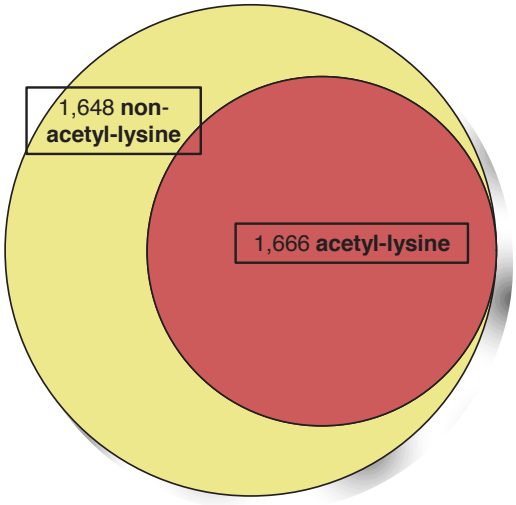


FIGURE 8. Comparison of the molecular function of acetyl-lysine peptides identified from immunoaffinity enriched E. coli sample (non acetyl-lysine peptides excluded) to non-enriched E. coli sample based on GO ontology classifications. The distribution of acetyl-lysine peptides closely matches the distribution of peptides identified from a non-enriched fraction.

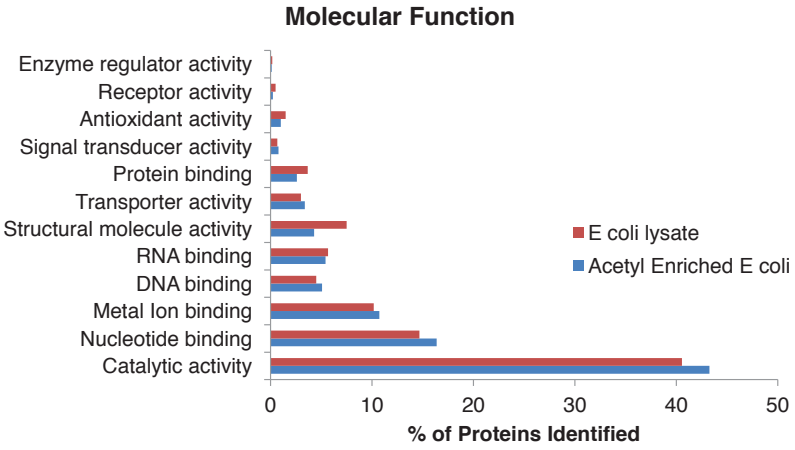
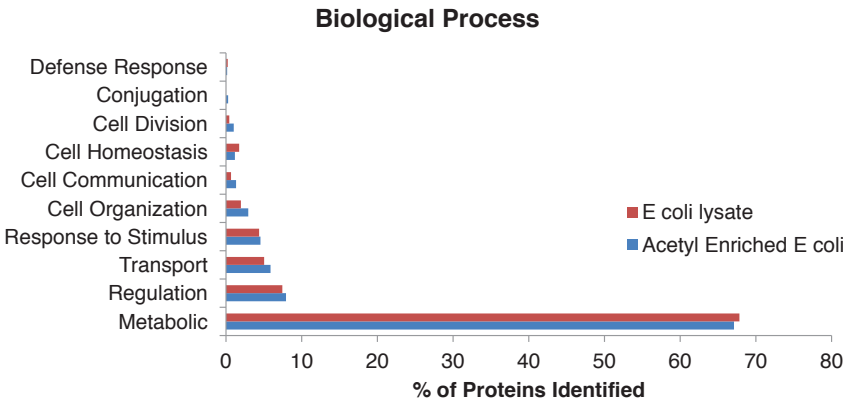


FIGURE 9. Comparison of the biological processes of acetyl-lysine peptides identified from immunoaffinity enriched E. coli sample (non acetyl-lysine peptides excluded) to non-enriched E. coli sample based on GO ontology classifications. The distribution of acetyl-lysine peptides closely matches the distribution of peptides identified from a non-enriched fraction.



Conclusion

- 3,452 unique proteins were identified from a *C. elegans* tryptic lysate from a single 55 min analytical LC gradient run, which constitutes at least 17% of the *C. elegans* proteome.
- The enhanced speed and sensitivity of the Orbitrap Fusion MS allows for shorter gradients to be utilized while characterizing a proteome.
- Through the combination of rapid immunoaffinity acetyl-lysine enrichment with a Orbitrap Fusion Mass Spectrometer we identified 1,666 acetyl-lysine modified peptides from a DH10B *E. coli* strain with a 45 min analytical LC gradient.
- The similar distribution of identified proteins for the acetyl-lysine enriched sample and the non-enriched sample indicate lysine acetylation is present throughout the *E. coli* proteome and involved in many molecular functions and biological processes.

References

1. Zhang, J. et al. Lysine Acetylation Is a Highly Abundant and Evolutionarily Conserved Modification in *Escherichia Coli*, MCP 2009.
2. Tran PV, et al. Fetal Iron Deficiency Alters Proteome of Adult Rat Hippocampal Synaptosomes, Am J Physiol Regul Integr Comp Physiol. 2013.
3. Kim, S. et al. Substrate and Functional Diversity of Lysine Acetylation Revealed by a Proteomics Survey, Mol Cell 2006
4. www.wormbase.org 10 August 2011.

www.thermofisher.com

Africa +43 1 333 50 34 0
Australia +61 3 9757 4300
Austria +43 810 282 206
Belgium +32 53 73 42 41
Canada +1 800 530 8447
China 800 810 5118 (free call domestic)
400 650 5118
Denmark +45 70 23 62 60
Europe-Other +43 1 333 50 34 0
Finland +358 9 3291 0200
France +33 1 60 92 48 00
Germany +49 6103 408 1014
India +91 22 6742 9494
Italy +39 02 950 591
Japan +81 45 453 9100
Latin America +1 561 688 8700
Middle East +43 1 333 50 34 0
Netherlands +31 76 579 55 55
New Zealand +64 9 980 6700
Norway +46 8 556 468 00
Russia/CIS +43 1 333 50 34 0
Singapore +65 6289 1190
Spain +34 914 845 965
Sweden +46 8 556 468 00
Switzerland +41 61 716 77 00
UK +44 1442 233555
USA +1 800 532 4752

Thermo SCIENTIFIC
A Thermo Fisher Scientific Brand

PN-64098-EN 0716S