Large Scale Targeted Protein Quantification Using WiSIM-DIA Workflow on a Orbitrap Fusion Tribrid Mass Spectrometer

Reiko Kiyonami,[1] Bhavin Patel,[3] Michael Senko,[1] Vlad Zabrouskov,[1] Jarrett Egertson,[2] Sonia Ting,[2] Michael MacCoss,[2] John Rogers,[3] Andreas FR Hühmer[1]
[1]Thermo Fisher Scientific, San Jose, CA; [2]University of Washington, Seattle, WA;
[3]Thermo Fisher Scientific, Rockford, IL

Thermo SCIENTIFIC

Overview

Purpose: Evaluate the quantitative performances for large scale targeted protein experiments using the unique data-independent acquisition workflow (WiSIM) which uses high resolution accurate mass of SIM scans for quantification and CID MS/MS spectra of MS/MS scans for simultaneous peptide sequence confirmation on a Thermo Scientific™ Orbitrap Fusion™ Tribrid™ mass spectrometer.

Methods: Orbitrap Fusion MS equipped with a Thermo Scientific™ EASY-Spray™ source is used for all experiments. For data-independent acquisition set up, three HRAM SIM scans (240,000 FWHM) with wide isolation windows (200 amu) were used to cover all precursor ions of 400–1000 *m/z*. In parallel with each SIM scan, 17 sequential ion trap MS/MS with 12 amu isolation windows were acquired to cover the associated 200 amu SIM mass range. Quantitative information for all precursor ions detected in three sequential SIM scans is recorded in a single run. Plus, all MS/MS fragment information over the mass range of 400–1000 *m/z* is recorded for sequence confirmation of any peptide of interest by querying specific fragment ions in the spectral library. The quantitative performances using this approach for large scale targeted protein quantification were evaluated using various samples.

Results: The data collected from SIM scans with extreme high resolving power provided high mass accuracy, which allowed unambiguous detection of targeted peptide peaks for highly reproducible and accurate quantitative results over thousands of peptides by applying targeted data extraction strategy after the data independent acquisition.

Introduction

Data independent acquisition (DIA) is an emerging high-throughput quantitative technique for quantifying large numbers of proteins in biological samples. Last year, we reported LODs as low as 10 attomole and over 4 orders of linear dynamic range for targeted peptides on an Orbitrap Fusion Tribrid mass spectrometer using a new WiSIM-DIA workflow[1]. WiSIM-DIA uses precursor ions from HR/AM SIM scans collected with wide isolation windows for quantification and ion trap CID MS/MS spectra collected in parallel with the SIM acquisition for confirmation[1]. Here we report that precise and reproducible quantitative results are also achieved while applying this unique WiSIM-DIA workflow to large scale quantitative studies in complex matrices.

Methods

Sample Preparation

Sample set 1: Three *E.coli* digest samples (250 ng/µL, 500 ng/µL, and 1000 ng/µL) were prepared using a commercially available *E.coli* digest (Waters).

Sample set 2: EGFR, AKT isoforms and PTEN targets were immunoprecipitated from A431 lysate by using biotinylated antibodies and Thermo Scientific™ Pierce™ Streptavidin Coated Magnetic Beads. IP enriched samples and A431 lysate were processed by in-solution digestion method. Prior to MS analysis, tryptic digest samples were desalted using the Thermo Scientific™ Pierce™ C18 Spin Tips.

Liquid Chromatography

A Thermo Scientific™ EASY-nLC™ 1000 nanoflow LC was used for all experiments.

Column:	EASY-Spray column; Thermo Scientific™ Acclaim™ column (PepMap™ C18, 2 µm, 75 µm × 50 cm)
Flow Rate:	300 nL/min; Buffer A: 0.1% FA/H2O; Buffer B: 0.1% FA/CAN
Gradient:	5% B to 25% B in 100 min; 25% B to 35% B in 20 min
Sample Loading:	Directly loaded on column
Injection Amount:	1µL

Mass Spectrometry

Orbitrap Fusion MS equipped with an Easy-Spray source is used for all experiments.

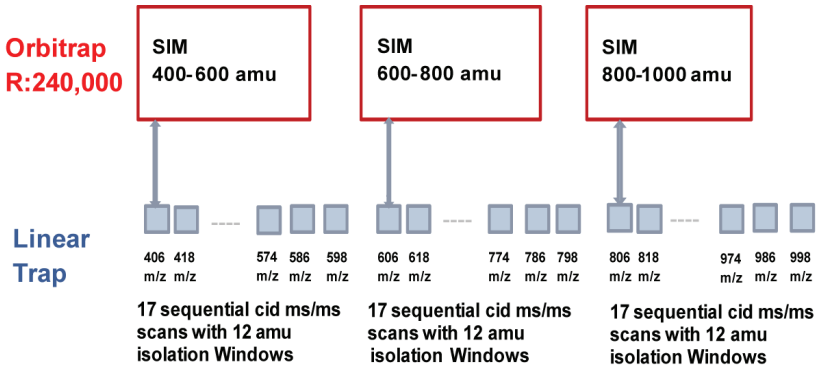
Capillary temperature:	275° C; Spray voltage: 1800 V
FT SIM: Resolution:	240,000; AGC target: 4+E04; Isolation width: 200 amu
CID MS/MS:	Rapid CID MS/MS; AGC target: 5+E04; Isolation width: 12 amu using the quadrupole for isolation

Six scan events: Three SIM scan events (scan 1, 3, and 5). Each SIM scan experiment is followed by one tMSn scan experiment which carries out 17 consecutive CID MS/MS events using predefined precursor ion inclusion list (Figure 1).

Data Analysis

Thermo Scientific™ Pinpoint™ 1.4 software is used for targeted qualitative and quantitative data extraction post data acquisition.

FIGURE 1. WiSIM-DIA workflow to collect HR/AM SIM and CID MS/MS in parallel.

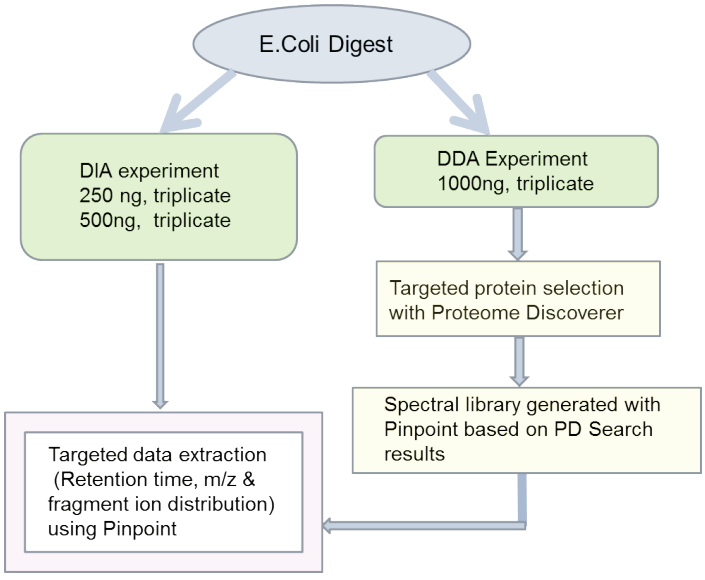


Results

Large-Scale Relative Quantification of *E.coli* Proteins
Spectral Library Generation and Targeted *E.coli* Protein Selection

For generating spectral library and identifying *E.coli* proteins list of interest, the 1000 ng *E.coli* digest was analyzed in triplicate using a standard shotgun DDA experiment first. Thermo Scientific™ Proteome Discoverer™ software version 1.4 was used for the database search of the DDA raw files against an *E.coli* database. The SEQUEST® HT search engine was used. A target decoy peptide spectral match (PSM) validator (0.01 FDR strict – 0.05 FDR relaxed) was used for PSM validation. The search results of the triplicate runs were combined and 1,100 unique *E.coli* proteins were identified, which included at least two identified peptides with high confidence, #1 peptide rank, and minimal cross-correlation scores (2.0 for charge 2, 2.25 for charge 3, and 2.5 for charge 4). These peptides were selected as the quantitative targets for targeted extraction in the large-scale quantitative comparison of *E.coli* digests using the WiSIM-DIA workflow. The search results were read in by Pinpoint software directly for generating the spectral library (Figure 2).

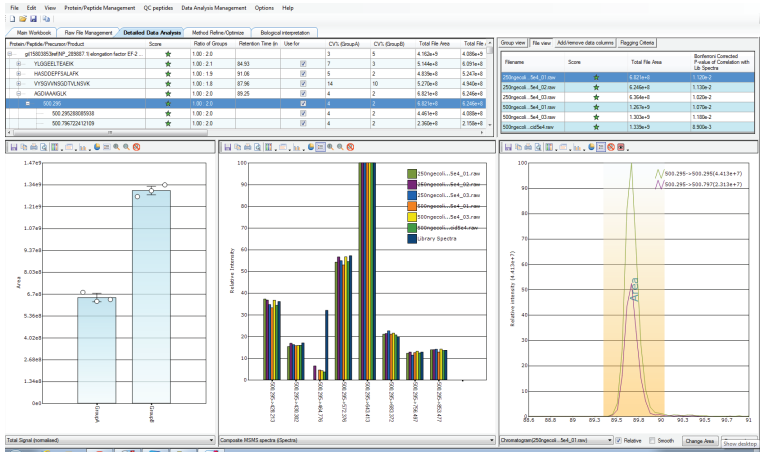
FIGURE 2. Experiment workflow for the large scale *E.coli* protein using WiSIM-DIA workflow.



Data Processing

Pinpoint software was used for all of the targeted data extraction. The Proteome Discoverer software search results were imported into Pinpoint software for generation of a sample-specific spectral library. Only the MS/MS spectra that passed the minimal cross-correlation scores were imported into the spectral library. All peptides in the spectral library were then used for the targeted data extraction of the WiSIM-DIA data. The XICs of the ^{12}C and ^{13}C isotopes for the precursor ions of each targeted peptide were used for quantification with a ±5 ppm window. The eight most-intense fragment ions (b and y types larger than 200 *m/z*) detected from the discovery data were used for peptide sequence confirmation through spectral library match within the 12 *m/z* CID MS/MS. A *p*-value (probability of random spectral matching) was calculated using standard statistical methods to compare the multiple product ion distribution extracted from the DIA MS/MS data with MS/MS data of a peptide stored in a spectra library[2]. A peptide with a *p*-value of less than 0.1 was considered to be identified with high confidence by the library match.

FIGURE 3. Simultaneous Qual/Quan through targeted data extraction.



Quantitative Summary of Targeted 1,100 *E.coli* Proteins

The data collected from 250 ng *E.coli* and 500 *E.coli* samples using WiSIM-DIA method were processed using Pinpoint software. 5,923 targeted *E.coli* peptides were identified with *p*-values less than 0.1 and %CVs ≤25%, and were used to determine the relative expression ratios between the two samples. These 5,923 identified peptides represent 1,090 *E.coli* proteins yielding a 98% success rate for quantifying a total of 1,100 targeted proteins (Figure 4). Over 97% of quantified proteins gave exceptional quantitative accuracy for the detection of the two-fold expression change expected between the two samples (Figure 4). In addition, 85% of the 5,923 quantified peptides gave %CVs less than 15% (Figure 5).

FIGURE 4. Quantification summary for quantifying 1,100 *E.coli* proteins.

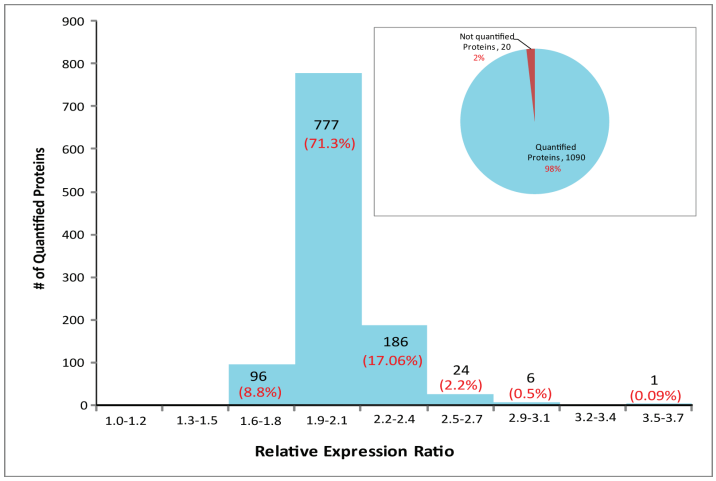
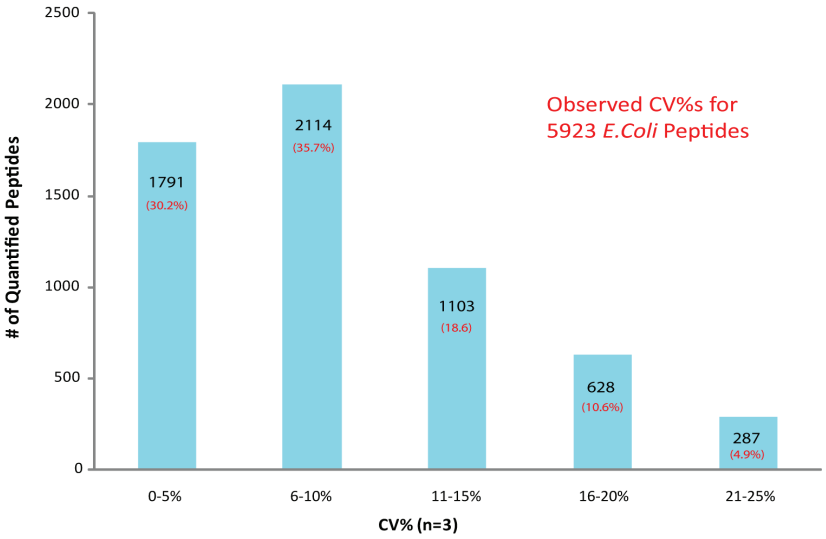


FIGURE 5. Coefficient of variation of 5,923 quantified *E.coli* peptides.



Relative Quantification of IP-enriched and Untreated A431 Lysate Samples

The WiSIM-DIA workflow was further applied to the cell lysate samples for evaluating immunocapture efficiency of EGFR, AKT2, AKT1, and PTEN from the A431 lysate. The EGFR, AKT isoforms, and PTEN were enriched by immunoprecipitation from A431 lysate with biotinylated antibodies and Streptavidin Coated Magnetic Beads[3].

The data collected from the untreated A431 lysate and IP-enriched A431 lysate samples using WiSIM-DIA method were processed using Pinpoint software. The spectral library was generated relying on the PD search results from the shotgun data dependent experiments of both the entire and IP-enriched A431 lysate samples. All identified peptides of EGFR, AKT2 , AKT1, and PTEN from the discovery experiments were imported into Pinpoint software and targeted extracted for simultaneous peptide confirmation and quantification. The Immunoprecipitation using biotinylated antibodies and Streptavidin Coated Magnetic Beads resulted in overall higher yield of EGFR, AKT2, and AKT1 target proteins from the complex A431 lysate matrices (Figure 6). Table 1 shows partial relative quantification results for targeted proteins.

FIGURE 6. XICs of AKT2 peptide SDGSFIGYK from A431 lysate samples.

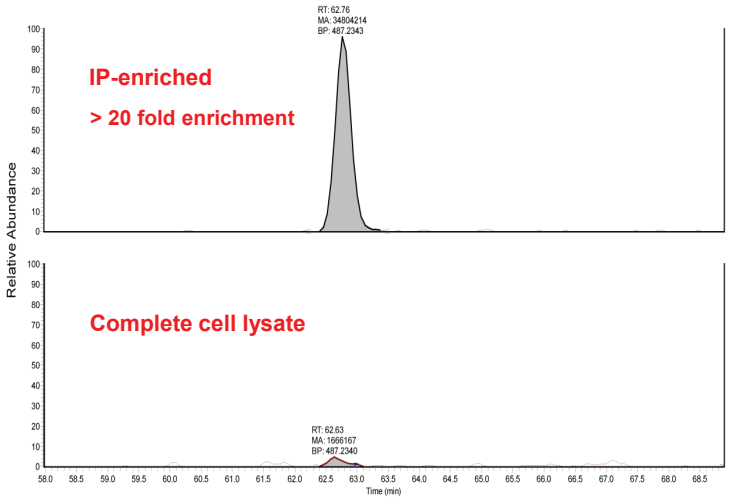


TABLE 1. Relative Quantification Summary for EGFR and AKT Proteins.

Proteins	Fold of enrichment IP-Enriched /Unenriched	Retention Time	Total File Area IP_A431	Total File Area A431
>sp|P00533|EGFR_HUMAN Epidermal growth facto	16.09		IP-Enriched	Unenriched
WM[Oxid]ALESILHR	5	91.6	5.65E+06	1.12E+06
ELIIEFSK	2.17	93.1	3.28E+07	1.50E+07
VLVIQGEDR	8.6	63.58	1.60E+08	1.86E+07
NYDLSFLK	16.7	100.6	6.14E+07	3.88E+06
GM[Oxid]NYLEDR	50	46.5	2.55E+07	4.63E+05
CNLLEGEPR	6.25	60.3	2.23E+07	3.50E+06
VAPQSSEFIGA	8.33	83.8	1.27E+08	1.51E+07
GPDNCIQCAHYIDGPHCVK	6.25	59.2	1.62E+07	2.56E+06
GSHQISLDNPDYQQDFFPK	5.26	97	4.41E+06	1.51E+07
VCNGIGIGEFK	10	80.1	3.12E+07	3.04E+06
VLGSGAFGTVYK	3.03	81.5	1.33E+08	4.34E+07
RPAGSVQNPVYHNQPLNPAPSR	6.25	53.7	1.86E+07	2.90E+06
IPLENLQIIR	5.88	103.8	5.95E+07	1.04E+07
NGLQSCPIKEDSFLQR	2.94	78.2	2.58E+07	8.76E+06
FSNNPALCNVESIQWR	50	105.8	5.33E+05	1.08E+04
EISDGDVIISGNK	20	70.4	8.80E+07	4.47E+06
ELVEPLTPSGEAPNQALLR	5.88	104.9	1.27E+07	4.93E+06
NYVVTDHGSCVR	16.67	39.3	1.37E+07	8.81E+05
CDPSCPNGSCWGAGEENCQK	10	60	2.10E+06	2.06E+05
MHLPSPTDSNFYR	4.76	82.4	2.50E+06	5.37E+05
ACGADSYEM[Oxid]EEDGVR	20	53.2	2.90E+06	1.35E+05
YSFGATCVK	72.98	61.22	7.59E+07	1.04E+06
ACGADSYEMEEDGVRK	33.3	50.1	6.88E+05	2.35E+04
>sp|P42566|EPS15_HUMAN Epidermal growth fact	8.175		IP-Enriched	Unenriched
LPVDILGR	4.54	91.4	4.21E+06	9.43E+05
VNNEDPFR	12.5	51.3	5.44E+06	4.55E+05
SATSSSVSNVVITK	11.11	57.4	5.92E+06	5.15E+05
LQQETAELEESVESGK	4.55	77.5	1.11E+06	2.48E+05
>sp|Q96GX5|MASTL_HUMAN Microtubule-associat	55.3		IP-Enriched	Unenriched
SANAIETK	14.51	16.64	1.09E+06	7.51E+04
SFNSHINASNNSEPSR	168	38.35	3.60E+06	2.14E+04
CLTSNLLQSR	2.5	72.9	1.88E+06	7.53E+05
DYLSSSFLCSDDDR	36.3	109.4	4.14E+05	1.14E+04
>sp|P31751|AKT2_HUMAN RAC-beta serine/threon	36.52		IP-Enriched	Unenriched
AIQM[Oxid]VANSLK	9.54	52.17	8.72E+06	9.14E+05
SLLAGLLK	n/a	102.9	1.14E+07	n.d.
DEVAHTVTESR	7.35	31.71	2.05E+06	2.79E+05
VTMNDFDYLK	92.67	98.72	2.15E+06	2.32E+04
>sp|P31749|AKT1_HUMAN RAC-alpha serine/threo	60.19		IP-Enriched	Unenriched
LGGGSEDAK	120.59	12.49	8.61E+05	7.14E+03
SDGSFIGYK	16.86	62.9	1.77E+07	1.05E+06
KQEEEEM[Oxid]DFR	n/a	28.53	4.14E+06	n.d.
NDGTFIGYK	95.5	63.66	2.34E+07	2.45E+05
EAPLNNFSVAQCQLMK	7.81	101.69	6.03E+05	7.72E+04

Conclusion

- The unique WiSIM data-independent acquisition workflow that collects HRAM wide-window SIM data and rapid CID MS/MS data in parallel on the new Orbitrap Fusion Tribrid MS enables reproducible and accurate quantitative results for large scale quantitative experiments.
- By quantifying using high-resolution (240,000) data that provides higher mass accuracy for separating precursor ions from background interferences, low limits of detection and high selectivity were achieved. Among 1,100 targeted *E.coli* proteins, over 97% of quantified proteins gave exceptional quantitative accuracy for the detection of the two-fold expression change expected between the two samples. In addition, 85% of the 5,923 quantified peptides gave %CVs less than 15%.
- WiSIM-DIA workflow was successfully applied for evaluating the immunocapture efficiency of Streptavidin Coated Magnetic Beads for enrichment of EGFR, AKT2, and AKT1 from A431 lysate matrices.

References

1. Kiyonami, R.; Senko, M.; Zabrouskov, V.; Egertson, J.; Ting, S.; MacCoss, M. J.; Hühmer, A. F. Thermo Scientific Application Note AN64026.
2. Prakash, A.; Tomazela, D. M.; Frewen, B.; Maclean, B.; Merrihew, G.; Peterman, S. M.; Maccoss, M. J. Expediting the development of targeted SRM assays: using data from shotgun proteomices to automate method development. *J. Proteome Res.* **2009**, *8*, 2733–2739.
3. Patel, B.; Meier, S.; Opperman, K.; Haney, P.; Kaboord, B.; Rogers, J. Enrichment of EGFR/PI3K/AKT/PTEN Proteins using Immunoprecipitation and Analysis with Mass Spectrometry-based Proteomics. ASMS 2014, Poster 514.

www.thermoscientific.com

Thermo Fisher Scientific, San Jose, CA USA is ISO 9001:2008 Certified.

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-64085-EN-0614S