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

Reiko Kiyonami, Michael Senko, Vlad Zabrouskov, Andreas F.R. Hähner, Thermo Fisher Scientific, San Jose, CA, USA, Jarrett Egerton, Sonia Ting, Michael MacCoss, University of Washington, Seattle, WA, USA

Key Words
DIA (data-independent acquisition), targeted protein extraction post data acquisition, Pinpoint software, HRAM SIM, rapid CID/MS/MS, simultaneous quantification, SIM (single-ion monitoring)

Goal
Develop a highly sensitive and highly selective data-independent acquisition (DIA) workflow for complete, reproducible, large-scale targeted protein quantification on the new Thermo Scientific™ Orbitrap Fusion™ Tribrid™ mass spectrometer.

Introduction
Proteomics studies are rapidly turning from qualitative to largely quantitative experiments in order to better understand the functions and interactions of proteins in biological systems and to verify long lists of putative biomarkers. The extreme complexity and large dynamic range of proteins in standard biological samples challenge traditional data-dependent workflows. They necessitate very fast tandem mass spectrometry (MS/MS) with simultaneous high sensitivity and reproducibility. Without these, it is difficult to achieve the depth of sample interrogation necessary to identify and quantify the same targets in replicate runs to accurately determine differences amongst samples.

Recently, several data-independent acquisition approaches have been explored to increase reproducibility and comprehensiveness for better quantification. One such approach acquires only high-resolution, accurate-mass (HRAM) MS/MS data (no full scans) generated with very wide isolation windows (~25 m/z). This data is then interrogated for quantification using extracted ion chromatograms of targeted fragment ions only. The quantitative performance from using only HRAM MS/MS data generated with wide isolation windows is potentially compromised by interfering product ions from co-eluted background compounds. For highly complex biological samples, a portion of the targeted fragment ions are potentially contaminated by fragment ions from nontargeted peptides that co-elute in the same retention time window and are co-isolated within a 25 m/z wide window. These interferences can result in higher limits of detection (LOD)/limits of quantitation (LOQ), less accurate quantitative results, and narrower dynamic range. The Orbitrap Fusion Tribrid MS is based on a mass resolving quadrupole, Orbitrap mass analyzer, collision cell, linear ion trap mass analyzer (Q-OT-qIT) architecture (Figures 1 and 2). The Orbitrap detector can collect data with a very high resolving power of 240,000 at a scan rate of 1.5 Hz. The linear ion trap can collect more than 20 high-quality collision-induced dissociation (CID) MS/MS spectra in 1 sec. The unique architecture of the Orbitrap Fusion MS enables parallel Orbitrap selected-ion monitoring (SIM) scanning with rapid and sensitive targeted ion trap MS/MS detection to generate data with the resolution and sensitivity required for accurate data-independent analyses.
A new DIA workflow was developed that takes advantage of the unique architecture of the Orbitrap Fusion Tribrid MS. It simultaneously collects HRAM SIM spectra with wide isolation windows in the Orbitrap detector and CID MS/MS spectra in the linear ion trap detector. This method is referred to as wide selected-ion monitoring, data-independent acquisition, or WiSIM-DIA. The WiSIM-DIA method consists of three SIM scans acquired with extremely high resolution and mass accuracy, and with wide isolation windows (240,000 resolution FWHM at m/z 200, and <3 ppm, respectively). It covers all precursor ions between m/z 400 and 1000 while increasing the sensitivity relative to standard full-scan analysis. The use of a wide SIM window for the collection of ions in the first step of this novel method leads to the selective enrichment of analytes in the selected SIM mass range and enables the detection and quantification of low abundance peptides in the final targeted extraction step. In parallel with each SIM spectrum, 17 sequential ion trap MS/MS spectra with 12 m/z isolation windows are acquired to detect fragment ions across each associated 200 m/z SIM mass range (Figure 3). Quantification of targeted proteins after data acquisition is carried out using extracted ion chromatograms of HRAM SIM data with a ±5 ppm window. Simultaneous peptide sequence confirmations are carried out using CID MS/MS relying on a spectral library.

The main advantages of the WiSIM-DIA workflow, relative to the wide-window MS/MS-only DIA workflows, are that there is enough time to acquire HRAM precursor data for accurate and sensitive detection and quantification, while the fast and sensitive MS/MS can be used for reliable confirmation. As a result, the quantitative performance of this method does not suffer from the drawbacks of wide-window MS/MS-only DIA workflows. All quantitative analyses are carried out using HRAM SIM data. The CID MS/MS data are used only for peptide confirmation. Therefore, the fragment ion contamination caused from nontargeted co-eluted peptides does not impact the quantitative results. This novel WiSIM-DIA approach increases selectivity, sensitivity, and reproducibility, ultimately improving the accuracy and throughput of the quantification experiments by using precursor ions collected with very high (240,000) resolving power. The fragment ions detected in the WiSIM-DIA approach, which are used only for confirmation of identity, are from narrower precursor populations (12 m/z), and are detected by a more sensitive mass analyzer. Therefore, this new WiSIM-DIA approach can quantify all detected peptide precursor peaks with high sensitivity and selectivity, yielding a complete and reproducible quantitative data set in a single sample injection. In addition, all MS/MS product ion information is recorded for sequence confirmation of any peptide of interest within the mass range of m/z 400–1000 by subsequent confirmatory matching of specific product ions of an easy-to-generate, sample-specific spectral library. The details of this unique WiSIM-DIA workflow, its quantitative performances, and easy, automatic data processing with Thermo Scientific™ Pinpoint™ software, are reported here.
Experimental

Sample Preparation
Sample 1 (a dilution series for evaluating detection limits and linear dynamic range): A mixture of 14 isotopically labeled synthetic peptides was spiked into a complex E. coli digest (500 ng/μL) at five different concentrations (0.01 fmol/μL, 0.1 fmol/μL, 1 fmol/μL, 10 fmol/μL, and 100 fmol/μL) to generate a dilution series.

Sample 2 (for evaluating detection and quantification of low to high abundance proteins in a complex mixture): A digest of six proteins with concentrations covering five orders of magnitude (0.01 fmol/μL, 0.1 fmol/μL, 1 fmol/μL, 10 fmol/μL, 100 fmol/μL, and 1000 fmol/μL) was spiked into an E. coli digest (500 ng/μL).

Sample 3 (a sample set for large-scale relative quantification comparison of separate complex mixtures): An E. coli digest was analyzed at 250 ng and 500 ng on column.

Nano-LC

<table>
<thead>
<tr>
<th>System</th>
<th>Thermo Scientific EASY-nLC 1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column</td>
<td>Thermo Scientific EASY-Spray PepMap C18 column (75 μm x 50 cm, 2 μm particle size)</td>
</tr>
<tr>
<td>Flow rate</td>
<td>300 nL/min</td>
</tr>
<tr>
<td>Buffer A</td>
<td>0.1% formic acid/water</td>
</tr>
<tr>
<td>Buffer B</td>
<td>0.1% formic acid/acetonitrile</td>
</tr>
<tr>
<td>Gradient</td>
<td>2% B in 5 min, 2% B to 20% B in 100 min, 20% B to 32% B in 20 min</td>
</tr>
<tr>
<td>Sample loading</td>
<td>Directly loaded on column</td>
</tr>
<tr>
<td>Injection amount</td>
<td>1 μL</td>
</tr>
</tbody>
</table>

MS

An Orbitrap Fusion MS equipped with a Thermo Scientific EASY-Spray source was used for all experiments.

<table>
<thead>
<tr>
<th>Capillary temperature</th>
<th>275 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spray voltage</td>
<td>1600 V</td>
</tr>
<tr>
<td>S-Lens RF level</td>
<td>60</td>
</tr>
</tbody>
</table>

FT SIM

<table>
<thead>
<tr>
<th>Resolution</th>
<th>240,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGC target</td>
<td>3 x 10⁴</td>
</tr>
<tr>
<td>Isolation width</td>
<td>200 m/z</td>
</tr>
<tr>
<td>Maximum ion injection time</td>
<td>50 ms</td>
</tr>
</tbody>
</table>

CID MS/MS

| Rapid CID MS/MS, AGC target | 5 x 10⁴ |
| Isolation width             | 12 m/z  |
| Maximum ion injection time  | 47 ms   |
| Collision energy (%)        | 30      |

Data Processing

Pinpoint software version 1.3 was used for targeted qualitative and quantitative data extraction after data acquisition. A spectral library containing precursor ion and CID MS/MS information was established using previous discovery data collected on an Orbitrap Fusion MS. The extracted-ion chromatograms (XICs) of isotope ¹³C and ¹²C precursor ions per targeted peptide were used for quantification with a ±5 ppm window. The eight most intense fragment ions (b and y types) detected from discovery data for each peptide were used for confirmation through spectral library matching.

Results and Discussion

Decreased Detection Limits with SIM Acquisitions Relative to Full MS Acquisitions

The newly developed WiSIM-DIA workflow provided supreme quantitative results with a label-free approach. Standard label-free experiments rely on full-scan data for quantification of the precursors; this often compromises the detection of low-intensity ions during the elution of more intense species. The WiSIM-DIA approach was developed to alleviate this issue through the use of sequential SIM acquisitions for quantification. Unlike a full-mass-range acquisition, the SIM acquisition with a 200 m/z isolation window effectively “enriches” all ions in that window while excluding all other ions outside the mass range of interest. As a result, the SIM acquisition, even with a 200 m/z isolation window, can provide much higher sensitivity for detecting low-abundance peptides compared to full-mass-range acquisitions. An example of this is shown in Figure 4 where simply using a SIM acquisition allows for a nearly 5-fold increase in signal-to-noise ratio. When applicable, the expected retention time per targeted peptide was also considered when extracting the quantitative information.
CONFIDENT PEPTIDE SEQUENCE VERIFICATION USING ION TRAP CID MS/MS

For peptide sequence confirmation, the data-independent CID MS/MS acquisitions with a 12 m/z window over each m/z 200 SIM window were used. Unlike data-dependent analysis that requires triggering of a specific precursor for fragmentation and detection, leading to a lack of reproducibility between replicates, data-independent MS/MS acquisitions are scheduled and cover the complete mass range in every analysis. Eight of the most intense fragment ions (b and y types) identified in discovery (data-dependent) experiments were targeted in the ion trap CID MS/MS data using a 10 ppm extraction window. They were then used for sequence verification through spectral library matching. 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. A smaller p-value represented a better match between the observed data and the spectral library. A peptide with a p-value of less than 0.1 was considered to be identified with high confidence by the library match. The ion trap CID MS/MS provided high sensitivity and good spectral quality without prior knowledge of precursor ion charge state, unlike collision cell fragmentation. It was able to confirm the targeted peptide sequences with sufficient selectivity through the spectral library matches (Figure 5).

EVALUATION OF DETECTION LIMITS AND LINEAR DYNAMIC RANGE OF THE WISIM-DIA WORKFLOW IN A COMPLEX MIXTURE

The detection limits and the quantitative dynamic range of the unique WISIM-DIA workflow were evaluated using a dilution series of a mixture of 14 isotopically labeled synthesized peptides spiked into E. coli digests (500 ng/μL) at five different concentrations (0.01 fmol/μL, 0.1 fmol/μL, 1 fmol/μL, 10 fmol/μL, and 100 fmol/μL). In addition to the decreased detection limit as a result of using SIM acquisitions, the very high resolving power of 240,000 enabled unambiguous detection of targeted peptide peaks from matrix interferences, even at the lowest 10 attomole concentration level. As a result, LODs down to 10 attomole were achieved. Table 1 shows the observed coefficient of variation (CV) of each spiked peptide from triplicate runs at each concentration level. All spiked peptides were detected at the lowest concentration level tested (10 attomole). The CVs of all spiked peptides at 100 attomole concentration were less than 15% with the exception of one peptide (22%) (Table 1). With the low LOD and high selectivity of the DIA workflow, each spiked peptide was able to be detected over four orders of linear dynamic range (Figure 6).
**Table 1. Coefficients of variation for the 14 isotopically labeled peptides spiked into a complex matrix over 5 orders of magnitude**

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Precursor Ion (m/z)</th>
<th>CV% (n=3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01 fmol</td>
<td>0.1 fmol</td>
</tr>
<tr>
<td>SSAAPPPPP[HeavyR]</td>
<td>493.768</td>
<td>31</td>
</tr>
<tr>
<td>HLVTS/GEX[HeavyK]</td>
<td>496.237</td>
<td>36</td>
</tr>
<tr>
<td>IGDYAGIK[HeavyK]</td>
<td>422.736</td>
<td>24</td>
</tr>
<tr>
<td>TASEFDCAAQDK[HeavyK]</td>
<td>695.832</td>
<td>9</td>
</tr>
<tr>
<td>SAAGAFPELSR[HeavyR]</td>
<td>588.800</td>
<td>48</td>
</tr>
<tr>
<td>ELGQGVDTYLR[HeavyK]</td>
<td>773.866</td>
<td>44</td>
</tr>
<tr>
<td>QVLVGGYGR[HeavyK]</td>
<td>558.328</td>
<td>40</td>
</tr>
<tr>
<td>GLVFGSGVSGGEEGR[HeavyR]</td>
<td>801.411</td>
<td>17</td>
</tr>
<tr>
<td>SFANQPLEVYVSK[HeavyK]</td>
<td>745.392</td>
<td>21</td>
</tr>
<tr>
<td>LTILELR[HeavyR]</td>
<td>458.802</td>
<td>12</td>
</tr>
<tr>
<td>GISNEGQNASI[HeavyK]</td>
<td>513.317</td>
<td>7</td>
</tr>
<tr>
<td>NQFILSFGPR[HeavyR]</td>
<td>573.302</td>
<td>49</td>
</tr>
<tr>
<td>ELASLSFVGFK[HeavyK]</td>
<td>660.373</td>
<td>36</td>
</tr>
<tr>
<td>LSEAPAFQDFDLK[HeavyK]</td>
<td>787.421</td>
<td>63</td>
</tr>
</tbody>
</table>

**Figure 6.** WSIM-DIA workflow on the Orbitrap Fusion MS provides 10 attomole LOD with four orders of magnitude linear dynamic range.

**Detection and Quantification of Low- to High-Abundance Proteins in a Complex Mixture in a Single Run with the WSIM-DIA Workflow**

To further evaluate the ability of the new WSIM-DIA workflow to quantify targeted peptides over a wide dynamic range in a single run, six bovine protein digests were spiked into a 500 ng E. coli matrix at an abundance level that spanned over five orders of dynamic range (low-abundance to high-abundance proteins). Among the six spiked proteins, bovine serum albumin (BSA) was spiked at the lowest abundance level (10 attomoles on column) (Table 2). The sample was run in triplicate using the WSIM-DIA workflow. The HRAM SIM acquisitions reproducibly provided low limits of detection, high selectivity, and wide dynamic range for all six spiked proteins over five orders of magnitude of dynamic range. Table 2 summarizes the detected peptides per spiked protein and observed %CVs. Ninety percent of quantified peptides gave %CVs less than 10%.
Table 2. Detection and quantification of six low to high abundance bovine proteins in a complex mixture over five orders of dynamic concentrations in a single experiment

<table>
<thead>
<tr>
<th>Retention Time (min)</th>
<th>CV% (n=3)</th>
<th>Run 1</th>
<th>Run 2</th>
<th>Run 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;p&gt;207659&lt;/p&gt;</td>
<td>&lt;p&gt;ALBU_BOVIN&lt;/p&gt;</td>
<td>Serum albumin (100 ng on column)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLVDERONILK</td>
<td>51.06</td>
<td>22</td>
<td>6.41E+05</td>
<td>1.32E+06</td>
</tr>
<tr>
<td>&lt;p&gt;202652&lt;/p&gt;</td>
<td>&lt;p&gt;CASA1&lt;/p&gt;</td>
<td>&lt;p&gt;BVIN Alpha-S1-casein (100 ng on column)</td>
<td>FFAPFPEVGK</td>
<td></td>
</tr>
<tr>
<td></td>
<td>110.66</td>
<td>8</td>
<td>2.32E+06</td>
<td>2.39E+06</td>
</tr>
<tr>
<td>&lt;p&gt;200509&lt;/p&gt;</td>
<td>&lt;p&gt;HE3&lt;/p&gt;</td>
<td>&lt;p&gt;vGtin glutamate dehydrogenase (1 pmol on column)</td>
<td>LQHGTILGFPPK</td>
<td></td>
</tr>
<tr>
<td></td>
<td>56.78</td>
<td>21</td>
<td>5.47E+06</td>
<td>5.55E+06</td>
</tr>
<tr>
<td>GVAEDVKEK</td>
<td>21.16</td>
<td>11</td>
<td>1.02E+07</td>
<td>8.96E+06</td>
</tr>
<tr>
<td>MVGFFDOR</td>
<td>88.94</td>
<td>7</td>
<td>6.30E+06</td>
<td>3.07E+06</td>
</tr>
<tr>
<td>RDDSWEVEVGYR</td>
<td>67.64</td>
<td>16</td>
<td>1.19E+06</td>
<td>1.10E+06</td>
</tr>
<tr>
<td>HGGTPIVPTAEPQDQR</td>
<td>68.8</td>
<td>1</td>
<td>3.38E+06</td>
<td>3.43E+06</td>
</tr>
<tr>
<td>C[Terboxyethyl]AVVGVFPGGK</td>
<td>73.73</td>
<td>14</td>
<td>1.63E+07</td>
<td>1.66E+07</td>
</tr>
<tr>
<td>YSTDVSVDEVK</td>
<td>46.35</td>
<td>10</td>
<td>2.11E+07</td>
<td>1.70E+07</td>
</tr>
<tr>
<td>&lt;p&gt;200921&lt;/p&gt;</td>
<td>&lt;p&gt;CAH2&lt;/p&gt;</td>
<td>&lt;p&gt;BOVIN Carbonic anhydrase (100 ng on column)</td>
<td>VLDALSDK</td>
<td></td>
</tr>
<tr>
<td></td>
<td>65.52</td>
<td>4</td>
<td>2.39E+06</td>
<td>2.34E+06</td>
</tr>
<tr>
<td>VGPNAPALOK</td>
<td>23.66</td>
<td>16</td>
<td>1.89E+06</td>
<td>1.66E+06</td>
</tr>
<tr>
<td>EPPSVVISQMLK</td>
<td>55.15</td>
<td>9</td>
<td>1.07E+07</td>
<td>8.92E+06</td>
</tr>
<tr>
<td>AVQDQPKVLAVGTEASR</td>
<td>91.67</td>
<td>1</td>
<td>3.46E+07</td>
<td>3.54E+07</td>
</tr>
<tr>
<td>YGDFGTAAGQPDGLAGVGFH</td>
<td>115.83</td>
<td>10</td>
<td>3.69E+06</td>
<td>4.54E+06</td>
</tr>
<tr>
<td>&lt;p&gt;200925&lt;/p&gt;</td>
<td>&lt;p&gt;PERL&lt;/p&gt;</td>
<td>&lt;p&gt;BOVIN Lactopeptidase (100 pmol on column)</td>
<td>SPALGAARN</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20.17</td>
<td>3</td>
<td>1.05E+09</td>
<td>9.81E+08</td>
</tr>
<tr>
<td>LFQPTHK</td>
<td>19.72</td>
<td>11</td>
<td>5.30E+06</td>
<td>4.64E+06</td>
</tr>
<tr>
<td>DGGIDPLVR</td>
<td>58.98</td>
<td>5</td>
<td>1.18E+06</td>
<td>1.08E+06</td>
</tr>
<tr>
<td>IHPDFLAANIQR</td>
<td>79.44</td>
<td>1</td>
<td>1.14E+06</td>
<td>1.13E+06</td>
</tr>
<tr>
<td>GFC[C]Terboxyethyl]GLGSPPK</td>
<td>49.46</td>
<td>9</td>
<td>1.57E+09</td>
<td>1.32E+09</td>
</tr>
<tr>
<td>RPSLGAARN</td>
<td>14.27</td>
<td>13</td>
<td>5.69E+06</td>
<td>5.69E+06</td>
</tr>
<tr>
<td>IGVYELDDGDIQN</td>
<td>76.67</td>
<td>1</td>
<td>3.49E+09</td>
<td>3.49E+09</td>
</tr>
<tr>
<td>VP[C]Terboxyethyl]FAGDFR</td>
<td>88.27</td>
<td>6</td>
<td>1.73E+09</td>
<td>1.88E+09</td>
</tr>
<tr>
<td>LK[C]Terboxyethyl]BNTKHK</td>
<td>33.78</td>
<td>3</td>
<td>3.56E+08</td>
<td>3.76E+08</td>
</tr>
<tr>
<td>DYYPLMDQK</td>
<td>146.67</td>
<td>7</td>
<td>1.18E+09</td>
<td>1.24E+08</td>
</tr>
<tr>
<td>FWENENQVTFK</td>
<td>102.01</td>
<td>5</td>
<td>4.59E+09</td>
<td>4.84E+09</td>
</tr>
<tr>
<td>AGFVC[C]Terboxyethyl]PTTPY</td>
<td>77.27</td>
<td>3</td>
<td>2.22E+09</td>
<td>2.24E+09</td>
</tr>
<tr>
<td>&lt;p&gt;202754&lt;/p&gt;</td>
<td>&lt;p&gt;KL&lt;/p&gt;</td>
<td>&lt;p&gt;Beta lactoglobulin (1 pmol on column)</td>
<td>IDALNENK</td>
<td></td>
</tr>
<tr>
<td></td>
<td>25.32</td>
<td>11</td>
<td>9.15E+09</td>
<td>8.75E+09</td>
</tr>
<tr>
<td>TPEVDEALEK</td>
<td>59.56</td>
<td>5</td>
<td>2.17E+10</td>
<td>2.13E+10</td>
</tr>
<tr>
<td>VLVDTTOYYK</td>
<td>46.50</td>
<td>5</td>
<td>1.29E+10</td>
<td>1.26E+10</td>
</tr>
<tr>
<td>TPEVDEALEK</td>
<td>41.33</td>
<td>6</td>
<td>4.28E+10</td>
<td>3.72E+10</td>
</tr>
<tr>
<td>WENC[CF]Terboxyethyl]AQK</td>
<td>22.78</td>
<td>5</td>
<td>3.64E+08</td>
<td>3.30E+08</td>
</tr>
<tr>
<td>LSNPTQLEEGC[CF]Terboxyethyl]</td>
<td>89.46</td>
<td>2</td>
<td>4.13E+10</td>
<td>4.16E+10</td>
</tr>
</tbody>
</table>

Large-Scale Relative Quantification Using the WiSIM-DIA Workflow

To further evaluate the analytical precision and quantitative accuracy of the WiSIM-DIA workflow when applied to large-scale quantitative experiments, a 500 ng on column E. coli digest and a 250 ng on column E. coli digest were compared. Each sample was run in triplicate using the WiSIM-DIA instrument method for quantification.

For targeted E. coli peptide selection and spectral library generation, a 1 µg E. coli sample was run in triplicate with a standard shotgun DDA experiment using Orbitrap analyzer full scan and ion trap CID MS/MS analyses on an Orbitrap Fusion instrument. 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.

Pinpoint software was used for all of the targeted data extraction. The Proteome Discoverer 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 C13 and C15 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 2 m/z CID MS/MS spectra.

Figure 7 shows the parameter setup for data processing using Pinpoint software. The relative expression ratios of targeted proteins between the 250 ng E. coli digest and 500 ng E. coli digest and CVs of each sample were calculated automatically. The p-value for each targeted peptide was also determined based on the spectrum match result. In this example, 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 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 8). Over 97% of quantified proteins gave exceptional quantitative accuracy for the detection of the two-fold expression change expected between the two samples (Figure 8). In addition, 85% of the 5,923 quantified peptides gave %CVs less than 15% (Figure 9).

By importing the 1,100 E. coli protein dataset into Thermo Scientific™ ProteinCenter™ software, functional GO annotation and pathway analysis was carried out for each targeted protein. Using the E. coli (G2) dataset as a reference, 240 targeted E. coli proteins were identified in an over-representation analysis, covering proteins of major primary metabolic pathways (Figure 10). The analytical precision observed for these 240 metabolic E. coli proteins was similarly distributed as peptides representing the entire dataset, which enabled accurate quantification of relevant proteins in further pathway studies (Figure 11).
Figure 7. WiSiM-DIA data processing set up using Pinpoint 1.3 software

Figure 8. Quantification summary for quantifying 1,100 E. coli proteins

Figure 9. Coefficient of variation of 5,923 quantified E. coli peptides

Figure 10. KEGG Pathways for the 1,100 targeted E. coli proteins

Figure 11. Coefficient of variation for 1,623 E. coli peptides representing major metabolic pathways
Conclusion
A unique data-independent acquisition workflow that collects HRAM wide-window SIM data and rapid CID MS/MS data in parallel on the new Orbitrap Fusion MS was developed. The method setup is simple, generic, and applicable to different types of samples. Any precursor ions detected by HRAM SIM acquisition can be quantified using an XIC with a ±5 ppm window and simultaneously confirmed using DIA-CID MS/MS by applying a targeted data extraction approach post-acquisition.

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. Ten atomole on-column LOD/LOQ and four orders of linear dynamic range were observed with the developed DIA workflow. The reproducible qualitative and quantitative record for each sample allowed large-scale quantification of over a thousand proteins of interest with excellent analytical precision and great quantitative accuracy.

References