A Novel Automated and Highly Selective Phosphopeptide Enrichment Strategy for Successful Phosphopeptide Identification and Phosphosite Localization

Shuai Wu1; Kenneth Newton1; Linfeng Wu1; Jordy J. Hsiao1; Valery G. Voinov2,3; Joseph S. Beckman2,3

1Agilent Technologies Inc., Santa Clara, CA
2e-Msion Inc., Corvallis, OR 3Oregon State University, Corvallis, OR
Phosphopeptide analysis has remained challenging due to reproducibility and selectivity variations caused by manual sample preparation. The labile phospho-site on peptides also makes it difficult to localize the phosphoryl group in a CID experiment. In the Phosphopeptide Challenge of the Human Proteome Project (HPP), Agilent AssayMAP Bravo platform provided a fully automated and highly selective phosphopeptide enrichment workflow using high capacity Fe(III)-NTA cartridges. Sample analysis was performed on a nano-LC/6550 iFunnel Q-TOF system with an electromagnetostatic (EMS) ECD cell. Peptide identification and phosphopeptide quantitation were performed using data-dependent acquisition (DDA) together with Skyline. Phosphosite localization was analyzed using ECD and Byonic software.

### Experimental

The AssayMAP Phosphopeptide Enrichment App was used for automated phosphopeptide enrichment using Fe(III)-NTA cartridges (Figure 1). The enriched phosphopeptides were cleaned up further using C18 cartridges. The buffer conditions on the AssayMAP and instrument parameters are listed in Table 1.

#### Peptide identification and phosphosite localization

The Agilent 1290 Infinity II LC system was converted to nanoflow LC with the Agilent Infinity UHPLC Nanodapter. This nanoflow LC was connected to the Agilent nanoESI source and coupled to the 6550 iFunnel Q-TOF LC/MS for peptide identification (Figure 2A). The LC parameters were listed in Table 2. Both "Phosphopeptide" sample and enriched "Phosphopeptide-Yeast" sample were resuspended in 50 µL of 5% ACN, 0.1% TFA.

### Table 2. Nano-LC Parameters

<table>
<thead>
<tr>
<th>LC Conditions</th>
<th>Guard Column</th>
<th>Analytical Column</th>
<th>Solvent A</th>
<th>Solvent B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PepMap C18, 75 μm x 2 cm at 60 °C</td>
<td>PepMap C18, 75 μm x 25 cm at 60 °C</td>
<td>0.1% Formic Acid, 0.1% deactivator in Water</td>
<td>0.1% Formic Acid in 90 % Acetonitrile</td>
</tr>
<tr>
<td>Flow rate</td>
<td>0.085 mL/min primary flow 300 mL/min on-column flow rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time (min)</td>
<td>0</td>
<td>3</td>
<td>97</td>
<td>70</td>
</tr>
<tr>
<td>B (%)</td>
<td>90</td>
<td>37</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>Time (min)</td>
<td>95</td>
<td>70</td>
<td>120</td>
<td>3</td>
</tr>
<tr>
<td>B (%)</td>
<td>Injection volume</td>
<td>2 µL for CID, 4 µL for ECD, 1 µL for MS1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For peptide identification, CID was performed using DDA with triplicate injections. Spectrum Mill was used to search against the customized database containing Swiss-Prot human proteins with a 1.5 % false discovery rate.

The EMS ECD cell was designed to merge with a shorter hexapole collision cell to replace the original hexapole collision cell in the Agilent Q-TOF system. For phosphosite localization, ECD experiments were performed using EMS ECD cell (Figure 2B). The ECD spectra were analyzed using Byonic software from Protein Metrics Inc. with the same Swiss-Prot human proteins database.
Results and Discussion

Peptide Identification and phosphosite localization

Figure 4A shows the total ion chromatogram (TIC) of “Phosphopeptide” sample with a 90-minute gradient. With triplicate sample injections using DDA, Spectrum Mill was able to identify 437 distinct peptides with 294 distinct phosphopeptides. All of the 89 non-phosphopeptides in the peptide sequence list provided by HUPO were identified. The phosphorylated counterparts were also identified without confirmation of the phosphosite location using CID. Figure 3A is the CID spectrum of peptide VVEAVNSDSDEFGIPK showing complete sequence coverage with b, y type ions. Figure 3B, 3C and 3D are the ECD spectra of peptides with the same sequence but with mono-, di- and tri- phosphorylation. Byonic confirmed the location of the phosphosites based on the c, z type ions in the ECD spectra. When peptide size increased with more phosphosites, the precursor ion with +3 charge state became more abundant leading to a better ECD spectrum with a higher sequence coverage (Figure 3C). The more abundant precursor ion with the same charge state also resulted in a better ECD spectrum (Figure 3C and 3D).

![Figure 3A](image1.png)

**Figure 3A.** CID spectrum of peptide VVEAVNSDSDEFGIPK showing complete sequence coverage with b, y type ions.

![Figure 3B](image2.png)

**Figure 3B.** ECD spectrum of peptide VVEAVNSDSDEFGIPK with mono-phosphorylation localized by c, z type ions.

![Figure 3C](image3.png)

**Figure 3C.** ECD spectrum of peptide VVEAVNSDSDEFGIPK with di-phosphorylation localized by c, z type ions. Note the presence of most of the sequence ions in the ECD spectra even with the expected low efficiency for low charge state.

![Figure 3D](image4.png)

**Figure 3D.** ECD spectrum of peptide VVEAVNSDSDEFGIPK with tri-phosphorylation localized by c, z type ions.
An automated phosphopeptide enrichment with both qualitative and quantitative analysis using LC/Q-TOF has been implemented for HUPO Phosphopeptide Challenge.

- Within "Phosphopeptide" sample, CID experiment identified 437 distinct peptides with 294 phosphopeptides. All the 89 non-phosphopeptides from HUPO sequence list were identified. ECD experiment determined the location of phosphosites for 96 phosphopeptides based on the 89 non-phosphopeptide sequences from HUPO peptide list.
- Within enriched "Phosphopeptide-Yeast" sample, 287 distinct peptides were identified, of which 264 were distinct phosphopeptides. The overall selectivity of the enrichment is about 92.0%.
- Moreover, 95 out of 96 phosphopeptides spiked into the yeast were still identified from the enriched "Phosphopeptide-Yeast" sample. This showed a really high recovery of the enrichment.

For Research Use Only. Not for use in diagnostic procedures.