Post-translational Modification Identification by Mass Spectrometry

Liwen Zhang
Mass Spectrometry and Proteomics Facility
The Ohio State University

Summer Workshop 2018
PTM: Chemical Modifications on Proteins
Increases the Functional Diversity of the Proteome
Regulate Activity, Localization and Interaction

Image from https://www.lifetechnologies.com/overview-post-translational-modification.html
# Post-translational Modifications

<table>
<thead>
<tr>
<th>Modification</th>
<th>Modified Residue</th>
<th>Mass Shift (Da)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disulphide bond formation</td>
<td>C</td>
<td>-2.0157</td>
<td>protein stability</td>
</tr>
<tr>
<td>Methylation</td>
<td>K/R</td>
<td>14.0157</td>
<td>regulation of gene expression</td>
</tr>
<tr>
<td>Hydroxylation</td>
<td>P</td>
<td>15.9949</td>
<td>Protein stability and protein-ligand interaction</td>
</tr>
<tr>
<td>Oxidation of Met</td>
<td>M</td>
<td>15.9949</td>
<td>usually introduced during digestion</td>
</tr>
<tr>
<td>Acetylation</td>
<td>K</td>
<td>42.0106</td>
<td>Protein stability, protection of N terminus, regulation of protein-DNA interactions</td>
</tr>
<tr>
<td>S-Glutathionylation</td>
<td>C</td>
<td>305.0682</td>
<td>oxidative stress; preventing irreversible oxidation of protein thiol; control of cell-signaling pathways by modulating protein function</td>
</tr>
<tr>
<td>Phosphorylation</td>
<td>S/T/Y</td>
<td>79.9663</td>
<td>Signaling, activation/inactivation of enzyme activities, modulation of molecular interactions</td>
</tr>
<tr>
<td>Nitration</td>
<td>Y</td>
<td>44.9851</td>
<td>oxidative damage during inflammation</td>
</tr>
<tr>
<td>Ubiquitination</td>
<td>K</td>
<td>114.0429</td>
<td>destruction signal</td>
</tr>
<tr>
<td>Oxidation of Cys</td>
<td>C</td>
<td>31.9721/47.9847</td>
<td>oxidative/nitrosative stress</td>
</tr>
<tr>
<td>Nitrosylation</td>
<td>C</td>
<td>28.9902</td>
<td>oxidative/nitrosative stress</td>
</tr>
<tr>
<td>Glycosylation</td>
<td>N/S/T</td>
<td>......</td>
<td>Excreted proteins, cell-cell recognition/signaling O-GlcNAc, reversible, regulatory functions</td>
</tr>
</tbody>
</table>
Method for PTM Detections

- Specific Fluorescence Dye: Pre-Q Diamond---Phosphorylation
  Pre-Q Emerald---Glycosylation
- Specific Antibodies: Nitration; Phosphorylation; Methylation;
  Ubiquitination; Acetylation
- Mass Shift/PI Change: Gel Approaches

Pros: High sensitivity, Favor the modification
Cons: Lack of accuracy
  Interference from other residues
  Lack the ability to locate the actual modification sites
Peptide Modifications Identification by MS/MS

- Virtually anything that shifts the mass can be determined by MS.
- Using MS/MS allows for identification of the modification AND location.
- You must have an idea of what modification you are looking for.
- First use protein stain to examine various modifications.
- Western analysis VS MS analysis.

Cons: Detection based on ionization efficiency and abundance of the peptides, thus may not favor modified peptides.
Identification of PTMs by MS:

- **Know the Modification**
  (Stability, Mass Shift, possible modification site)

- High Samples Purity

- High Concentration

- High Mass Accuracy/Resolution
Example 1 — Phosphorylation Identification by MS/MS

$m/z=850.54^{2+}$

Ser Val Asn Thr Val

Ser(PO$_4$)

Val Val Asn Asn

Ser(PO$_4$)-Val

Ser(PO$_4$)

Thr Asn Val Ser Val

M-H$_3$PO$_4$

Ser Val Asn Thr Val

266.02

(87+80+99)

Ser$_{(PO_4)}$-Val

167.02

(87+80)

Ser$_{(PO_4)}$
Example 3 — Identification of DMPO Adduct by MS/MS

Theoretical $m/z = 859.94^{2+}$
Observed $m/z = 859.89^{2+}$

DMPO = $C_6H_9NO$ (111.0684 Da)
Difficulties for PTMs Identification

- Generating Peptides containing modified residue
  --Different Enzymes/Digestion Condition

- Low population of modifications (usually less than 1%)
  --Enrichment
  --Targeted Method

- Modification not stable
  --Experiment Condition (reducing reagents, enzymes)
  --Fresh sample

- Modification labile in MS analysis
  --Neutral Loss Method
  --Different Fragmentation Method
### Enzymes for Proteome Research

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Cleavage Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trypsin</td>
<td>K-X and R-X</td>
</tr>
<tr>
<td>Endoprotease Lys-C</td>
<td>K-X except when X = P</td>
</tr>
<tr>
<td>Endoprotease Arg-C</td>
<td>R-X except when X = P</td>
</tr>
<tr>
<td>Endoprotease Asp-N</td>
<td>X-D</td>
</tr>
<tr>
<td>Endoprotease Glu-C</td>
<td>E-X except when X = P</td>
</tr>
<tr>
<td>Chymotrypsin</td>
<td>L-X, F-X, Y-X and W-X</td>
</tr>
<tr>
<td>Cyanogen Bromide</td>
<td>X-M</td>
</tr>
</tbody>
</table>
Using the CORRECT Enzyme

Trypsin: 309 AALEFEDGMYANLGIPILLASNFISPNMVTHLQSEGVLGLGPYPLK356
Chymotrypsin: 329 ASNF332,
GluC: 315 DGMYANLGIPILLASNFISPNMVTHLQSE344
Optimize Digestion Conditions

<table>
<thead>
<tr>
<th>Start</th>
<th>Observed</th>
<th>Mr(expt)</th>
<th>Mr(calc)</th>
<th>ppm</th>
<th>M</th>
<th>Score</th>
<th>Peptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA2-20</td>
<td>932.9369</td>
<td>1863.8592</td>
<td>1863.9</td>
<td>-0.34</td>
<td>0</td>
<td>104</td>
<td>M.ADESSDAAGEPQPAPAPVR.R</td>
</tr>
<tr>
<td>AA2-21</td>
<td>674.3265</td>
<td>2019.9576</td>
<td>2020</td>
<td>-1.67</td>
<td>1</td>
<td>95</td>
<td>M.ADESSDAAGEPQPAPAPVR.R</td>
</tr>
<tr>
<td>AA2-22</td>
<td>726.3605</td>
<td>2176.0598</td>
<td>2176</td>
<td>-1.06</td>
<td>2</td>
<td>66</td>
<td>M.ADESSDAAGEPQPAPAPVR.R</td>
</tr>
<tr>
<td>AA23-37</td>
<td>555.9385</td>
<td>1664.7936</td>
<td>1664.8</td>
<td>1.78</td>
<td>1</td>
<td>55</td>
<td>R.SSANYRAYATEPHAK.K</td>
</tr>
<tr>
<td>AA29-37</td>
<td>494.2502</td>
<td>986.4858</td>
<td>986.48</td>
<td>3.69</td>
<td>0</td>
<td>26</td>
<td>R.AYATEPHAK.K</td>
</tr>
<tr>
<td>AA29-39</td>
<td>622.3464</td>
<td>1242.6782</td>
<td>1242.7</td>
<td>4.97</td>
<td>2</td>
<td>44</td>
<td>R.AYATEPHAKK.S</td>
</tr>
<tr>
<td>AA29-41</td>
<td>486.9404</td>
<td>1457.7994</td>
<td>1457.8</td>
<td>0.25</td>
<td>3</td>
<td>35</td>
<td>R.AYATEPHAKKSK.I</td>
</tr>
<tr>
<td>AA47-59</td>
<td>382.7495</td>
<td>1526.9691</td>
<td>1527</td>
<td>7.14</td>
<td>2</td>
<td>35</td>
<td>R.KLQLKTLMLQIAK.Q</td>
</tr>
<tr>
<td>AA52-64</td>
<td>530.952</td>
<td>1589.8341</td>
<td>1589.8</td>
<td>4.51</td>
<td>1</td>
<td>85</td>
<td>K.TLMQIAKQEMERE.E</td>
</tr>
<tr>
<td>AA52-69</td>
<td>735.7083</td>
<td>2204.1029</td>
<td>2204.1</td>
<td>3.59</td>
<td>3</td>
<td>44</td>
<td>K.TLMQIAKQEMEREAEER.R</td>
</tr>
<tr>
<td>AA60-69</td>
<td>436.1964</td>
<td>1305.5675</td>
<td>1305.6</td>
<td>4.31</td>
<td>1</td>
<td>25</td>
<td>K.QEMEREAEER.R</td>
</tr>
<tr>
<td>AA76-99</td>
<td>730.1263</td>
<td>2916.476</td>
<td>2916.5</td>
<td>5.98</td>
<td>2</td>
<td>31</td>
<td>R.VLRTRCQPLEDGLGFEELQDLCR.Q</td>
</tr>
<tr>
<td>AA79-99</td>
<td>850.4134</td>
<td>2548.2183</td>
<td>2548.2</td>
<td>5.24</td>
<td>2</td>
<td>95</td>
<td>R.TRCQPLEDGLGFEELQDLCR.Q</td>
</tr>
<tr>
<td>AA81-99</td>
<td>1146.5381</td>
<td>2291.0616</td>
<td>2291</td>
<td>2.36</td>
<td>0</td>
<td>115</td>
<td>R.CQPLEDGLGFEELQDLCR.Q</td>
</tr>
<tr>
<td>AA81-104</td>
<td>725.1068</td>
<td>2896.3981</td>
<td>2896.4</td>
<td>0.75</td>
<td>1</td>
<td>69</td>
<td>R.CQPLEDGLGFEELQDLCRQHLAR.V</td>
</tr>
<tr>
<td>AA105-118</td>
<td>565.6145</td>
<td>1693.8217</td>
<td>1693.8</td>
<td>3.43</td>
<td>2</td>
<td>41</td>
<td>R.VDKVDEERVDEAK.V</td>
</tr>
<tr>
<td>AA105-121</td>
<td>1012.028</td>
<td>2022.0415</td>
<td>2022</td>
<td>7.17</td>
<td>3</td>
<td>39</td>
<td>R.VDKVDEERVDEAKVT.K</td>
</tr>
<tr>
<td>AA119-132</td>
<td>787.4489</td>
<td>1572.8833</td>
<td>1572.9</td>
<td>-2.94</td>
<td>0</td>
<td>107</td>
<td>R.ISADAMMQALLGTR.A</td>
</tr>
<tr>
<td>AA119-137</td>
<td>745.4208</td>
<td>2233.2405</td>
<td>2233.2</td>
<td>3.91</td>
<td>2</td>
<td>65</td>
<td>K.VTKNITEIALDLTQKIYDLR.G</td>
</tr>
<tr>
<td>AA119-139</td>
<td>605.5953</td>
<td>2418.3522</td>
<td>2418.3</td>
<td>1.68</td>
<td>3</td>
<td>54</td>
<td>K.VTKNITEIALDLTQKIDLRG.K</td>
</tr>
<tr>
<td>AA122-132</td>
<td>623.3372</td>
<td>1244.6599</td>
<td>1244.7</td>
<td>-1.06</td>
<td>0</td>
<td>79</td>
<td>K.NITEIALDLTKI</td>
</tr>
<tr>
<td>AA122-139</td>
<td>697.7211</td>
<td>2090.1414</td>
<td>2090.1</td>
<td>2.03</td>
<td>2</td>
<td>48</td>
<td>K.NITEIALDLTKIYDLR.G</td>
</tr>
<tr>
<td>AA138-146</td>
<td>368.2345</td>
<td>1101.6816</td>
<td>1101.7</td>
<td>4.08</td>
<td>2</td>
<td>26</td>
<td>R.GFKKRPTLR.R</td>
</tr>
<tr>
<td>AA138-147</td>
<td>420.2665</td>
<td>1257.7776</td>
<td>1257.8</td>
<td>-0.46</td>
<td>3</td>
<td>26</td>
<td>R.GFKKRPTLR.R</td>
</tr>
<tr>
<td>AA148-163</td>
<td>866.964</td>
<td>1731.9134</td>
<td>1731.9</td>
<td>0.61</td>
<td>1</td>
<td>79</td>
<td>R.VRISADAMMQALLGTR.A</td>
</tr>
<tr>
<td>AA150-163</td>
<td>739.3765</td>
<td>1476.7385</td>
<td>1476.7</td>
<td>-2.94</td>
<td>0</td>
<td>107</td>
<td>R.ISADAMMQALLGTR.A</td>
</tr>
<tr>
<td>AA164-171</td>
<td>466.2661</td>
<td>930.5176</td>
<td>930.51</td>
<td>4.52</td>
<td>1</td>
<td>58</td>
<td>R.AKESDLR.A</td>
</tr>
<tr>
<td>AA176-187</td>
<td>505.9445</td>
<td>1514.8118</td>
<td>1514.8</td>
<td>4.36</td>
<td>3</td>
<td>59</td>
<td>K.QVKKDIEKENR.E</td>
</tr>
<tr>
<td>AA188-205</td>
<td>678.3385</td>
<td>2031.9937</td>
<td>2032</td>
<td>6.93</td>
<td>2</td>
<td>43</td>
<td>R.EVGDWRKNIDALSGMEGR.K</td>
</tr>
<tr>
<td>AA194-205</td>
<td>645.8306</td>
<td>1289.6466</td>
<td>1289.6</td>
<td>5.29</td>
<td>1</td>
<td>78</td>
<td>R.KNIDALSGMEGR.K</td>
</tr>
<tr>
<td>AA195-205</td>
<td>581.7822</td>
<td>1161.5499</td>
<td>1161.5</td>
<td>4.39</td>
<td>0</td>
<td>95</td>
<td>R.KNIDALSGMEGR.K</td>
</tr>
<tr>
<td>AA195-206</td>
<td>645.8309</td>
<td>1289.6472</td>
<td>1289.6</td>
<td>5.77</td>
<td>1</td>
<td>54</td>
<td>R.KNIDALSGMEGR.K</td>
</tr>
<tr>
<td>AA195-207</td>
<td>709.8787</td>
<td>1417.7414</td>
<td>1417.7</td>
<td>4.74</td>
<td>2</td>
<td>46</td>
<td>R.KNIDALSGMEGR.K</td>
</tr>
<tr>
<td>AA195-211</td>
<td>627.329</td>
<td>1878.9651</td>
<td>1879</td>
<td>1.58</td>
<td>4</td>
<td>29</td>
<td>R.KNIDALSGMEGRK.KKFFEG.-</td>
</tr>
</tbody>
</table>

Trypsin; 1:200; RT; 20 min Digestion

Sequence coverage: 97.2%
Adding Ion Pairing Reagent

<table>
<thead>
<tr>
<th>AA Start - End</th>
<th>Observed</th>
<th>Mr(expt)</th>
<th>Mr(calc)</th>
<th>ppm</th>
<th>M</th>
<th>Score</th>
<th>Peptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>22-28</td>
<td>467.2008</td>
<td>932.387</td>
<td>932.387</td>
<td>0.53</td>
<td>1</td>
<td>34</td>
<td>R.RSSANYR.A + Phospho (ST)</td>
</tr>
<tr>
<td>23-28</td>
<td>349.1678</td>
<td>696.321</td>
<td>696.319</td>
<td>2.7</td>
<td>0</td>
<td>42</td>
<td>R.SSANYR.A</td>
</tr>
<tr>
<td>23-28</td>
<td>349.1684</td>
<td>696.3222</td>
<td>696.319</td>
<td>4.54</td>
<td>0</td>
<td>43</td>
<td>R.SSANYR.A</td>
</tr>
</tbody>
</table>

MSMS for RSS(PO4)ANYR

MSMS for SSANYR
Difficulties for PTMs Identification

- Generating Peptides containing modified residue
  -- Different Enzymes/Digestion Condition

- Low population of modifications
  -- Enrichment
  -- Targeted Method (For better MSMS)
PTM Enrichment Techniques (Start with mg of samples)

Acetylation: Anti-acetyl lysine polyclonal antibody

Phosphopeptides: Immobilized Metal Affinity Chromatography (Fe3+)
Metal Oxide Affinity Chromatography (TiO2, ZrO2)
Reversible Covalent Binding


Phospho-Threonine Antibody (P-Thr-Polyclonal)

Nitration: Anti-Nitrotyrosine polyclonal antibody

Ubiquitination: K-ε-GG–specific antibody (enrichment has to be done prior to other treatment on lysine)

Glycopeptides: Lectin affinity enrichment
Covalent Interactions
Chromatographic separation

(Glycopeptide enrichment and separation for protein glycosylation analysis., *J Sep Sci.* 2012 Sep;35(18):2341-72.)

Methylation: Anti-methyl lysine/aragginine antibody
Enrichment for Phosphorylation

- mLKB1
- extract cells
- cell lysis
- In-solution digestion
- peptides
- TiO₂ beads
- protein

Slide courtesy of Nilini S. Ranbaduge
Multi-dimensional Separation

Phosphopeptides/proteins Identification

Slide courtesy of Nilini S. Ranbaduge
Example---Targeted MS Scan (for Known Modifications)

Full Scan
100:1

SIC of 516.7793\(^{2+}\) (SLVLC\textsubscript{(CAM)}TPSR)

SIC of 479.2730\(^{2+}\) (SLVL\textsubscript{f}GlyTPSR)

MS/MS of 516.7793\(^{2+}\) (SLVLC\textsubscript{(CAM)}TPSR)

MS/MS of 479.2730\(^{2+}\) (SLVL\textsubscript{f}GlyTPSR)

NOT Observed!!!!
Example---Targeted MS Scan (for Known Modifications)

Full Scan

Targeted @ 479.2730²⁺(SLVFgTyrTSR)/516.7793²⁺(SLVLC(CAM)TPSR)

MS/MS of 516.7793²⁺ (SLVLC(CAM)TPSR)
Difficulties for PTMs Identification

- Generating Peptides containing modified residue
  --Different Enzymes/Digestion Condition

- Low population of modifications
  --Enrichment
  --Targeted Method

- Modification not stable
  --Experiment Condition (reducing reagents, enzymes)
  --Fresh sample
Adjust Experiment Condition to Preserve Modifications

Chain A

185N A C187 G S G Y D F D V F V V R199

Chain B

138L V E G C142 L V G G R146

A

274.25 y2A 373.31 y3A 520.42 y4A 619.48 y5A 734.51 y6A 996.57 y8A 1159.62 y9A 1216.57 y10A 1303.64 y11A 1360.66 y12A

B

408.17 465.20 552.31 609.30 772.30 1034.35 1149.39 1248.41 1395.53 1494.50 1593.59

289.23 342.33 501.10 520.42 619.48 734.51 996.57 1159.62 1216.57 1303.64 1360.66

m/z
Using Different Enzyme to Protect Modifications

Theoretical m/z = 783.4089²⁺, Observed m/z = 783.4080²⁺, Mass Error = 1.15ppm

O=Pyrrolysine

Using Different Enzyme to Protect Modifications
## LTQ DATA—Fresh Sample

<table>
<thead>
<tr>
<th>Start</th>
<th>Peptide</th>
<th>Observed Mr</th>
<th>Mr(calc)</th>
<th>Delta</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>92-112</td>
<td>K.SVSTPSEAGS,QSGDGAVGSR.T</td>
<td>1950.856</td>
<td>1949.856</td>
<td>0.729</td>
<td>108</td>
</tr>
<tr>
<td>92-112</td>
<td>K.SVSTPSEAGS,QSGDGAVGSR.T + Deamidated (NQ)</td>
<td>1950.84</td>
<td>1950.84</td>
<td>0.025</td>
<td>94</td>
</tr>
<tr>
<td>92-112</td>
<td>K.SVSTPSEAGS,QSGDGAVGSR.T + Deamidated (NQ)</td>
<td>1950.84</td>
<td>1950.84</td>
<td>0.025</td>
<td>94</td>
</tr>
<tr>
<td>92-112</td>
<td>K.SVSTPSEAGS,QSGDGAVGSR.T + Deamidated (NQ); Phospho (ST)</td>
<td>2030.807</td>
<td>2030.807</td>
<td>-0.261</td>
<td>86</td>
</tr>
<tr>
<td>92-112</td>
<td>K.SVSTPSEAGS,QSGDGAVGSR.T + Deamidated (NQ); Phospho (ST)</td>
<td>2030.807</td>
<td>2030.807</td>
<td>-0.261</td>
<td>86</td>
</tr>
</tbody>
</table>

Sequence Coverage 39%; T95 was phosphorylated
Orbitrap XL Data---Two Days old, Stored at 4C

<table>
<thead>
<tr>
<th>Start</th>
<th>Observed</th>
<th>Mr(expt)</th>
<th>Mr(calc)</th>
<th>ppm</th>
<th>Score</th>
<th>Peptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>92-112</td>
<td>975.9321</td>
<td>1949.8496</td>
<td>1949.8563</td>
<td>-3.43</td>
<td>138</td>
<td>K.SVSTPSEAGSQDSGDGAVGSR.T</td>
</tr>
<tr>
<td>92-112</td>
<td>975.9351</td>
<td>1949.8556</td>
<td>1949.8563</td>
<td>-0.36</td>
<td>149</td>
<td>K.SVSTPSEAGSQDSGDGAVGSR.T</td>
</tr>
<tr>
<td>92-112</td>
<td>975.9354</td>
<td>1949.8562</td>
<td>1949.8563</td>
<td>-0.052</td>
<td>149</td>
<td>K.SVSTPSEAGSQDSGDGAVGSR.T</td>
</tr>
<tr>
<td>92-112</td>
<td>975.9377</td>
<td>1949.8608</td>
<td>1949.8563</td>
<td>2.33</td>
<td>143</td>
<td>K.SVSTPSEAGSQDSGDGAVGSR.T</td>
</tr>
<tr>
<td>92-112</td>
<td>975.9377</td>
<td>1949.8609</td>
<td>1949.8563</td>
<td>2.39</td>
<td>130</td>
<td>K.SVSTPSEAGSQDSGDGAVGSR.T</td>
</tr>
<tr>
<td>92-112</td>
<td>975.9386</td>
<td>1949.8626</td>
<td>1949.8563</td>
<td>3.27</td>
<td>125</td>
<td>K.SVSTPSEAGSQDSGDGAVGSR.T</td>
</tr>
<tr>
<td>92-112</td>
<td>975.9403</td>
<td>1949.8661</td>
<td>1949.8563</td>
<td>5.02</td>
<td>142</td>
<td>K.SVSTPSEAGSQDSGDGAVGSR.T</td>
</tr>
<tr>
<td>92-112</td>
<td>975.9407</td>
<td>1949.8668</td>
<td>1949.8563</td>
<td>5.39</td>
<td>135</td>
<td>K.SVSTPSEAGSQDSGDGAVGSR.T</td>
</tr>
</tbody>
</table>

Sequence Coverage 53%; No Phosphorylation Detected

Detected 30 times!!!
Difficulties for PTMs Identification

- Generating Peptides containing modified residue
  --Different Enzymes/Digestion Condition

- Low population of modifications
  --Enrichment
  --Targeted Method

- Modification not stable
  --Experiment Condition (reducing reagents, enzymes)
  --Fresh sample

- Modification labile in MS analysis
  --Neutral Loss Method
  --Different Fragmentation Method
Neutral Loss Scan

Scan Dissociate Scan

A

PTM-specific ion signal

Stable PTM

MS/MS

\( \Delta m \) (PTM, position)

B

PTM-specific neutral loss

Labile PTM

MS/MS

\( \Delta m \) (PTM = \( \text{PTM} \) )

(aa)

*C-terminal amino acid residue
- Amino acid residue modified with stable PTM
*N-terminal amino acid residue
- Amino acid residue modified with labile PTM

Hydrothermal synthesis of $\alpha$-Fe$_2$O$_3$@SnO$_2$ core–shell nanotubes for highly selective enrichment of phosphopeptides for mass spectrometry analysis *Nanoscale*, 2010, 2, 1892-1900

- Help identify the sequence but don’t give information about the location of the modification if multiple sites coexisting.
- Sensitive not ideal

**Alpha Casein**

$\text{YKVPQLEIVPNSAEER}$

Most abundant peak represents a loss of phosphoric acid (98 Daltons/2 = 49 due to the fact that the peptide has a +2 charge state).

$\text{NL of Phosphorylation: 98 (+1), 49 (+2)…}$
ECD/ETD: Keep the Phosphorylation Intact!!!

H$_3$PO$_4$ Group may Leave upon CID Dissociation; Actual Amino Acid Location of H$_3$PO$_4$ Group May not be Clear if >2 Sites in the Sequence

**CID:** a, b and y ion, loss of H$_2$O, NH$_3$, side-chain

**ECD/ETD:** c, y and z ions

Electron Capture Dissociation (ECD, ICR)/ Electron Transfer Dissociation (ETD, Traps) will Generate Peptide Fragments without Losing the Sidechain Phosphorylation

_Sweet, Anal. Chem. 2006, 78, 7563-7569_
Unknown Truncation Sites Identification

Separate Suspected Truncation Product(s) from Intact Protein

Intact MS Measurement

Top down Sequencing

Separation on 1D SDS PAGE

Cut the band (truncated protein)

Digestion

LC/MSMS

Peptide Sequencing

MASCOT (trypsin OR no enzyme)

Protein sequence

Suggested truncation sites

Search against new sequences

Validation
Example---Unknown Truncation Sites Identification

MASCOT RESULT:

1  MVLWILWRPF  GFSGRFLKLE  SHSITESSKL  IPVAWTSLTQ  MLLEAPGIFL
51  LGQRKRFSTM  PETETHERET  ELFSPPSDVR  GMTKLDRTAF  KKTVNIPVLK
101 VRKEIVSKLM  RSLKRAALQR  PGIIRVIEDP  EDKESRLIIM  DPYKIFTTHDS
151 FEKAELSVLE  QLNVSPQISK  YNLELYEHF  KSEEILRAVL  PEGQDVTSGF
201 SRIIGHIAHNL  LRDHQLPFKH  LIGQVMIDKN  PGITSAVNKI  NNIDNMRRNF
251 QMEVLSGEQN  MMTKVRENHY  TYEFDFSKVY  WNPRLSTEHS  RITELLKPGD
301 VLFDVFAGVG  PFAIPVAKKN  CTVFANDLNP  ESHKWLLYNC  KLNKVDQKV
351 VFNLDGKDQL  QGPVKEELMQ  LLGLSKEKRP  SVHVVMLPA  KAIEFLSAFK
401 WLLDGQPCSS  EFLPIVHCYS  FSKDANPAED  VRQRAGAVLG  ISLEACSVH
451 LVRNVAPNKE  MLCITFQIPA  SVLYKNQTRN  PENHEDPLLK  RQTAEAFSD
501 EKTQIVSNT

~260AA =~27-28kDa

Conservation of structure and mechanism by Trm5 enzymes *RNA* 2013 Sep; 19(9): 1192–1199
Constructed new sequences (a total of 24) and searched the data against these sequences:

>16502 Sequence 1 (241-509)

NNIDNMYRNFQMEVLSGEQNMMTKVRENNTYEFDSKVYWNPRPLSTEHSRITELLKPGDVLF DVFAGVGPGPAIPVAKKNCTVFANDLNPE SHKWLLYNCKLNVKDVQKVVFNLDRGDQLQGPKVEELMQLLGLSKERKPSVHVMNLPA KAIEFLSAKWLDDQPCSEFLPIVHCYSFSDKANPAEDVRQRAGAVLGISLEACSSVHLVR NVAPNKEMLCITFQIPASVLYKNQTRNPENHEDPPLKRQR TAEAFSDEKTQIVSNT <

>16502 Sequence 2 (242-509)

OIIDNMYRNFQMEVLSGEQNMMTKVRENNTYEFDSKVYWNPRPLSTEHSRITELLKPGDVLF DVFAGVGPGPAIPVAKKNCTVFANDLNPE SHKWLLYNCKLNVKDVQKVVFNLDRGDQLQGPKVEELMQLLGLSKERKPSVHVMNLPAKAIEFLSAKWLDDQPCSEFLPIVHCYSFSDKANPAEDVRQRAGAVLGISLEACSSVHLVR NVAPNKEMLCITFQIPASVLYKNQTRNPENHEDPPLKRQR TAEAFSDEKTQIVSNT <

>16502 Sequence 23 (263-509)

TKVRENNTYEFDSKVYWNPRPLSTEHSRITELLKPGDVLF DVFAGVGPGPAIPVAKKNCTVFANDLNPE SHKWLLYNCKLNVKDVQKVVFNLDRGDQLQGPKVEELMQLLGLSKERKPSVHVMNLPAKAIEFLSAKWLDDQPCSEFLPIVHCYSFSDKANPAEDVRQRAGAVLGISLEACSSVHLVR NVAPNKEMLCITFQIPASVLYKNQTRNPENHEDPPLKRQR TAEAFSDEKTQIVSNT <

>16502 Sequence 24 (264-509)

KVRENNTYEFDSKVYWNPRPLSTEHSRITELLKPGDVLF DVFAGVGPGPAIPVAKKNCTVFANDLNPE SHKWLLYNCKLNVKDVQKVVFNLDRGDQLQGPKVEELMQLLGLSKERKPSVHVMNLPAKAIEFLSAKWLDDQPCSEFLPIVHCYSFSDKANPAEDVRQRAGAVLGISLEACSSVHLVR NVAPNKEMLCITFQIPASVLYKNQTRNPENHEDPPLKRQR TAEAFSDEKTQIVSNT <

MASSMATRIX Results:

<table>
<thead>
<tr>
<th>hit#</th>
<th>score</th>
<th>decoy%</th>
<th>protein description</th>
</tr>
</thead>
<tbody>
<tr>
<td>hit1</td>
<td>3533</td>
<td>0.00%</td>
<td>16502 Sequence 12 (252-509)</td>
</tr>
<tr>
<td>hit2</td>
<td>3522</td>
<td>0.00%</td>
<td>16502 Sequence 23 (263-509)</td>
</tr>
<tr>
<td>hit3</td>
<td>3519</td>
<td>0.00%</td>
<td>16502 Sequence 14 (254-509)</td>
</tr>
<tr>
<td>hit4</td>
<td>3512</td>
<td>0.00%</td>
<td>16502 Sequence 21 (261-509)</td>
</tr>
</tbody>
</table>
Conservation of structure and mechanism by Trm5 enzymes RNA. 2013 Sep; 19(9): 1192–1199
Conservation of structure and mechanism by Trm5 enzymes. RNA. 2013 Sep; 19(9): 1192–1199

Observed m/z = 768.3573±
Theoretical m/z = 768.3539±
Mass Error = 4.42ppm
Glycosylation—A Different Story

“The cell surface landscape is richly decorated with oligosaccharides anchored to proteins or lipids within the plasma membrane. Cell surface oligosaccharides mediate the interactions of cells with each other and with extracellular matrix components.” [Science 291:2337]

Why Glycosylation Investigation is SO HARD???

- Only Protein Modification Requiring Detailed Structural Characterization (sugar chain/branched)
- Poor Ionization efficiency (large mass increase, sometimes has negative charge, hydrophobicity..)
- Heterogeneity
- Large Size
Determination of site-specific glycan heterogeneity on glycoproteins


- Enrichment (Lectin Affinity Binding)
- Derivatization (Permethylation)
- LC Separation (PGC Column for Glycans, HILIC Column for Glycopeptides)
- Both MALDI and ESI will Work
- High Mass Accuracy Desired
- Negative Mode Works Better
- ETD Improve the MSMS Identification