Some considerations on the analysis of post-translational modifications (PTM) by MS

PW, MQ, 30.9.2013

Here we take the example of **phosphorylation** to explain some general concepts relative to the analysis of PTM's by mass spectrometry.

Identification and localization of a PTM (here: phosphorylation)

In the majority of cases, proteins are first cleaved into peptide fragments with a specific protease (Trypsin; more rarely Glu-C, Asp-N, Chymotrypsin). The obtained peptides are analysed by MS to obtain i) accurate mass of the intact peptides ii) peptide fragmentation ("CID" or "MS/MS") patterns which provide (partial) sequence information. Both the mass of "precursor" peptides and the fragmentation patterns are modified by the PTM and this theoretically allows determining the exact site of modification (Figure 1).

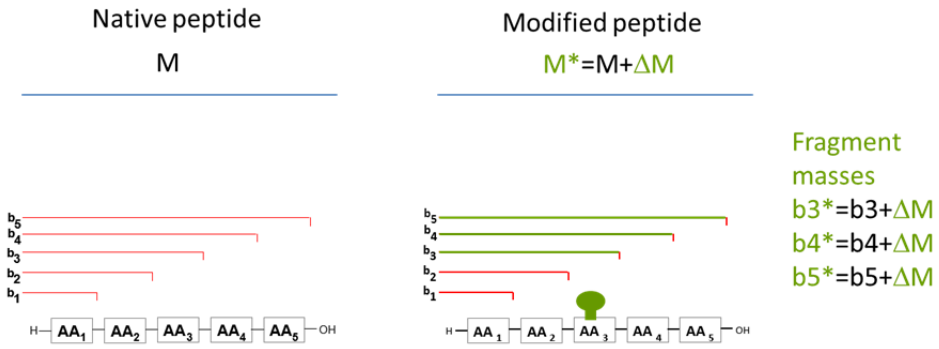

Figure 1. *A PTM induces a change in both the intact mass M of a peptide and the mass of CID fragments (here as an example the so-called b-fragments) generated upon Collision-Induced Dissociation (CID). Both types of information are then used to assign the modification to the peptide and, if possible, to localize the modified site in the sequence.*

In practice, however, things are not always this simple because of the complexity of the possible modification patterns and some variability in the quality of the data. To exploit and evaluate the obtained results it is essential to consider several facts.

1) the **AA sequence** is the primary determinant of the protein sequence coverage obtained with a certain enzyme. If cleavage sites are too far apart or too close to each other the generated peptides are very long or very short; in both cases their identification becomes difficult. When a stretch of sequence is not covered, no information can be gained on PTM's in this stretch.

2) for simplicity we can consider that the data analysis proceeds through the above mentioned steps:

1) identification of phospho-peptides and determination of number of modifications carried by each peptide

2) attempt to localize the modification site(s).

3) In most cases, **saying with a certain confidence that a peptide is phosphorylated is easier than locating the exact modification site** (obvious exception : when only one S,T,Y is present in the sequence). The software used for data analysis tries to assign a localization score based on the fragmentation pattern. However, especially when S,T,Y residues are close by in the sequence, the fragmentation information does often not allow the unambiguous assignment of the phospho-site to one residue (figure 2).

Score	Mr(calc)	Delta	Sequence	Site Analysis
62.6	1943.7884	0.0030	KETESEAEDNLDDLEK	Phospho S5 79.48%
56.8	1943.7884	0.0030	KETESEAEDNLDDLEK	Phospho T3 20.52%
11.5	1943.8054	-0.0140	EAYTHAQISRNNELK	
4.3	1943.8054	-0.0140	EAYTHAQISRNNELK	

Figure 2*. Mascot probability assignment for a phospho-peptide sequence containing two possible modification sites.*

4) **multi-phosphorylated peptides** (which do exist) are more difficult to detect than singly phosphorylated ones

5) for peptides containing >1 phosphorylatable residue, one should also consider the **possibility that the modification is spread on more than one residue**, due to the action of several kinases or a single promiscuous kinase. Two isoforms of a peptide which carry a phosphate at two different positions have the same MW and may co-elute in HPLC, thus giving rise to ambiguous information on site localization.

Quantitation (without isotope labeling)

6) MS information can allow to quantitate phosphopeptides, using their MS signal intensity. This is often performed by plotting the signal of the exact mass of the phospho-peptide as a function of time in the HPLC gradient (Figure 3). If only one [S,T,Y] is present or if site localization is clear this allows to infer quantitation for the site. If the localization of the site within the peptide is ambiguous, its quantification will be subjected to the same uncertainty.

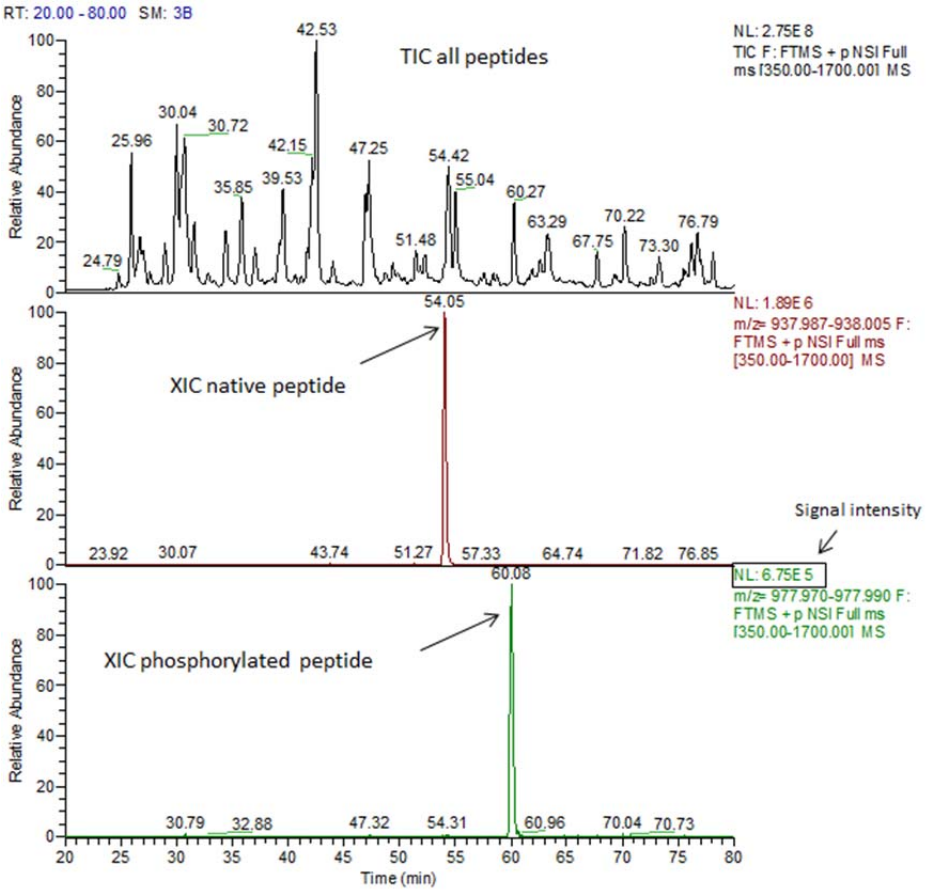

Figure 3*. Extracted ion current (XIC) of m/z corresponding to native and phosphorylated peptides are used for quantitation (TIC: total ion current).*

7) If two or more sites are identified in a phosphopeptide resulting in isoforms, these can be quantitated only if they are separated during the HPLC gradient.

8) a useful internal normalization of phospho-peptide quantitation can be carried out by using the signal intensity of the native, non-modified peptide. However one complication is that in some cases phosphorylation decreases the efficiency of protease cleavage of a given sequence.

9) due to a certain variability in the selection of peptide precursors for fragmentation, it can happen that a given phospho-peptide is present in low amounts in a sample but it is not identified by MS/MS. If the peptide was identified in a related sample, we try to detect and quantitate it also in those samples where it was not explicitly identified, using its accurate mass and HPLC elution time.

10) it is possible to do **enrichment** for phospho-peptides by various means (IMAC, TiO2, Antibodies). Although this is applied widely for large scale studies, it may not be essential when studying a single protein and indeed we only use it as a second possible strategy after a direct analysis. Among the drawbacks of enrichment are the loss of the unmodified peptides and thus the loss of quantitation references.

Quantitation (with isotope labeling)

11) For large-scale experiments or when accurate quantitation of changes is required, it is possible to couple analysis of PTM's with stable isotope labeling using various strategies (SILAC, iTRAQ, ..). These require a dedicated workflow and are not discussed here.