Phosphoproteomics

Francesco Lonardoni and Alessandra Maria Bossi
Verona University,
Italy

1. Introduction

Phosphorylation is the most widespread and studied Post Translational Modification (PTM) in proteins [Collins et al. 2007]. It is involved in almost all cell functions: metabolism, osmoregulation, transcription, translation, cell cycle progression, cytoskeletal rearrangement, cell movement, apoptosis, differentiation, regulation of the signal transduction pathways, intercellular communication during the development and functioning of the nervous system [Graves & Krebs, 1999; Hunter, 2000; Sickmann & Mayer, 2001].

The phosphorylation/dephosphorylation process is regulated by the switch kinases/phosphatases. Kinases add a phosphate group to a receptive side chain of an aminoacid; phosphatases catalyze instead the hydrolysis of a phosphoester bond [Raggiaschi et al., 2005; Thingholm et al., 2009a]. The effect of the addition or subtraction of a phosphate group is the modification of enzymatic activity, protein-protein interaction and cellular localization.

Phosphorylation is not an unique process: often a single protein can display more than a single site suitable for the process, often catalyzed by different kinases. For example, glycogen synthase contains at least 9 phosphorylation sites, and its modulation is performed by at least 5 protein kinases acting on different sites of the protein [Nelson & Cox, 2004].

A misregulation of the phosphorylation processes can cause severe damage to the cells, leading to diseases like cancer, diabetes or neurodegeneration [Clevenger, 2004; Zhu et al., 2002].

The most common kind of phosphorylation in eukaryotes is O-phosphorylation, on serine, threonine and tyrosine with a ratio of 1800/200/1 [Grønborg et al., 2002; Kersten et al., 2006]. Other sites of phosphorylation can be histidine, lysine, arginine, glutamic acid, aspartic acid and cysteine [Sickmann & Mayer, 2001], even though are less studied due to the lability of the chemical bond and the subsequent necessity to use very special techniques to analyse them.

It is esteemed that 2-4% of eukaryotic genes are associated with kinases and phosphatases (there are about 500 kinase and 100 phosphatase genes in the human genome) [Manning et al., 2002; Twyman, 2004; Venter et al., 2001]. Around 100,000 phosphorylation sites may exist in the human proteome, the majority of which are presently unknown [H. Zhang et al., 2002]. The importance of studying the phosphorylation was marked by the success of the
cancer drug Gleevec, the first to inhibit a specific kinase, which gave definitely an impulse to the research on kinases and their substrates as potential drug targets [Manning et al., 2002].

The comprehensive analysis of the protein phosphorylation should include: identification of phosphorylated proteins and of their sites of phosphorylation, how these phosphorylations modify the biological activity of the protein and kinases and phosphatases involved in the process.

2. A delicate analysis

Working in phosphoproteomics presents a series of hurdles in the analytical strategy.

The first issue concerns the reversible nature of the phosphorylation. The study of the phosphoproteins necessitates their isolation from a cell extract or a sub-cellular compartment. Subsequently to the cell lysis, however, many enzymes like phosphatases become active, determining the degradation of the proteins and detachment of phosphate groups from their sites. Also kinases can express their action, confusing the picture of which phosphate groups are biologically relevant [Raggiaschi et al., 2005]. Working at low temperature helps significantly in slowing down these processes, but it’s not enough. In order to stop the action of these enzymes it is essential to add to the cell extracts a specific mix of inhibitors of proteases and phosphatases [Hemmings, 1996; Reinders & Sickmann, 2005; Schmidt et al., 2007; Thingholm et al., 2008a], while to inhibit kinases EDTA, EGTA or kinase inhibitors are added. It is also important to choose an inhibition mix that doesn’t interfere with the downstream analytical methods, like the phospho-specific enrichment methods.

Another issue relates with phosphoproteins characterization, mostly performed through Mass Spectrometry (MS) methods after enzymatic digesting. The detection of phosphopeptides (FPs from now on) with MS is hampered by the presence of the non phosphorylated partners; moreover, the efficiency of ionization is higher for the latter ones, also generally more present in the sample (this fact is referred to as “low stoichiometry” of phosphorylation).

It follows that enrichment methods are necessary to extract the phosphoproteins or the FPs from the sample. There are various methods to choose from, depending to the kind of sample and the aims of the study [Kalume et al., 2003; Mann et al., 2002].

3. Detection of phosphoproteins

The detection of phosphoproteins in a sample still relies on optimized “classical” methods.

3.1 Isotopic labeling of phosphoproteins

One of the oldest methods used to study phosphoproteins is the metabolic labeling with $^{32}$P and $^{33}$P. It consists in nourishing living cells or organisms with substances labeled with these radioactive isotopes which are incorporated in the synthesised proteins. Following lysis of the cells, the protein population is isolated through 1-DE or 2-DE, visualized on the gels with autoradiography or acquired digitally with Phosphorimagery systems. This method is still largely employed, because of its simplicity and reliability when somebody works with alive systems in vitro or in vivo [Eymann et al., 2007; Su et al., 2007].
A comparison between the performances of $^{32}\text{P}$ and $^{33}\text{P}$ in labeling the proteins was made [Guy et al., 1994], indicating that $^{33}\text{P}$ gives more neat image and higher resolution, even though after a longer exposition time, respect $^{32}\text{P}$.

Apart from the safety and environmental implications in using radioactive isotopes, this method has other drawbacks. First of all it is not compatible with some downstream methods, like MS. Furthermore it can only be applied on viable cells, since the radioactive isotopes have to be taken from the media and metabolized: it cannot be therefore applied on post-mortem tissues or biopsies.

In *in vivo* studies, cells are incubated with $^{32}\text{P}$, however the presence of ATP reservoirs inside the cells can interfere with the labeling, reducing the efficiency of the method [Steen et al., 2002]. Furthermore, $^{32}\text{P}$ is toxic for the cells, and over time can cause damages [Hu & Heikka, 2000; Hu et al., 2001; Yeargin & Haas, 1995].

In *in vitro* studies, proteins are incubated with specific kinases in the presence of $[^\gamma-^{32}\text{P}]-\text{ATP}$ and, under specific conditions, the radioactive atom is incorporated into aminoacidic residues. Due to the unnatural presence of kinase respect the target protein, however, it is frequent the phosphorylation of a different target instead of the natural one (promiscuity) [Graham et al., 2007].

### 3.2 Western blotting employing phosphospecific antibodies

Western blotting is a quite old technique. It is based on the selective binding of an antibody to a protein, transferred from a 1D or 2D gel to a nitrocellulose or polyvinylidene difluoride (PVDF) membrane support, and the subsequent revealing of the antibody marked spot with some visual method [Magi et al., 1999; Towbin et al., 1979]. The key role is played by the antibody, that should be specific for the protein epitope of interest: in this case epitopes are phosphoserine, phosphothreonine and phosphotyrosine. The selectivity and affinity characteristics of the antibody are of major importance, to perform specific recognition and limit false positives. While excellent anti-phosphotyrosine antibodies have been developed (e.g. (PY)20, (PY)100 and 4G10 hybridoma clones), better antibodies are still needed for phosphoserine and phosphothreonine. Antibodies generated against $\text{pSer}$ and $\text{pThr}$, in fact, very often necessitate of a consensus sequence flanking the phosphoaminoacid; this might be due to the lower immunogenicity of $\text{pSer}$/$\text{pThr}$ compared to $\text{pTyr}$ [Schmidt et al., 2007].

Grønborg et al. performed a test for specificity and reliability of anti-phosphoserine and anti-phosphothreonine antibodies [Grønborg et al., 2002]. They made a large scale differential analysis of phosphorylated proteins, succeeding in identifying phosphorylation sites and FPs not identified with dedicated prediction software. The combination of high resolution 2-DE techniques and the Enhanced Chemiluminescence (ECL) system give improved sensitivity to the method, i.e. intensification of around 1000 times of the light emitted from a spot [Buonocore et al., 1999; Kaufmann et al., 2001].

Although Western blotting is an efficient technique to reveal even small amount of protein ($10^{-10}$ mol), its use in quantitative analysis is limited by the variability of the amount of proteins that can be transferred to the membrane.
3.3 Direct staining of phosphoproteins

The easiest way to detect phosphoproteins in a sample is the direct staining with a phosphospecific dye after a SDS-PAGE gel. Many attempts have been done from the 1970’s [Steinberg et al., 2003], but all these methods face problems in terms of sensibility and of non specificity, e.g. the inability to discriminate between phosphorylated and non phosphorylated proteins or to detect phosphotyrosine. Quite recently a novel fluorescent dye has been introduced: Pro-Q Diamond [Schulenberg et al., 2003]. This dye selectively binds to phosphoproteins requiring a very simple experimental protocol. The sensitivity of the stain depends on the number of phosphorylated residues of the proteins: the detection limit was 16 ng for pepsin (1 phosphorylated residue) and 2 ng for casein (8 phosphorylated residues) [Schulenberg et al., 2003]. Thus, the method is very useful for a preliminary screening, but still not sufficient for a comprehensive analysis of the phosphoproteome.

The dye is also compatible with MS and with Multiplex Proteomics (MP), i.e. detecting simultaneously phosphoproteins and total proteins (e.g. these latter stained with SYPRO Ruby dye) on the same 2D gel. The combination of the two staining methods permits to distinguish low represented but highly phosphorylated proteins from highly represented but poorly phosphorylated ones, comparing the results from the two different colorations.

3.4 Detection of phosphoproteins employing protein phosphatases

The presence or absence of a phosphate group on a protein is enough to change its pI and subsequently its position in the 1st dimension of a 2D gel. That’s why, by employing a phosphatase enzyme, it is possible to modify the position of a protein spot on a map and thus determine its nature comparing the maps before and after the treatment. Softwares have been also developed that can predict the pI shift due to the addition/removal of a phosphate group, that can be of 1-2 pH units [Kumar et al., 2004]. As an example, Yamagata et al. exploited the specific enzymatic activity of k-phosphatase (kPPase) on phosphoserine, phosphothreonine, phosphotyrosine and phosphohistidine residues to identify novel phosphoproteins in cultured rat fibroblasts [Yamagata et al., 2002].

The methods employing phosphatases, however, are not suitable for quantitative analysis, mainly because of the complexity of the analysis and the variable efficacy of the enzymatic action.

4. Selective enrichment of phosphoproteins and FPs

The identification of the PTM sites on a protein is generally performed by using MS approaches. On most occasions, the only enrichment of the sample in phosphoproteins followed by protease-specific digestion and MS analysis is not sufficient to identify the sites of phosphorylation present (due to the general low stoichiometry of the phosphorylation), thus a second enrichment step at the FP level is often also required.

Some commercial kits for phosphoprotein and FP enrichment are available, offering ease of use and reproducibility; nevertheless it has been clearly demonstrated that the different methods available differ in the specificity of isolation and in the set of phosphoproteins and FPs isolated [Bodenmiller et al., 2007], strongly suggesting that no single method is sufficient for a comprehensive phosphoproteome analysis. Strategies for phospho-specific enrichment are shown in fig.1.
Fig. 1. Selective enrichment of phosphoproteins and phosphopeptides.
A) Immunoprecipitation: phosphoproteins or phosphopeptides are selectively precipitated through the use of appropriate anti-phospho antibodies; at the moment only the use of anti-pTyr antibodies has proven to be robust. B) Affinity chromatography: a resin with immobilized chelated metal or TiO$_2$ can selectively bind the phosphoric group of peptides and also proteins in IMAC (par. 4.2). A combined approach is SIMAC (IMAC + TiO$_2$). An alternative technique could be the use of Molecularly Imprinted Polymers (MIPs, par.4.10). C) Chemical derivatization: the phosphoric group reacts with the aminogroup of a tag in PAC or can be subdued to β-elimination and linking of a suitable tag through Michael addition (par. 4.11).

4.1 Phosphoprotein enrichment by Immunoprecipitation

Phospho-specific antibodies can be used to selectively immunoprecipitate phosphorylated proteins depending on the specificity of the antibody. As for Western blotting (see above) anti-phosphotyrosine antibodies are the most reliably and widely used in order to enrich tyrosine-phosphorylated proteins from complex mixtures.

After immunoprecipitation (IP) the phosphotyrosine enriched sample can be analyzed with different analytical methods such as 1-DE and 2-DE [Blagoev et al., 2004; Stannard et al., 2003].

Also in this case, variations of protein phosphorylation levels are very difficult to characterize unless in combination with a particular protein labeling (Stable Isotope Labeling with Aminoacids in Cell culture: SILAC, see par.6.1) with stable isotopes ($^{13}$C or $^{15}$N).
Protein Purification

$^{15}$N) is used [Ong et al., 2002]. This strategy allowed a quantitative and temporal investigation of tyrosine phosphorylation events of proteins involved in signaling pathways after stimulation with Epidermal Growth Factor (EGF) [Blagoev et al., 2004].

4.2 Phosphopeptide and phosphoprotein enrichment using Immobilized Metal Affinity Chromatography (IMAC)

IMAC exploits the material formed through chelation of a di-, tri- or tetravalent metal by nitrilotriacetic acid (NTA), iminodiacetic acid (IDA) or Tris(carboxymethyl)ethylenediamine (TED) immobilized on a solid support, like porous silica beads [Porath et al., 1975]. The most commonly used resins make use of Fe$^{3+}$, Ga$^{3+}$ and Al$^{3+}$, even though Zn$^{2+}$ and Zr$^{4+}$ are used as well [Feng et al., 2007].

This method is routinely employed in FPs enrichment prior to MS analysis, nevertheless it has some limits; the most evident is its undesired ability to bind acidic peptides. This limitation has been largely surpassed by the acidification of the media (to protonate the carboxylic groups) [Posewitz & Tempst, 1999] and esterification of the carboxylic moieties with methanolic HCl before the enrichment step, even if this method introduces complexity due to the variable yield of methylation [Ficarro et al., 2002].

A second limitation is the net bias of the method towards monophosphorylated peptides [Ficarro et al., 2002].

The method is particularly effective when used in combination with an enrichment step at the protein level. This operation can be carried out with methods like IMAC itself [Collins et al., 2005], exploiting however a more suitable solid support, like Sepharose beads. Moreover, secondary interactions have to be reduced, e.g. with the use of denaturing conditions (6M urea). The detachment of the potentially many phosphate groups from the column imposes instead the use of strong eluting buffer as 0.1-0.2 M EDTA [Collins et al., 2005].

Other phosphoprotein enrichment methods are phosphotyrosine IP [Ficarro et al., 2003], Strong Anion eXchange chromatography (SAX) [Trinidad et al., 2006], Strong Cation eXchange chromatography (SCX) [Nühse et al., 2003] and SDS-PAGE [Villen et al., 2007].

4.3 Metal Oxide Affinity Chromatography (MOAC)

Metal oxide chromatography (MOAC) employs mainly Ti, Zr or Al oxides, in the form of solid or coated beads, as chromatographic media to sequestrate FPs. Many different crystalline forms and nanostructured materials have been devised [Leitner, 2010].

The first report about the potentialities of these materials regarded the use of TiO$_2$-based columns to sequestrate phosphate ions from the water [Connor & McQuillan, 1999; Ikeguchi & Nakamura, 2000].

In 2004, Pinkse et al. published a paper on the use of this material to bind FPs. They devised a 2D-LC-MS online strategy with TiO$_2$ beads (Titansphere) as first dimension and RP as second one [Pinkse et al., 2004]. Acidic conditions (pH 2.9) promoted the adhesion of FPs to the first column, leaving the non-phosphorylated ones to flow through and to be analysed with nano-LC-RP- MS/MS. Basic conditions (pH 9.0) eluted FPs in a second step. The method was tested on a 153kDa homo-dimeric cGMP-dependent kinase, promoting the discovery a total of 8 phosphosites, 2 of which were novel.
Larsen’s group then devised an off-line strategy to bind FPs to a TiO$_2$ material. The use of additives as 2,5- dihydroxybenzoic acid (DHB), phthalic acid or glycolic acid largely reduced the aspecific attachment of acidic peptides [Jensen & Larsen, 2007; Larsen et al., 2005; Thingholm et al., 2006]. This technique has been named HAMMOC, for hydroxy acid modified metal oxide chromatography.

A significative advantage of TiO$_2$ is its tolerance towards most buffers and salts used in biochemistry and cell biology laboratories [Jensen & Larsen, 2007]. This is one of the reasons which made TiO$_2$ so common in large scale phosphoproteomics studies [Dengjel et al., 2007; Molina et al., 2007; Olsen et al., 2006; Olsen et al., 2007; Thingholm et al., 2008a].

Not only titania has been employed as metal oxide for FPs enrichment. Natural substitutes can be oxides belonging to the same group. For example, ZrO$_2$ microtips have been recently introduced as mean for FPs enrichment. This oxide shows a preference towards singly phosphorylated FPs, while TiO$_2$ preferentially retains multiply phosphorylated ones [Kweon & Hakansson, 2006]. However, subsequent large-scale studies demonstrated also a lower selectivity versus acidic peptides [Sugiyama et al., 2008], suggesting the necessity of further improvements.

4.4 Sequential elution from IMAC (SIMAC)

The identification of multiply phosphorylated peptides has proven to be a hard task, mostly because of their difficult ionization and subsequent low signal compared to singly- and not-FPs, so to be not selected for the subsequent fragmentation in the mass spectrometer.

To address this issue, in 2007 Martin Larsen’s group presented an analytical strategy for sequential elution of mono- and multiphosphorylated peptides. The sequential elution from IMAC (SIMAC) exploits the complementary characteristics of IMAC and TiO$_2$ in enriching the sample respectively in multiply and singly phosphorylated peptides [Thingholm et al., 2008b].

In particular, the elution from IMAC in acidic conditions (pH 1.0) enriches the sample in mono-FPs, while the basic conditions (pH 11.3) elute subsequently the multi-FPs [Thingholm et al., 2008a]. A further enrichment with TiO$_2$ chromatography is needed only in the acidic fraction, because the basic one is enough depleted of non-FPs.

The separation of singly and multiply FPs permits then their analysis with pdMS$^3$ (phosphorylated directed fragmentation) in optimized settings for each group [Raggiaschi et al., 2005; Thingholm et al., 2008b].

The method was tested on a 120µg whole-cell extract from human mesenchimal stem cells (hMSCs) and the results compared with those from TiO$_2$ enrichment alone. A total of 716 phosphosites was identified with SIMAC, while 350 with TiO$_2$. Moreover the number of multiply phosphorylated peptides was significantly increased [Thingholm et al., 2008b].

Recently, a new intriguing method for multiply phosphorylated peptides enrichment based on polyarginine-coated diamond nanoparticles was presented, however it has still to be tested on large scale samples and where a low amount of starting material (micrograms) is available [Chang et al., 2008].
4.5 Magnetic beads

Starting from the TiO$_2$-based chemistry with the idea to find an easier way to perform the extraction of FPs, Chen and Chen [Chen & Chen, 2005] thought to conjugate the properties of magnetic materials with those of TiO$_2$, coating Fe$_3$O$_4$ nanoparticles with TiO$_2$ through a silanic bridge. The nanobeads are mixed with a trypsic digest, vortexed and captured with a magnet. The captured FPs are subsequently analysed through a laser desorption/ionization from the inorganic particles and a run in MS. The method was named SALDI-MS, from Surface Assisted Laser Desorption/Ionization Mass Spectrometry [Schürenberg et al., 1999; Sunner et al., 1995].

There were subsequent improvements of the method, using for example Fe$_3$O$_4@C@SnO$_2$ core-shell microspheres (the symbol @ means “coated by”), with which scientists were able to detect 77 phosphorylation sites in rat liver cells [Qi et al., 2009].

4.6 Calcium Phosphate Precipitation (CPP)

In 1994, Reynolds et al. presented a strategy for FPs enrichment through Ca$^{2+}$ ions and 50% ethanol precipitation. The precipitated peptides from a trypsic digest of casein mostly contained multiple phosphoserines [Reynolds et al., 1994].

Zhang et al. tested the strategy by using calcium chloride (CaCl$_2$), ammonia solution (NH$_3$.H$_2$O) and disodium phosphate (Na$_2$HPO$_4$) on a rice embryo preparation [X. Zhang et al., 2007]. The dissolved and desalted pellet was then enriched through IMAC. In total, 227 non-redundant phosphorylation sites were identified, of which 213 on serine residues and 14 on threonine.

Through phosphate precipitation directly coupled to LC-MS/MS, Qiangwei Xia et al. identified 466 unique phosphorylation sites (379 on serine and 87 on threonine) in post-mortem Alzheimer disease brain tissue, 70% of which were not reported in Phospho.ELM database [Xia et al., 2008].

In both studies no tyrosine phosphorylated peptides were identified: it is not clear if this is due to the low abundance of these or to the poor selectivity of the method. Anyway the method offers the advantage of being rather simple, column-free and straightforward.

4.7 Ion exchange chromatography (IC)

The simple enrichment of FPs with IMAC, TiO$_2$ or SIMAC has generally proven to be not enough productive when a deep knowledge of the phosphoproteome is required [Rigbolt et al., 2011]. A fractionation step is also needed. Chromatography techniques based on charge interaction, hydrophilic interaction or a combination of both have been employed for this purpose [Zarei et al., 2011].

Anion exchange chromatography exploits the generally higher affinity of FPs for the positively charged stationary phase due to the intrinsic high negative charge carried by the phosphate group. Strong anion chromatography (SAX) has been employed both as fractionating technique before a FPs enrichment step (e.g. with IMAC or TiO$_2$) [Nühse et al., 2003; K. Zhang, 2006] and also as sole fractionating technique before LC-MS/MS.
<table>
<thead>
<tr>
<th>Method</th>
<th>Target</th>
<th>Sample type and amount</th>
<th>Strategy</th>
<th>Results</th>
<th>Reference</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMAC (Immobilized Metal Affinity Chromatography)</td>
<td>protein, peptide</td>
<td>H1 stem cells proteins, 10 mg</td>
<td>SCX-IMAC-RPLC-MS² (ETD-OT)</td>
<td>10844 phosphites at 1% FDR</td>
<td>Swaney et al., 2009</td>
<td>Proven to be effective in large scale analysis.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D. melanogaster lysate Kc167, 10 mg</td>
<td>SCX-IMAC-RPLC-MS² (LTQ-OT)</td>
<td>13720 phosphites at 0.63% FDR</td>
<td>Zhai et al., 2008</td>
<td>Limited capacity and specificity, directed towards multiply phosphorylated peptides, affinity for acidic peptides.</td>
</tr>
<tr>
<td>MOAC (Metal Oxide Affinity Chromatography)</td>
<td>peptide</td>
<td>HeLa lysate (amount not given)</td>
<td>SCX-TiO₂-RP-MS² &amp; MS³ (LIT-FT-ICR)</td>
<td>6600 phosphites, 2244 proteins at &lt;1% FPR</td>
<td>Olsen et al., 2006</td>
<td>Robust and chemically inert material, high capacity and fast absorption.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>K562 lysate, 400 µg</td>
<td>SCX-TiO₂/Nb₂O₅-RP-MS² (MALDI-TOF)</td>
<td>622/642/834 phospho peptides (Ti/Nb/all*) at 4% FPR</td>
<td>Ficarro et al., 2008</td>
<td>Slow flow desorption, most effective for singly phosphorylated peptides.</td>
</tr>
<tr>
<td>SCX (Strong Cation Exchange)</td>
<td>peptide</td>
<td>HeLa cell lysate, 300 µg</td>
<td>SCX-RP-LC-MS² &amp; MS³ (IT)</td>
<td>2002 phosphites</td>
<td>Beausoleil et al., 2004</td>
<td>Coelution with other acidic peptides.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HEK 293 cell lysate, 1 mg</td>
<td>SCX-RP- LC-MS² (ETD,LIT)</td>
<td>&gt;5000 unique phosphopeptides 1% FDR</td>
<td>Mohammed &amp; Heck, 2011</td>
<td>Coelution with other acidic peptides.</td>
</tr>
</tbody>
</table>

www.intechopen.com
<table>
<thead>
<tr>
<th>Method</th>
<th>Target</th>
<th>Sample type and amount</th>
<th>Strategy</th>
<th>Results</th>
<th>Reference</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAX (Strong Anion Exchange)</td>
<td>peptide</td>
<td>Human liver protein digest, 100 µg</td>
<td>SAX-RPLC-MS²&amp;MS³ (LTQ)</td>
<td>274 phosphoites at 0.96% FDR</td>
<td>Guanghui et al., 2008</td>
<td>Useful for peptide fractionation Useable as pre-enrichment technique</td>
</tr>
<tr>
<td>Ion Exchange Chromatography.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphopeptides more retained by stationary phase due to their higher negative charge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HILIC (Hydrophilic Interaction Chromatography)</td>
<td>peptide</td>
<td>HeLa cell lysate, 1 mg</td>
<td>HILIC-RP-MS &amp; MS²(IT)</td>
<td>1004 phosphoites</td>
<td>McNulty &amp; Annan, 2008</td>
<td>Coelution with other acidic peptides. Useable as pre-enrichment technique</td>
</tr>
<tr>
<td>Phosphopeptides more retained by stationary phase due to their higher polarity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemical derivatisation</td>
<td>peptide</td>
<td>D.melanogaster Kc167 tryptic digest, 1.5 mg</td>
<td>PAC-RP-LC-MS²</td>
<td>535 phosphoites</td>
<td>Bodenmiller et al., 2007a</td>
<td>Modulation of peptide properties for MS purposes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Jurkat T cells lysate (amount not reported)</td>
<td>PAC-RP-LC-MS² (LTQ)</td>
<td>79 tyrosine phospho proteins</td>
<td>Tao et al., 2005</td>
<td>Reaction yield, side reactions, large amount of sample required.</td>
</tr>
<tr>
<td>Method</td>
<td>Target</td>
<td>Sample type and amount</td>
<td>Strategy</td>
<td>Results</td>
<td>Reference</td>
<td>Comments</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------------------------</td>
<td>------------------------</td>
<td>-----------------------------------</td>
<td>------------------------</td>
<td>----------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Immuno-precipitation</td>
<td>protein, peptide</td>
<td>Jurkat T cells lysate</td>
<td>IP(pY-100 antibody) -RP-LC-MS²</td>
<td>194 phosphotyrosine sites</td>
<td>Rush et al., 2005</td>
<td>Mostly directed to p-Tyr phosphosites</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(amount not reported)</td>
<td>(LCQ-IT)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IP (pY-100 antibody) -RP-LC-MS²</td>
<td>481 phosphotyrosine sites at FDR 1%</td>
<td>Heibeck et al., 2009</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HME Cells lysate, 4mg</td>
<td>(LTQ-OT)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*N*superimposition without redundancies of Ti and Zr detected phosphosites.

Table 1. Comparation of enrichment and fractionation methods for phosphopeptides.

Abbreviations: CID, Collision-Induced Dissociation; DHB, Dihydroxybenzoic acid; ECD, Electron Capture Dissociation; ETD, Electron Transfer Dissociation; FDR, False Discovery Rate; FPR, False positive rate; FT-ICR, Fourier Transform Ion Cyclotron Resonance; b-HPA, b-hydroxypropanoic acid; LIT, Linear Ion Trap; OT, Orbitrap; Q-TOF, Quadrupole-Time-Of-Flight; RP, Reversed Phase; SCX, Strong Cation Exchange; LTQ, Linear Trap Quadrupole; LCQ, Liquid Chromatography Quadrupole; IP, Immunoaffinity Purification.

Nühse et al. for example identified more than 300 phosphorylation sites in the plasma membrane fractions of *Arabidopsis thaliana* using a SAX + IMAC approach [Nühse et al., 2004], while Han and co-workers identified 274 phosphorylation sites in human liver tissue extract without the enrichment step [Han et al., 2008].

A drawback of this method has been reported by Thingholm et al. [Thingholm et al., 2009a], who noted a strong attachment of FPs to the SAX resin, from which they are difficultly recovered.

Its counterpart, strong cation chromatography (SCX), has been largely employed as prefractioning technique for proteins and peptides. The pioneering work on FPs was carried out by Beausoleil et al. in 2004 [Beausoleil et al., 2004]. Tryptic peptides were acidified at pH 2.7, where most of the peptides carry a +2 charge due to C-terminal lysine or arginine and the N-terminal aminogroup. FPs, instead, carry a reduced charge due to the phosphate group, e.g. +1 in monoFPs. The reduced affinity of the resin should therefore leaving FPs to flow more easily through the column. This idea was confirmed by the results, which brought to the identification of 2000 phosphosites in 8mg of nuclear extract of HeLa cells lysate.
Trinidad and co-workers evaluated the efficiency of SCX as prefractionation method prior to IMAC to the efficiency of IMAC and SCX alone, finding a three-fold increased FPs identification in the combined approach respect both methods [Trinidad et al., 2006].

The strength of SCX as prefractioning system was further confirmed by Olsen et al., who identified 6600 phosphosites in 2244 proteins in EGF-stimulated HeLa cells through SCX + TiO$_2$ [Olsen et al., 2006]. In more recent reports 23000 and 36000 phosphorylation sites have been identified respectively with SCX followed by IMAC and TiO$_2$ respectively [Huttlin et al., 2010; Rigbolt et al., 2011]. This remarkable increase in number of detected phosphosites is mainly due to instrumental and software improvements [Zarei et al., 2011].

It remains to be assessed if the fractionation/enrichment approach can be suitable in samples where a small amount of starting material is available.

### 4.8 Hydrophilic interaction chromatography

Hydrophilic interaction chromatography (HILIC) is a separation technique for biomolecules [Alpert, 1990]. The method relies on the interaction of the analytes, e.g. peptides, with a neutral stationary phase through hydrogen bonding. They are eluted from the column with a decreasingly organic mobile phase according to their polarities (hydrophilicity).

McNulty and Annan introduced this method as prefractionation step before IMAC [McNulty & Annan, 2008]. They tested the FPs enrichment capacity of HILIC compared to SCX and IMAC alone, as well as the fractionation ability of HILIC before and after IMAC treatment. The analysis of 1 mg of tryptic digest of HeLa cells evidenced the prefractionation vocaction of the method. The use of IMAC prior to HILIC gave in fact a high level of nonspecific binding due to not FPs, while the reversed approach improved the selectivity for FPs to more than 99%.

ERLIC (Electrostatic Repulsion - Hydrophilic Interaction Chromatography), or Ion Pair normal phase, is a newly developed chromatographic method able to enrich and fractionate FPs in a single step [Alpert, 2008].

At pH 2.0 the C-termini and the carboxylic side chains of Asp and Glu are neutral, thus the peptides are generally positively charged and are not retained by a positively charged stationary phase. The presence of a phosphate group, however, reduces the net charge of FPs; this is not enough to overcome the overall repulsion for the stationary phase due to the basic residues, however an organic mobile phase, e.g. acetonitrile 70%, promotes their hydrophilic interaction with the column. The attraction for the column is enhanced in multiply phosphorylated peptides.

Zarei et al. evidenced in fact a higher efficiency of ERLIC compared to SCX and HILIC in fractionating multiply phosphorylated peptides [Zarei et al., 2011], more retained by the stationary phase.

In a recent article [Chen et al., 2011] Xi Chen et al. compared the FPs profiles obtained from a HeLa cell lysate by using 4 HPLC methods after a phospho-enrichment with IMAC. Even though with any of the four methods (SCX, HILIC, ERLIC with volatile and not volatile solvent) 4-5000 peptides could be identified, the combination of all the methods gave a total number of 9069 unique FPs, with a considerable amount of non-overlapping unique FPs. The four methods are thus complementary to get a full coverage of the phosphoproteome.
4.9 Hydroxyapatite (HAP)

Firstly introduced by Tiselius and co-workers in 1956 [Tiselius et al., 1956], hydroxyapatite (HAP) chromatography has been very popular until the 1980s, when new matrices have been introduced. HAP is a crystalline form of calcium phosphate with formula $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$, which binds FP in virtue of both anionic and cathionic interactions. In particular $\text{Ca}^{2+}$ binds phosphoric groups, and more strongly than it does with other acidic groups. The binding to the matrix is thus proportional to the number of phosphogroups on the peptides and can be exploited for a sequential elution of them according to their degree of phosphorylation. Mamone et al. [Mamone et al., 2010] exploited this material to analyse a tryptic digest of standard proteins. The advantages of HAP are the low cost of the material and the possibility to elute stepwise the peptides according to the degree of phosphorylation. Yet the system has to be tested on more complex biological samples.

4.10 Molecularly Imprinted Polymers (MIPs)

Recently [Emgenbroich et al., 2008; Helling et al., 2011], Borje Sellergren’s group has proposed a method for the depletion of phosphoryl-containing peptides based on Molecularly Imprinted Polymers (MIPs). FPs imprinting was performed by an epitope approach [Nishino et al., 2006; Rachkov & Minoura, 2001; Titirici et al., 2003], i.e. using a part of the analyte of interest (in this case the N- and C- protected phosphorylserine) as a template to prepare a MIP able to fish FPs. The Solid Phase Extraction analysis showed a 18-fold preferential retention for phosphorylated angiotensin II respect the non phosphorylated one, preference not shown in the not imprinted polymer. Moreover, the material showed a preference for a triply phosphorylated peptide over a monophosphorylated one, showing opposite behaviour comparing to $\text{TiO}_2$ material. Of course a more extensive analysis and improvements are needed in order to tackle more complex samples.

4.11 Chemical derivatisation strategies

The chemical derivatisation strategies exploit the typical properties of the phosphate groups in the peptides, like the lability of the phosphoesteric bond and the subsequent easy substitution of the phosphate with a nucleophilic tag.

A different widely used approach is the methyl esterification of the carboxylic groups, notoriously competitive with the phosphate groups due to their negative charge (fig.2).

4.11.1 Methyl esterification of carboxyl-groups

In IMAC strategy carboxylic groups influence the elution of FPs, due to their charge attracted by the positively charged metal of the resin. For this reason their esterification could prove useful. Ficarro et al. first applied this strategy for the analysis of the cell extract from $S$.Cerevisiae [Ficarro et al., 2002] through IMAC approach: 333 sites of phosphorylation were detected with this method. Some characteristics of this approach have already been depicted in par. 4.2 and will be no further analysed.

Methyl esterification can be also exploited as a mean of isotopic labeling (with CH$_3$OH and CD$_3$OH for two different cellular states) for quantification purposes [Ficarro et al., 2003].
4.11.2 Biotin tagging by β-elimination and Michael addition

Many chemical derivatisation strategies have been devised to displace the phosphoryl group and binding a tag to the “naked” peptide. One of these methods [Jaffe et al., 1998] relies on the β-elimination of the phosphate from phosphoserine and phosphothreonine producing dehydroalanine and β-aminobutyric acid respectively. These products can be directly detected using tandem MS or derivatised, for example by Michael addition of a reactive thiol and subsequent binding of a tag. One common tag is biotin, notoriously having great affinity for avidin, liable of immobilization to an affinity column [Oda et al., 2001].

This technique shows the limits of not being applicable to phosphotyrosine containing peptides and of the occurrence of side reactions like tagging of non phosphorylated serine.

Zhou et al. [Zhou et al., 2001] proposed another method of derivatisation applicable to all residues, even if it involves many steps and it has not been tested on complex samples yet.

A) Methyl esterification of acidic groups

![Diagram A](https://www.intechopen.com)

B) β-elimination / Michael addition

![Diagram B](https://www.intechopen.com)

C) Phosphoramidate derivatisation

![Diagram C](https://www.intechopen.com)

D) Conversion to aminoethylcysteine

![Diagram D](https://www.intechopen.com)

Fig. 2. Chemical derivatisation methods for phosphorylation capture and analysis.
A. IMAC suffers of aspecific binding of acidic peptides to the resin. Esterification of carboxylic groups with Methanol/HCl is a strategy to overcome this problem.

B. In presence of a strong base the β-elimination at phosphoserine and phosphothreonine produces dehydroalanine and β-aminobutyric acid. These can be derivatised with a thiol through a Michael-like addition and a tag can be added. If the tag is biotin, its affinity with avidin can be exploited.

C. FPs can be derivatised to phosphoramidates after esterification of carboxylic groups. If a dendrimer is employed the derivatised peptides can be separated through size-selective methods.

D. β-elimination and Michael-like addition of cysteamine converts phosphoserine and phosphothreonine in lysine analogues that specific enzymes leave in the C-terminus.

4.11.3 Phosphoramidate conversion

In 2005 Aebersold’s group proposed a derivatisation procedure based on the carboxyl protection through methyl esterification followed by conjugation to a soluble polymer with phosphoramidate chemistry (PAC) [Tao et al., 2005].

The mixture of peptides is first converted to the corresponding methyl esters. In this step two cellular states can also be differentially labeled for quantitative analysis. Subsequently, the methylated peptides are combined and put to react with EDC, imidazole and a polyamine dendrimer. Phosphopeptides are converted in the corresponding phosphoramidates, easily separated from the non FPs through size selective methods. FPs are recovered with a brief acid hydrolysis and sent to the MS analysis.

When coupled with pervanadate stimulation and an initial antiphosphotyrosine protein precipitation step, this method allowed the identification and quantification of all known plus previously unknown phosphorylation sites in 97 tyrosine proteins in Jurkat T cells.

A modification of this method was proposed [Bodenmiller et al., 2007], exploiting the reaction of the phosphate groups with cystamine and a reducing agent instead of the dendrimer. The –SH group of cystamine reacts with maleimide-activated glass beads, immobilizing the FPs on a solid phase. This method allowed the identification of 229 FPs in the cytosolic proteome of Drosophila melanogaster Kc167 cells without any pre-enrichment step.

4.11.4 Conversion to aminoethylcysteine

Even after a good preconcentration step it is difficult the exact assignation of a phosphorylation site, due to the lability of the phosphate group, often lost during the backbone fragmentation in a MS collision, and to the intrinsic low abundance of phosphorylated peptides.

To address this problem, Knight et al. [Knight et al., 2003] devised a derivatisation method based on β-elimination and Michael addition of cysteamine to convert phosphoserine and phosphothreonine in aminoethylcysteine (Aec) and β-methylaminoethylcysteine
respectively. Due to the resemblance of Aec to lysine, the use of proteases that recognize this aminoacid (e.g. Lys-C and lysyl endopeptidase) cleaves proteins leaving it in the C-terminus. The system works also with β-methylaminoethylcysteine and permits to identify the exact site of phosphorylation.

The limit of the method is the racemization in the addition step, converting only 50% of the phosphoaminoacids in the appropriate enzyme substrate. In the case of multiply phosphorylated peptides this fact greatly increases the complexity of the peptidic mix arising from the protein.

4.12 Comparison of enrichment methods

From this survey of enrichment methods emerges that no single technique is able to tackle the entire phosphoproteome. Some methods work in the direction of enriching only some species, like the antibodies for phosphotyrosine or the combination β-elimination/Michael addition selective for phosphoserine and phosphothreonine. Other methods, like calcium precipitation, SAX, SCX and HILIC, work better as preseparation techniques to reduce sample complexity before more specific enrichment methods like IMAC, MOAC, SIMAC and PAC.

Every enrichment technique presents advantages and disadvantages, but also different specificities. Usually, MOAC is more specific for monophosphorylated peptides, due to the strong affinity for the multiply phosphorylated ones, not enough eluted. On the contrary, IMAC is more specific for multiply phosphorylated peptides, but has a low capacity and selectivity when used with highly complex samples. The combined approaches, like SIMAC, seem to be promising but reveal the necessity of a pre-enrichment step [Han et al., 2008; Thingholm et al., 2008a].

A systematic comparison of methods was made by Bodenmiller and co-workers (fig.3). They examined the reproducibility, specificity and efficiency of IMAC, PAC and two protocols for TiO₂ chromatography: pTiO₂ (phthalic acid in the loading buffer to quench nonspecific binding) and dhbTiO₂ (2,5 dihydrobenzoic acid, quencher too) [Bodenmiller et al., 2007]. Each method was tested through the injection of 1.5 mg of tryptic digest from cytosolic fractions from Drosophila melanogaster cells. The authors found a very good reproducibility of all the methods, making them suitable for quantitative analysis. Moreover, none of the methods was able to reveal the entire phosphoproteome, but they show partial overlapping results between each other.

In general a simple and straightforward strategy is desired, with few preparation steps and little sample handling in order to avoid loss of FPs. It is of course critical also the amount of starting material and the expertise of the people performing the extractions. For this reason detailed protocols are needed [Goto & Inagaki, 2007; Thingholm et al., 2006; Thingholm et al., 2009b; Turk et al., 2006].

The graph shows the efficiency and selectivity of IMAC, PAC and TiO₂ applied on a tryptic digest of a cytosolic protein extract of D.melanogaster cells [Bodenmiller et al., 2007]. In the starting material no FPs were detected, while the best selectivity in terms of P vs. not-P sites was IMAC.
5. **MS-based strategies for phosphoproteome analysis**

Mass spectrometry has become the preferential method for peptide and protein identification following the separation steps, also in the PTM analysis [Bennett et al., 2002; Domon & Aebersold, 2006; Loyet et al., 2005].

The first step of a typical MS analysis consists in the cleavage of a single protein or a mixture by using a dedicated enzyme, usually trypsin, which preferentially cleaves the peptide bonds after arginin or lysine. Moreover, the tryptic fragments’ weight is 700-3500 Da, a size suitable for MS analysis. The peptides are thus separated by nanoLC and vaporized/ionized through an ESI source. Their mass is evaluated and a second fragmentation, generating MS/MS spectra, permits to evaluate also their aminoacidic sequence. This is possible because of the higher lability of the bonds between aminoacids. Depending on the position of the cleavage along the peptide chain, the MS/MS fragments are classified in a,b,c (starting from the N-terminal) or x,y,z (starting from the C-terminal) according to Roepstorff and Fohlman [Biemann, 1988; Roepstorff & Fohlman, 1984] (fig.4). Only the highest abundance peptides are submitted to MS/MS. This creates a hurdle in phosphoproteomics, because of the lower abundance and difficult ionization of FPs compared to the co-present not-FPs, thus introducing the need of an enrichment step, as explained in section 4.

![Peptide fragmentation](image)

**Immonium ion**

\[
\text{Immonium ion} = \text{H}_2\text{N}^+ - \text{C-R} - \text{H}
\]

**Phosphotyrosine immonium ion**

\[
\text{Phosphotyrosine immonium ion} = \text{H}_2\text{N}^+ - \text{C-CH}_2-\text{O}-\text{PO}_3^-
\]

![Fig. 4. Common nomenclature of peptide fragment ions.](image)
The information about the peptide sequence is submitted to database-digging softwares as MASCOT [Perkins et al., 1999] or SEQUEST [Ducret et al., 1998], which explore protein databases to find a sequence match with previously annotated proteins and rank the correlations through a probability score.

The peptide fragmentation in MS/MS mostly breaks the inter-residue bonds to generate fragment series. CID generates preferentially y and b ions, while mostly z and c ions are originated by ECD and ETD. Phosphotyrosine immonium ion is diagnostic of tyrosine phosphorylation.

5.1 Collision Induced Dissociation (CID)

The most established method to induce a secondary fragmentation in peptides is the collision induced dissociation (CID). Basically, the peptide ion collides with an inert gas (He or Ar) which transfers its kinetic energy, subsequently redistributed between the atoms bringing to the breaking of the bonds. When a phosphoserine or phosphothreonine is present in the peptide sequence, the phosphoesteric bond is by far the most labile, thus a neutral loss of phosphoric acid $\text{H}_3\text{PO}_4$ (98 Da) takes place, originating respectively dehydroalanine and dehydroaminobutiric acid. Given that most part of the energy is employed to break the phosphoesteric bond, far less energy is available for the subsequent fragmentation of the peptide chain [Larsen et al., 2005]. This drains information when the identification of the phosphate group position is needed: only the bare presence or absence of a phosphate is assessed.

To overcome this issue several strategies have been applied. The first one is the introduction of a tertiary fragmentation, specifically directed towards peptides where a phosphate loss is detected. This strategy is named pdMS$^3$ (phosphorylation directed MS3) [Reinders & Sickmann, 2005]. The information due to the alternative fragmentations of the precursor ion are in this case lost, but they can be kept through another approach, named Multi-Stage Activation (MSA) [Steen et al., 2001]. In this case the ion trap, filled with the selected ion coming from neutral loss, is filled again with the original peptide and both are fragmented at the same time, originating a superimposed MS$^2$ / MS$^3$ spectra more information-rich.

Partial neutral loss happens also on phosphotyrosine residues, which leave a $\text{HPO}_3$ group (80Da) originating a characteristic phosphotyrosine immonium ion at m/z 216 (fig.4). The phosphoesteric bond is however in this case more stable, thus not compromising the information collection. Steen et al., for example, used the diagnostic fragment at m/z 216 for the selective detection of phosphotyrosine-containing peptides in chicken ovalbumin and murine MAP-kinase 2 [Steen et al., 2001].

5.2 Electron Capture Dissociation (ECD)

The limit of the peptide backbone poor fragmentation in the presence of a phosphate group was overcome in 1998 with a new fragmentation strategy. Electron Capture Dissociation (ECD) is a method developed by Zubarev and colleagues to improve the fragmentation of multiply charged protein and peptide ions [Zubarev et al., 1998]. These ions capture easily a thermal electron (<0.2 eV), which induces a non ergodic fragmentation, i.e. without vibrational energy redistribution like in CID. The result is a fragmentation mostly at S-S and
N-Cα backbone bonds, leaving intact the PTM bonds. The generated ions are c and z type (fig.4) [Kleinnijenhuis et al., 2007, Stensballe et al., 2000]. The method has some drawbacks, like a bigger affinity for disulphide bonds and a difficult fragmentation of N-terminal proline, which has two bonds to break. Moreover it can be carried out only with expensive FT-ICR instruments (up to $1 million) to generate the static magnetic field for the electrons, which reduces its wide scale diffusion.

5.3 Electron Transfer Dissociation (ETD)

The efforts to find an ECD-like method without the need of expensive instruments brought to the advent of ETD (Electron Transfer Dissociation). In this approach the electron is transferred to multiply charged peptides (charge >2+) through a radical anion with low electron affinity, like anthracene or azobenzene [Schroeder et al., 2005; Syka et al., 2004]. The method can be implemented on linear quadrupole ITs, with the natural drawback of a reduced resolution and accuracy [Syka et al., 2004]. Molina et al. carried out a large scale analysis of human embryonic kidney 293T cells, identifying 1435 phosphosites, 80% of which were novel. Moreover, they identified 60% more FPs with ETD compared to CID, mainly due to the 40% more fragment ions [Molina et al., 2007]. It has to be remarked the little overlap between the two fragmentation techniques, that was exploited to develop an integrated approach. Since ETD works better with high charge peptides, Lys-C was thought to give better results than trypsin, cleaving the peptides only at C-terminal lysine. Surprisingly the results didn't match the expectations, probably due to the high number of missed cleavages in the tryptic lysate [Molina et al., 2007]. Another way to generate highly charged peptides was attempted by Larsen et al, who added 0.1% m-nitrobenzyl alcohol (m-NBA) to the LC-MS solvent [Kjeldsen et al., 2007]. This approach increased the predominant charge from 2+ to 3+, improving the ETD results. The approach is currently being tested on more complex samples. Another fact to remark is the evolution of the software, born for the CID approach, in the direction of meeting the features of spectra generated by new enzymes and fragmentations [Kim et al., 2010, http://www.matrixscience.com]

<table>
<thead>
<tr>
<th>Method</th>
<th>Fragmentation agent</th>
<th>Generated ions</th>
<th>Instruments</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>CID</td>
<td>Inert gas (He, Ar)</td>
<td>y, b</td>
<td>ESI-MS</td>
<td>Better fragmentation of low charge peptides (2+)</td>
<td>Fragmentation mostly at the phosphogroup</td>
</tr>
<tr>
<td>ECD</td>
<td>Thermal electron</td>
<td>z, c</td>
<td>FT-ICR</td>
<td>Fragmentation only along the peptide bond</td>
<td>Need of expensive FT-ICR</td>
</tr>
<tr>
<td>ETD</td>
<td>Low electron affinity anion (e.g. anthracene)</td>
<td>z, c</td>
<td>IT, Q-TOF</td>
<td>Fragmentation only along the peptide bond</td>
<td>Less sensitive than CID</td>
</tr>
</tbody>
</table>

Table 2. Comparison of fragmentation methods.

All the methods show good results with a class of peptides, suggesting that an integrated approach CID/ECD or CID/ETD could be more effective [Molina et al., 2007].
6. Quantitative approaches for phosphoproteome analysis

In order to take a dynamic picture of the phosphorylation events in a particular pathway, it is desirable monitoring which sites are phosphorylated and to which extent following a stimulus. To achieve this goal some quantification methods are available and can be classified on the basis of the analysis step in which the quantitative information is generated: a differential isotopic label can be introduced in the cell culture, e.g. with labeled aminoacids (SILAC), in the protein mixture (ICAT), in the enzymatic digestion (\(^{18}\)O labeled water), or in the peptide mixture (iTRAQ), otherwise, in label-free experiments, the quantitative information is extracted at the MS level (fig.7). A thorough review about quantitation strategies has been published by Bantscheff et al. [Bantscheff et al., 2007].

<table>
<thead>
<tr>
<th>METHOD</th>
<th>LEVEL</th>
<th>CELL</th>
<th>PROTEINS</th>
<th>PEPTIDES</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SILAC</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(^{18})O</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICAT</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>iTRAQ</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>LABEL-free</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Fig. 7. Strategies for quantitative analysis of protein phosphorylation.

An isotopic label can be introduced in different moments of the analysis or not at all, in label-free experiments.

6.1 Metabolic labeling

Metabolic labeling was first described in 1999 [Oda et al., 1999]. In 2002 Mann and coworkers introduced the term Stable Isotope Labeling by Aminoacids in Cell culture (SILAC) and used it for quantitative analysis of protein phosphorylation in 2003 [Ibarrola et al., 2003]. The typical experiment consists of growing a cell population in a medium containing an essential aminoacid labeled with a stable isotope (\(^{15}\)N or \(^{13}\)C), and growing in parallel another cell population in a medium on non-labeled aminoacid. Usually labeled arginine and lysine are used, in order to ensure that every peptide from one culture contains a label after tryptic digestion. After several doublings, the cells are harvested from both cultures, and the protein extracts mixed together. After proteolysis, peptides can be analysed by MS and differences in the abundance of a peptide in the two cell extracts are shown through the different heights of two mass shifted peaks.

Recently, with this method Olsen et al. [Bodenmiller et al., 2007] reported the most comprehensive analysis of the effects of EGF stimulation on phosphoproteome dynamics in HeLa cells. This strategy has allowed the drawing of some detailed maps of time-resolved signaling pathways [Blagoev et al., 2004; Bose et al., 2006; Goss et al., 2006; Olsen et al., 2010]. The major limitation of SILAC stays in the cost of labeled aminoacids.
6.2 Protein and peptide labeling

Post-biosynthetic labeling of proteins and peptides is performed by chemical or enzymatic derivatization in vitro.

Enzymatic labeling exploits the incorporation of $^{18}$O atoms from marked water during protein digestion. Trypsin and Glu-C introduce two heavy oxygen atoms, resulting in a 4 Da mass shift, generally sufficient for the differentiation of isotopomers. This method has been applied for quantitative proteomic purposes [Dengjel et al., 2007], but complete labeling is difficult to obtain.

Chemical modification can be carried out at protein or peptide level introducing a tag on a chemically reactive side chain of an amino acid [Ong et al., 2005], in practice only cysteine and lysine are used for this purpose. A group of labeling reagents targets the N-terminus and the ε-aminogroup in the lysine side chain. They mostly exploit the N-hydroxysuccinimide (NHS) chemistry or other active esters and acid anhydrides, like in the Isotope-Coded Protein Label (ICPL) [Schmidt et al., 2005], isotope Tags for Relative and Absolute Quantification (iTRAQ) [Ross et al., 2004], Tandem Mass Tags (TMT) [Thompson et al., 2003] and acetic/succinic anhydride [Che & Fricker, 2002; Glocker et al., 1994; X.Zhang et al., 2002].

iTRAQ is a commercially available reagent, allowing to follow the evolution of biological systems over multiple time points. It was used, for example, to quantify 222 tyrosine phosphorylation sites across seven time points following EGF stimulation [Wolf-Yadlin et al., 2007].

Carboxylic groups of side chains of aspartic and glutamic acid as well as of the C-termini of peptidic chains can be isotopically labeled by esterification using deuterated alcohols, for example d0 and d3 methanolic HCl [Goodlett et al., 2001; Syka et al., 2004]. This reaction is particularly interesting, because the methylation is also a step used in the IMAC enrichment method to reduce aspecific binding of acidic peptides to the resin (par. 4.2).

General drawbacks of the chemical derivatization methods are the production of not desired side products, that negatively influence the quantification results and the cost of some of the mentioned reagents.

6.3 Absolute quantification using internal standards

The use of isotope-labeled internal standards in the field of proteomics is known with the name AQUA: Absolute QUAntification of proteins [Gerber et al.,2003].

The simplest protocol requires adding a known amount of a stable isotope-labeled peptide to the protein digest and in comparing the signal of it in the mass spectra respect the other peak areas [Pan et al., 2005].

There are some drawbacks with this approach. First of all the high dynamic range of concentrations of peptides makes difficult to find an appropriate concentration of standard for every analyte; second, it’s likely to find an isobaric peptide to our standard in the peptide mixture, therefore limiting its specificity. These problems, however, have been addressed with the approach called Multiple Reaction Monitoring (MRM) [Kirkpatrick et al., 2005], in which the triple quadrupole MS monitors both peptide and its fragments mass.
during the experiment. The combination of retention time, peptide mass and fragment mass practically eliminates the ambiguities, extending the dynamic range to 4-5 orders of magnitude [Bondarenko et al., 2002]. The real value of the quantification through the AQUA approach is naturally biased by the manipulation of sample before adding the standard: the amount of protein determined may therefore not reflect its actual expression level in the cell.

6.4 Label-free quantification

There are two approaches for label-free quantification of proteins. The first one relies on the measure of the area of a MS peak of a peptide related to a protein: the increase of this area means also an increased amount of the protein. This approach is called eXtracted Ion Chromatogram (XIC), because a single ion peak area is extracted from a plot of signal intensities against time in the chromatogram [Bondarenko et al., 2002; Wang et al., 2006]. Signal intensities of the same peptide in different experiments is then compared to extract quantitative information, for example the stoichiometry of phosphorylation [Steen et al., 2005].

The other approach measures the amount of peptides generated from a protein: the more is the amount of a protein the more are the tandem-MS generated peptides. Relative quantification is thus achieved by comparing the number of spectra generated from a protein in different experiments. It is necessary a normalization, for example depending from the protein mass, creating therefore Protein Abundance Indexes (PAIs) [Rappsilber et al., 2002]. The relationship between number of peptides observed and protein amount had been found to be logarithmic (emPAI) [Ishihama et al., 2005; Lu et al., 2007].

7. Non-MS approaches to elucidate cellular signaling networks

7.1 Antibody-based approaches

In order to monitor previously identified phosphorylation sites, arrays employing phosphospecific antibodies have been used to investigate dozens of phosphorylation sites simultaneously [Sheenan et al., 2005; Belluco et al., 2005]. The general hurdle of these techniques is the limited availability of dedicated antibodies, however further improvements could extend the use of microarray technology in phosphoprotein studies [Schmelzle & White, 2006].

Methods were developed to monitor the phosphorylation status of tyrosine [Gembitsky et al., 2004] and the kinetics of phosphorylation [Khan et al., 2006] in proteins in a multiplex format.

In order to evaluate the phosphorylation dynamics on a cellular scale, flow citometry approaches have been also devised to monitor up to 11 phosphorylation events in parallel [Irish et al., 2004; Krutzik et al., 2005; Sachs et al., 2005]. Again, the main limit of this approach is the availability of suitable fluorescent-labeled antibodies.

7.2 Interaction of phosphoproteins and phosphorylated sites

The phosphorylation-related events include also protein-protein interactions in the cell signaling network. To investigate these phenomena, Jones et al. [Jones et al., 2006] devised a
protein array to study the binary interactions between 61 fluorescent-labeled, tyrosine phosphorylated peptides from EGFR receptors with approximately 150 SH2 and PTB domains. By measuring the fluorescence at different titration points they determined the $K_D$ values for every peptide-receptor couple.

Another approach was followed by Yaoi et al. [Yaoi et al., 2006], that immobilized SH2 domains on microspheres to extract interacting proteins and phosphoproteins from a complex mixture of different cell lines.

Both approaches revealed new insights in the cellular signaling networks.

### 7.3 Kinase screening on peptide and protein arrays

Peptide microarrays consist of synthetic peptide sequences deposited onto glass slides or attached to a derivatized surface, usually in triplicate, with peptides having substitutions in the phosphorylation sites as controls. The *in vitro* phosphorylation reaction is performed in the presence of radiolabeled ATP, the array exposed to a film and the image captured. The method assumes that phosphorylation of peptides should be in most of the cases similar to that of the same sequence in the intact protein, due to the fact that many phosphorylation sites are in accessible and flexible regions of the protein structure [Nühse et al., 2004].

Collins et al. used this approach for phosphorylation investigation of synaptic proteins, finding 28 unique phosphorylation sites [Collins et al., 2005].

The *in vitro* phosphorylation can naturally be different from the *in vivo* action, but the screening can select and give priority to some phosphorylation sites for further investigation.

The same approach can be used for immobilized proteins or protein domains.

Ptacek et al. [Ptacek et al., 2005] immobilized yeast proteins on high density (4400 proteins in duplicate) arrays on glass slides. They screened 87 kinases, finding that each kinase recognized up to 256 substrates, with a media of 47 substrates per kinase. These data allowed the construction of a global kinase-substrate interaction network. There is of course a concern about non specific phosphorylation, but also the perspective of a high throughput analysis for mapping phosphorylation networks.

### 8. Bioinformatics

The knowledge discovery process in proteomics has been greatly boosted in the last years by the introduction of new bioinformatic tools.

Widely developed phosphoproteomics databases are for example PhosphoSite [Hornbeck et al., 2004], containing around 100000 non-redundant phosphorylation sites (as well as other modifications, given that the cell signaling is not exclusively phosphocentric), and Phosida [Gnad et al., 2007], containing temporal phosphorylation data from cell stimulation in time-course experiments.

These databases permit not only the data mining, but also the interpretation of the data in the context of biological regulation, diseases, tissues, subcellular localization, protein domains, sequences, motifs, etc. [http://www.phosphosite.org]
9. Conclusion

Phosphoproteomics is a rapidly growing field, owing this evolution to the importance of the protein phosphorylation in many biological processes and its alteration in many diseases.

The analysis is usually performed with MS-based methods, supported by enrichment steps at the protein or peptide level. The improvement of MS has been enormous, with increase in resolution, mass accuracy, larger dynamic range and more sensitivity and speed, driving the progress in this field. Of course it must be mentioned also the evolution in bioinformatics, with the developing of adequate software for literature mining, prediction algorithms, post-analysis annotation and so on. The aim of phosphoproteomics is not only to find phosphorylation sites, but also monitor the dynamic of phosphorylation following stimuli to characterize completely signaling networks.

Nowadays the phosphoproteome of highly complex samples has been tackled [Nühse et al., 2003; Olsen et al., 2006; Trinidad et al., 2006; Villen et al., 2007], but further development of the methods, including of the bioinformatic tools to integrate complex databases, is needed for a more thorough knowledge of the mechanisms of phosphorylation networks.

10. References


absolute protein amount in proteomics by the number of sequenced peptides per protein. Molecular and Cellular Proteomics, Vol.4, No.9 (Sept 2005), pp.1265–1272


Protein Purification


Protein Purification


The current volume entitled Protein Purification is designed to facilitate rapid access to valuable information about various methodologies. It aims as well to provide an overview of state-of-art techniques for the purification, analysis and quantification of proteins in complex samples using different enrichment strategies.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
