Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Plant Biology

Proteomics approaches to understand protein phosphorylation in pathway modulation

Waltraud X Schulze

Signaling pathways in all organisms consist of series of phosphorylation and dephosphorylation events that define directionality and allow different levels of feedback-regulation. Mass spectrometry-based proteomic analyses in recent years have led to a proteome-wide identification of thousands of phosphorylation sites in various plant species. Given this magnitude of mostly qualitative information about protein phosphorylation, discovery of specific phosphoproteins with regulatory functions represents a major challenge. In future large-scale experiments, combinations of data-driven modeling strategies based on quantitative data, targeted kinase–substrate screens, and verification in biochemical and genetic experiments are required to specifically spot phosphorylation sites with specific roles in signaling pathway modulation.

Address
Max Planck Institut für Molekulare Pflanzenpyhsiologie, Am Mühlenberg 1, 14476 Golm, Germany

Corresponding author: Schulze, Waltraud (wschulze@mpimp-golm.mpg.de)

Current Opinion in Plant Biology 2010, **13**:280–287

This review comes from a themed issue on
Physiology and metabolism
Edited by Uwe Sonnewald and Wolf B. Fromm

Available online 22nd January 2010

1369-5266/$ – see front matter

DOI 10.1016/j.pbi.2009.12.008

Introduction

Apart from regulation of protein abundance by transcription and translation, direct fine-tuned control of protein activity, protein abundance, or protein localization is achieved at the post-translational level. Phosphorylation of serine, threonine, tyrosine, and also of histidine and aspartate are considered as most important regulatory post-translational protein modifications in all organisms. Kinases and phosphatases counterbalance phosphorylation of their target proteins, thereby achieving specificity and fine control in signaling pathways through multiple regulatory feedback loops [1]. Protein kinases make up about 5.5% of the Arabidopsis genome [2]. This fraction is nearly twice as high as that in mammals [3], suggesting a particular complex network and high specificity of phosphorylation events in plants. The classic view of signal transduction pathways leads from receptor proteins at the plasma membrane to transcription factor proteins in the nucleus. However, signaling pathways also involve intracellular metabolite receptors transmitting information to effector enzymes.

Protein phosphorylation often leads to a structural change of the protein that can directly modulate protein activity, and induce changes in interaction partners or subcellular localization. Owing to the tight spatial and temporal control observed in signaling pathways, and as a means of achieving directionality in signaling networks, protein phosphorylation events with regulatory functions are often of low stoichiometry and transient nature [4]. Phosphorylation stoichiometry was found between <10% and 90% depending on the phosphorylation site and protein, but is rarely being determined on a proteome-wide scale, as this requires splitting of the sample, dephosphorylation of one set, and accurate comparative quantitation [5]. Especially in enzymes and kinases phosphorylation sites with about 5% phosphorylation can trigger functional effects [5].

In plants, regulation of proteins through phosphorylation has been studied extensively on purified proteins and by site-directed mutagenesis of phosphorylation sites. Especially regulatory phosphorylation on metabolic enzymes, such as nitrate reductase or sucrose phosphate synthase were studied thoroughly [6,7]. In recent years, the identification of protein phosphorylation sites has become routine through detection of phosphorylated peptides by mass spectrometry [8]. The breakthrough in efficient proteome-wide analysis of phosphorylation sites came with development of suitable enrichment methods for phosphoproteins or phosphopeptides from complex protein digests (Table 1). At the same time, technical advances of mass spectrometer sensitivity to subfemtomolar detection limits, increased mass accuracy to less than one ppm, as well as higher resolution and dynamic range increased confidence in protein identification. The development of suitable data-dependent ion scan procedures [9,10], soft ion fragmentation methods [11] (Table 2), and identification of typical caveats in fragmentation spectra interpretation [12] finally resulted in a boost in the hunt for protein phosphorylation sites.

This review covers large-scale datasets collected over the past three years and addresses the need for future experimental and computational strategies to better understand the biological role of protein phosphorylation in signaling pathways on a proteome-wide scale.

Table 1

Methods of phosphopeptide enrichment are biased toward selective phosphoproteome subsets. Overlap of identified phosphopeptides between two different methods can be as low as 30% [57••]. For best phosphopeptide coverage, more than one method need to be used

Resin type	Metal ion	Efficiency (%)	Quenching agent	Comments	Examples
Affinity purification					
IMAC	Fe^{3+}, Ga^{3+}, Zn^{2+}	50–90	Sometimes chemical derivatization (methyl-esterification)	Non-specific binding of acidic peptides to the matrix. Specificity can be increased through derivatization, but this is often involved with general loss of sample. More efficiently for multiply-phosphorylated peptides.	[14,15•, 23•,24•]
TiO_2	Ti^{2+}	30–60	2,5-Dihydroxybenzoic acid	Non-specific binding of acidic peptides can be quenched by different acidic chemicals.	[57••]
		60–80	Phthalic acid	More efficiently for singly phosphorylated peptides.	[15•]
		60–80	Lactic acid		[14]
ZrO_2	Zr^{2+}	60	ß-Hydroxypropanoic acid		[57••]
$Al(OH)_3$	Al^{3+}	30		Can be used also to purify phosphoproteins.	[58]
Chemical purification					
Phosphoramidate chemistry (PAC)		70		Phosphopeptides are coupled to solid-phase matrix; elution under acidic conditions. Phosphate group remains attached to peptide, efficient for singly phosphorylated peptides.	[57••]
Beta elimination		80–90		Resulting double bond can react with nucleophilic functional groups for selective purification. Possible side reaction. Phosphate group is removed from peptide.	[59]

Large-scale mass spectrometry-based phosphoproteomics datasets

Qualitative datasets

Numerous plant phosphoproteomics studies have been published identifying hundreds to thousands of phosphorylation sites in various plant species [13–27]. In all of these phosphoproteomic profiling studies, phosphopeptide enrichment using metal oxide chromatography has been applied, and in most studies prefractionation to specific organelles (chloroplast, tonoplast, nucleus, and plasma membrane), or peptide fractionation by ion exchange chromatography was carried out. Among the studies focusing on Arabidopsis, 6% of all identified phosphorylation sites were found by more than two independent experiments, but 83% of the published phosphorylation sites were identified only once [28].

Despite impressive advances in identification of thousands of plant phosphorylation sites, they probably still represent a rather incomplete subset of the entire phosphoproteome. It is not yet clear, how many of the about two million potentially phosphorylatable sites in Arabidopsis (1 142 488 S, 664 750 T, 369 122 Y) are actually *in vivo* being accessible to modification and are then used under specific conditions. A high confident positive phosphorylation site prediction was obtained for nearly 500 000 of these residues (203 622 S, 174 301 T, and 120 983 Y), and experimental evidence from biochemistry or mass spectrometry was yet obtained for about 12 000 residues (9406 S, 2352 T, 699 Y) covering about 5000 proteins in Arabidopsis [28]. Also, it remains open to which extent phosphoproteome compositions will vary among the many different cell and tissue types in multicellular organisms [29]. Although different growth conditions and tissues have been analyzed [28], functional conclusions based on qualitative data are yet difficult to draw. This is due to biases in isolated phosphoproteomes by different enrichment methods and due to differences in phosphopeptide identification depending on instrumentation used (Tables 1 and 2). In addition, key regulatory proteins such as signaling proteins and transcription factors are often of low cellular abundance [30].

The increasing qualitative information about protein phosphorylation in various databases (Table 3), has led to a disconnection from functional aspects, as the majority of sites remains uncharacterized. It has in fact been suggested that large numbers of the phosphorylation sites identified by proteomic methods could be non-functional [31] making an important question apparent: What are the appropriate screens in finding phosphorylation sites which have regulatory roles? Given the difficulties described above, detecting 'relevant' phosphorylation sites by unbiased qualitative profiling methods may not be so straight forward. Nevertheless, simple experimental evidence of specific phosphorylation sites has in many cases helped defining experimental targets. For example, T881 was confirmed as additional regulatory site in the