Phosphoproteomics as a tool to unravel plant regulatory mechanisms

Reversible phosphorylation of proteins plays a key role in many regulatory processes that lie at the basis of life. With plants, much research has focused on protein kinases that are involved in the adaptation to different stress conditions, such as pathogen attack and cold. However, the substrates of these kinases are mostly unknown. With the recent advances in phosphoproteomic techniques, the large-scale identification of kinase substrates, including their phosphorylation sites, is finally possible. Studies in mainly non-plant systems have demonstrated the high potential of this method by uncovering numerous novel phosphorylation events. In this minireview, we focus on recent developments in the field of phosphoproteomics that are based on phosphopeptide isolation from complex mixtures by immobilized metal-affinity chromatography coupled to sequence identification by mass spectrometry. Combination of these methods with labelling techniques now allows quantitative analysis of phosphorylation between different samples. We discuss the potential of this technology to uncover entire phosphoproteomes and signalling pathways in plants in the future.

Introduction

Reversible phosphorylation is one of the most important and diverse post-translational modifications of protein function. It can influence multiple characteristics of proteins, including the enzymatic activity, subcellular localization, protein–protein interaction network and half-life. As much as 30% of all proteins may be phosphorylated at any time (Hubbard and Cohen 1993), indicating that the phosphoproteome of each multicellular organism is immense. Numerous cellular signalling pathways are based on the sequential phosphorylation of an array of proteins. Therefore, the analysis of signalling pathways in plants has often focused on protein kinases. Most of these studies, however, described the phosphorylation of single substrates by a particular kinase. Approximately 1000 genes in the Arabidopsis...
thaliana genome are predicted to encode protein kinases (The Arabidopsis Genome Initiative 2000). An intriguing difference with animal genomes is the absence of clearly recognizable tyrosine kinases. Moreover, the kinase complement of A. thaliana is more complex than that of humans, which consists of about 500 genes (Manning et al. 2002). Many plant protein kinases have been identified that are crucial in resistance to a variety of stresses including cold stress and pathogen invasion. Determination of the molecular events occurring during adaptation to stresses will enhance our understanding of biological processes in plants.

The identification of phosphorylation sites has been technically very difficult in the past. Technological advances on the basis of phosphopeptide isolation by immobilized metal-affinity chromatography (IMAC) followed by tandem mass spectrometry [MS/MS (MS2) or MS/MS/MS (MS3)] have emerged as a novel tool to determine phosphorylation sites in yeast, animals and plants (Chen and White 2004, Laugesen et al. 2004). Contrasting with the rapid expansion of large-scale phosphoproteomic studies performed on yeast and animals, similar work on plants is still limited to a few studies (Heintz et al. 2004, Gilinski and Weckwerth 2005, Nühse et al. 2003a). In this minireview, we briefly describe the role of phosphorylation in plant stress illustrated by examples of abiotic and biotic stresses: cold and pathogen stress. The remaining part focuses on the recent developments in the field of phosphoproteomics and its promise to unravel plant phosphoproteomes and their dynamics during stress.

**Reversible phosphorylation in cold stress signalling**

The transcripts of several protein kinases and phosphatases are upregulated during cold stress (Kreps et al. 2002, Seki et al. 2002), and thus adaptation to this stress may be a concerted action of different members of these families. The involvement of phosphorylation in cold signalling has been clearly established by the role of protein kinases and phosphatases in this stress. Several rice calcium-dependent protein kinases are induced by low temperature (Abbasi et al. 2004, Martin and Busconi 2001, Saijo et al. 2000) and overexpression of their genes enhances cold tolerance (Abbasi et al. 2004, Saijo et al. 2000).

Calcium influx and the hormone abscisic acid (ABA) are essential factors in cold signalling as well as other signalling cascades. The A. thaliana protein Ser/Thr kinase AtCIPK3 is transcriptionally activated by different stresses and is involved in the downstream activation of cold-, hyperosmotic- and ABA-related genes (Kim et al. 2003). AtCIPK3 might be activated by cold-induced calcium levels through interaction with a calcium calci-neurin B-like calcium sensor (Kim et al. 2003). The ABA-activated protein kinase (AAPK) phosphorylates the AAPK-interacting protein 1 (AKIP1), a heterogeneous nuclear RNA-binding protein. Phosphorylation of AKIP1 increases its affinity for dehydrin mRNA, and the binding stabilizes the mRNA, providing a mechanism of regulating protein level of dehydrin (Li et al. 2002). Dehydrins are implicated in defence against stresses. The A. thaliana dehydrin ERD14 itself is also phosphorylated upon cold treatment, which increases its binding affinity for calcium (Alsheikh et al. 2003). The phosphorylation of rice calreticulin and maize ribosomal protein S6 is increased upon cold treatment (Li et al. 2003, Williams et al. 2003), and are thus also substrates of the phosphorylation cascade activated by cold. In contrast to the positive role of kinases in cold signalling, the C-terminal domain phosphatase-like AtCPL1 negatively regulates cold-induced gene expression (Koiwa et al. 2002).

A common feature of plant stress-induced pathways is the signalling through so-called mitogen-activated protein kinase (MAPK) modules (Nakagami et al. 2005). They consist of MAPK kinase kinases (MAPKKKs) that phosphorylate and thereby activate MAPK kinases (MAPKKs) that in their turn phosphorylate and activate MAPKs. Recently, the MAPKK MKK2 has been recognized as an important mediator of cold stress signalling in A. thaliana (Teige et al. 2004). MKK2 is activated by cold and several other stresses, and overexpressor and knock-out lines show enhanced and diminished tolerance to cold treatment, respectively. Also MKK1 is activated by cold stress (Matsuoka et al. 2002). The stress-activated MAPK, SAMK is activated within 10 min after cold (and other stress) treatment (Jonak et al. 1996). The induction of cold-activated gene expression is mediated by cytoskeleton destabilization and increased membrane rigidity and depends on calcium (Sangwan et al. 2001). Strikingly, all these factors are required for cold-induced SAMK activity (Sangwan et al. 2002). The SAMK pathway is repressed by the alfalfa MAPK phosphatase 2C (Meskiene et al. 1998). Altogether, these findings show that MAPK signalling plays an important role in cold stress adaptation in plants.

**Reversible phosphorylation in plant defence signalling**

Pathogen-derived elicitors induce many phosphorylation events, and among the early targets are two cytosolic proteins and a plasma membrane syntaxin (Nühse
et al. 2003b, Peck et al. 2001). Candidate kinases are for instance calcium-dependent protein kinases, which are rapidly activated upon specific activation of defence responses (Romeis et al. 1999). Pto, which is an intracellular protein Ser/Thr kinase involved in resistance against bacterial speck disease, interacts with and phosphorylates both the protein Ser/Thr kinase Pti1 and the transcription factor Pti4 (Martin et al. 2003). Phosphorylation of Pti4 increases its binding to promoter elements of genes that are activated during disease-resistance responses.

As general stress regulators, MAPK pathways are also involved in defence responses in plants. Complete MAPK cascades have been identified in A. thaliana, tomato and tobacco, downstream of pathogen receptors (Nakagami et al. 2005). These MAPK modules act as positive regulators of signalling. In contrast, components of MAPK modules can also be negative regulators of defence responses (Nakagami et al. 2005). The recently discovered protein kinase OXI1 is an upstream activator of MAPKs and is essential for basal resistance against virulent pathogens (Rentel et al. 2004).

Besides protein phosphorylation, also protein dephosphorylation is important during defence signalling. A protein Ser/Thr phosphatase inhibitor mimics part of the response to pathogen-derived elicitors (Felix et al. 1994), suggesting that during the absence of pathogens, phosphatases may dephosphorylate positive regulators of defence responses. A genetic approach showed that protein Ser/Thr phosphatase 2A (PP2A) acts as a negative regulator of defence responses (He et al. 2004). Protein phosphatases are also involved in downstream signalling by the receptor-like kinase (RLK) FLS2 (Asai et al. 2002, Gómez-Gómez and Boller 2000). In conclusion, evidence is accumulating for the involvement of reversible protein phosphorylation in plant disease resistance.

**Approaches to study phosphorylation events in plants**

Although many plant protein kinases have been shown to be involved in signalling cascades, the number of their substrates that have been identified is still limited. For instance, although MAPKs are activated in response to numerous stress conditions, only two studies have described the identification of their in vivo substrates in plants (Andreasson et al. 2005, Liu and Zhang 2004). In addition, laborious site-directed mutagenesis of potential phosphorylation sites has often been used to determine the targeted sites within each substrate. More global studies are clearly needed to fully understand plant-signalling pathways. Recently, a protein array-based study has suggested a set of potential MAPK substrates (Feilner et al. 2005). Although the potential of this technique is evident, it still requires in vivo studies to verify the findings as well as to identify the phosphorylation sites.

Focused analyses of the roles of specific kinases and substrates have given a fragmented picture of plant-signalling pathways. As a global approach, a novel field has emerged that has the potential to speed up the unravelling of phosphorylation-dependent processes: phosphoproteomics. In spite of its technical drawbacks (for instance, the need of a fully sequenced genome, high equipment costs and high expertise required), phosphoproteomics has already firmly established itself as a method to uncover phosphorylation sites on hundreds of proteins in yeast, animals and A. thaliana.

**Enrichment strategies for mass spectrometry-based phosphoproteomic analysis**

Several problems are encountered during the examination of phosphopeptides by MS (Mann et al. 2002). Most importantly, phosphoproteins are generally of extremely low abundance within the proteome. These features have prompted the development of phosphopeptide (or phosphoprotein)-purification methods from complex mixtures. Several methods are based on the chemical substitution of the phosphate moiety of phosphopeptides by a stable group that can then be used to specifically purify the peptide (Chen and White 2004, Laugesen et al. 2004). Affinity purification of phosphoproteins with phospho-specific antibodies prior to MS has been tested in several studies. Although success with anti-phosphoserine/threonine antibodies is limited (Grønborg et al. 2002), antiphosphotyrosine antibodies are suitable for specific affinity purification as illustrated by many studies. Because of the low occurrence of tyrosine phosphorylation in animal cells (about 0.05% against more than 99% on serine and threonine residues), phosphotyrosine-containing peptides/proteins need to be specifically purified from complex mixtures. Immunoprecipitation with antiphosphotyrosine antibodies coupled to liquid chromatography-mass spectrometry (LC-MS) identified 57 (Ficarro et al. 2005), 64 (Salomon et al. 2003), 66 (Brill et al. 2004), 104 (Zhang et al. 2005) and 628 phosphotyrosine sites (Rush et al. 2005). In plants, such a study has never been performed, but a few studies have suggested the occurrence of tyrosine phosphorylation (for instance of MAPKs). Whether tyrosine phosphorylation is more abundant than assumed in plants remains to be shown.

A few successful studies have described the isolation of phosphopeptides from complex mixtures with strong
cationic exchange (SCX) chromatography (Ballif et al. 2004, Beausoleil et al. 2004) or strong anionic exchange (SAX) chromatography followed by IMAC (Nühse et al. 2003a). Coupling these techniques to MS-based peptide identification proved to be very successful. SCX was used to identify spectacular numbers of >500 and 200 phosphosites from mammalian cells. However, a drawback of SAX or SCX is the failure to efficiently capture multiply phosphorylated peptides (Beausoleil et al. 2004, Nühse et al. 2003a).

As a phosphopeptide-purification method, IMAC has already been used for two decades (Andersson and Porath 1986). IMAC is based on the high affinity of trivalent metal ions (Fe\(^{3+}\)) for phosphate groups. However, only recently, IMAC has been successfully used in large-scale phosphoproteomic analysis, leading to the identification of numerous phosphorylation sites (Ficarro et al. 2002). An important improvement to the protocol added by Ficarro et al. (2002) was the conversion of carboxylic acid groups to methyl esters prior to IMAC purification of peptides. This step avoids the non-specific binding of peptides containing acidic residues to the IMAC material. Thus, the addition of this step greatly enhanced the specificity for phosphorylated over non-phosphorylated peptides (Brill et al. 2004, Ficarro et al. 2002), although this chemical modification may not be a prerequisite for high specificity (Nühse et al. 2003a). In our hands, the purity increased substantially by introducing the esterification step. In conclusion, IMAC promises to be a valuable tool to specifically isolate phosphopeptides from complex mixtures and recent studies fruitfully described the use of this method for the large-scale analysis of phosphorylation sites.

**Global identification of in vivo phosphorylation sites using IMAC coupled to mass spectrometric technology**

Ficarro et al. (2002) successfully implemented enrichment of phosphopeptides using IMAC from yeast whole-cell protein extracts that have been digested with trypsin. Coupling of esterification-IMAC to subsequent identification of peptide sequences by tandem MS allowed the identification of 383 phosphorylation sites in yeast peptides (Ficarro et al. 2002) and 238 sites in human peptides (Kim et al. 2005). The power of the combined use of IMAC and MS for plant work has been demonstrated by the large-scale analysis of phosphopeptides from *A. thaliana* plasma membrane extracts (Nühse et al. 2004). This first plant global phosphoproteomic study identified more than 300 phosphorylation sites, of which only a few were known. In a study of the phosphoproteome of the moss *Physcomitrella patens*, 253 phosphopeptides were detected during the esterification-IMAC-MS approach, but the sequence of only a few could be determined (Heintz et al. 2004). In a small-scale study of *A. thaliana* thylakoid membrane-associated proteins, Hansson and Vener (2003) identified several novel phosphorylation sites using the same approach. The work of Nühse et al. (2004) provided the first insights into the nature of protein phosphorylation in plants. For instance, the extensive phosphorylation of RLKs showed complex regulation of these proteins, suggesting that phosphorylation sites may be key determinants of signalling specificity of these receptors. The analysis of plasma membrane phosphoproteins also showed that regions surrounding the phosphorylation sites can often be grouped into conserved phosphorylation motifs that may be targeted by similar kinases. Such analyses will accelerate the identification of the responsible kinase targeting the subset of phosphorylation sites determined in studies aiming at the identification of the total phosphoproteome.

We have adapted the technology as described by Ficarro et al. (2002) to identify in vivo phosphorylation sites of intracellular *A. thaliana* proteins. Our experimental setup consisted of the rough separation of intracellular proteins into a cytosolic and a nuclear fraction (Fig. 1). The aim was to unravel the phosphoproteome within these extracts. In a complex mixture containing up to 300 \(\mu\)g of protein, polypeptides were cleaved with trypsin and esterified. Subsequently, phosphopeptides were isolated by Fe\(^{3+}\)-IMAC. In Fig. 2, the efficiency of phosphopeptide isolation from a complex mixture by IMAC is depicted. It is clear that most (abundant) peptides are not captured by IMAC, but only the low abundant phosphorylation sites. Sequence information was provided by fragmentation of each peptide by the collision-induced dissociation (CID) on a quadrupole linear ion trap mass spectrometer (Finnigan LTQ), yielding tandem MS (MS/MS) spectra. In addition to the MS/MS spectra, the Finnigan LTQ enables a data-dependent third stage of MS (MS/MS/MS; specifically of peptides that are selected when a loss of phosphate is observed during MS/MS), leading to spectra with more sequence-specific peptide-fragmentation patterns. This is an advantage since phosphopeptides generally give poor-quality fragmentation patterns when fragmented by CID. To circumvent this problem, phosphopeptides can be dephosphorylated before MS analysis. However, this leads to loss of information about the location and amount of phosphorylation sites. Each phosphopeptide fragmentation spectrum has to be manually verified, which is often labour intensive. From MS/MS and MS/MS/MS spectra obtained in a single run, we could make confident assignments for more than 130...
phosphopeptides. Of these phosphopeptides, 79% was singly phosphorylated, 19% was doubly phosphorylated and 2% was triply phosphorylated. The functional classification of all the 151 phosphoproteins identified so far in our studies is shown in Fig. 3.

It has been stated that prefractionation (with for instance SAX) is required for the efficient recovery of monophosphorylated peptides from complex mixtures when using IMAC (Nühse et al. 2003a). However, we recover a high percentage of these phosphopeptides without any further prefractionation of the complex cytosolic mixture. Nonetheless, more elaborate prefractionation will be required for more complete coverage of phosphorylation sites and efficient identification of the lower abundant phosphopeptides. In our analysis of the intracellular phosphoproteome, we identified several described phosphosites of abundant metabolic enzymes and ribosomal proteins, which represent main groups among the phosphoproteins (Fig. 3). For instance, we determined a conserved phosphorylation site in sucrose–phosphate synthase (SPS) in two *A. thaliana* isoforms that is analogous to the known pSer-158 (lower case p indicates phosphorylated residue) of spinach SPS (Fig. 4). In addition to this pSer, we identified several novel phosphorylation sites (Fig. 4). A major target of phosphorylation is represented by the mRNA splicing machinery (S. de la Fuente van Bentem, D. Anrather, E. Roitinger, D. Lecourieux and H. Hirt, manuscript in preparation). In addition to abundant phosphoproteins, also phosphorylation sites of low abundant signalling proteins were detected in our studies.

**IMAC-based quantitative phosphoproteomics as a tool to elucidate plant signalling pathways**

Recent developments in MS-based phosphoproteomics promised to be an excellent instrument to study dynamic phosphoprotein profiling of signalling networks. In animal and yeast systems, several such studies have already been performed and identified many novel stress-induced phosphorylation events (Ballif et al. 2005, Blagoev et al. 2004, Cutillas et al. 2005, Gruhler et al. 2005a, Ibarrola et al. 2004, Kratchmarova et al. 2005, Zhang et al. 2005). A recently developed technique to measure differences between proteins in two distinct samples by MS is Stable Isotope Labelling by Amino acids in Cell culture (SILAC; Ong et al. 2002). This in vivo method is based on the feeding of separate cultures with distinct isotopically labelled amino acids, which can later be differentiated during MS analysis. SILAC has been used as an effective tool to measure relative phosphorylation differences of proteins extracted from cell cultures grown under different conditions. SILAC followed by IMAC and tandem MS has been used for differential phosphorylation profiling of human and yeast-signalling pathways (Blagoev et al. 2004, Gruhler et al. 2005a, Kratchmarova et al. 2005). SILAC has been also successfully applied to *A. thaliana* for quantitative proteomics and may therefore be used for quantitative analysis of the phosphoproteome during plant signalling (Gruhler et al. 2005b). A drawback of SILAC is that only phosphopeptides containing the labelled amino acid can be considered in the analysis. Fortunately, *A. thaliana*, labelling was most efficient with $^{13}$C$_6$-Arg (Gruhler et al. 2005b), which can be used conveniently in combination with trypsin digestion since this protease cleaves after basic amino acids.

An alternative approach to determine relative differences between different samples is the in vitro labelling of peptides before or after isolation from a complex mixture. Peptides can be labelled during esterification with deuterated methanol before IMAC. However, deuterated peptides appear to behave differently in the LC
run than the corresponding non-deuterated derivative (Chen and White 2004). The development of a multi-plex set of reagents allows the incorporation of mass labels at the N-termini and lysine side chains of peptides in a digest mixture (Ross et al. 2004). Using four iso-forms of this so-called iTRAQ reagent, Zhang et al. (2005) gave the first insights into the dynamic changes of phosphorylation profiles during EGFR tyrosine kinase signalling. Regarding in vitro peptide labelling methods, the success largely depends on the incorporation efficiency, and mass spectrometers with high mass accuracy are required for the relative quantification of phosphopeptide differences. However, compared with in vivo labelling methods, this requires additional handling time.

Conclusions and future prospects

The current MS-based phosphoproteomic technology has established itself as an invaluable tool in the
identification at novel phosphorylation sites. Future improvements of this technique will be used to decipher the global phosphoproteome and ultimately the dynamic behaviour of the complete phosphoproteome in plants. The pre-purification of single organelles or prefractionation of complex mixtures by, for instance, SAX/SCX will increase the coverage of the phosphoproteome in each study. The use of a protease panel to extend the sequence coverage of the phosphoproteome is a simple and efficient method (Rush et al. 2005). Automated and sensitive phosphorylation site mapping by the esterification-IMAC procedure combined with LC-MS will enhance the identification of complete phosphoproteomes (Ficarro et al. 2005). The recent development of specific labelling techniques greatly aids the quantification of phosphorylation profiles and their stress-induced changes in time. Especially iTRAQ labelling and SILAC have shown to be successful in combination with IMAC and MS. These studies will reveal hints at novel signalling pathways and regulatory processes that are dependent on phosphorylation. Although the scale of studies on signalling cascades is increasing rapidly, improvements are required to generate entire signalling webs. This will require better software for both the automated identification of phosphopeptide sequences by MS and efficient structuring of the wealth of data becoming available. The development of a mass spectrometer with a high mass accuracy and a suitable method for phosphopeptide fragmentation will advance global analysis of signalling pathways. Promising fragmentation techniques such as electron-transfer dissociation, which produces a high degree of sequence information of phosphopeptides (Syka et al. 2004), will have to prove their applicability for large-scale phosphopeptide analysis. These phosphoproteomic studies will induce future experiments to determine the function of the phosphorylation sites.

In summary, future studies aiming at global phosphoproteomics will greatly benefit from the recent developments in the field of MS-based technology. As discussed here, these methods are becoming equally applicable for plant studies. The rapidly increasing methodology for quantitative phosphoproteomics is about to revolutionize our conceptual understanding of plant biology and should uncover many unexpected links within the signalling network in plants.

Acknowledgements – This work was supported by grants from the Austrian Science Foundation, the Vienna Science and Technology Fund and the European Union.

References


abiotic stress signaling, growth, and development. Proc Natl Acad Sci USA 99: 10893–10898


