Dissertation

Submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola-University of Heidelberg, Germany

for the Degree of

Doctor of Natural Sciences

Presented by

Dipl.-Biol. Florian Wolschin

Born in Leimen

Oral Examination: 11.08.2006

Characterisation of protein phosphorylation in Arabidopsis thaliana and Chlamydomonas reinhardtii based on affinity chromatography and mass spectrometry

Referees: Prof. Dr. Michael Wink

Prof. Dr. Wolf D. Lehmann

Acknowledgements:

Hiermit möchte ich mich bei allen Personen bedanken, die mir diese Arbeit ermöglicht haben, mich unterstützt haben und dafür gesorgt haben, dass ich die letzten Jahre immer in guter Erinnerung behalten werde.

Bei Dr. Wolfram Weckwerth bedanke ich mich herzlich für seine andauernde Unterstützung, seine Diskussionsbereitschaft und Inspiration. Für die Begutachtung der Arbeit, die kompetente Hilfe und entscheidende Gedankenanstöße möchte ich mich bei Prof. Wink und Prof. Lehmann herzlich bedanken. Prof. Mark Stitt danke ich für die Bereitstellung von Mitteln und Geräten, ohne die eine erfolgreiche Arbeit nicht möglich gewesen wäre. Dr. Julia Kehr und Prof. Mark Stitt danke ich für Ihre Bereitschaft zu konstruktiven Evaluierungsgesprächen.

Außerdem bedanke ich mich herzlich bei:

……. meinen Eltern, meiner Schwester und dem Rest der Familie, die mich immer unterstützt haben und ausdauerndes Interesse an meiner Arbeit zeigen.

……. Antonieta por apoyarme durante todos estos años y por todos los momentos felizes que hemos pasado juntos.

……. Ralf Krüger für hervorragende Zusammenarbeit und die Durchführung von ICP-MS Analysen.

……. Stanislav Laptev und Regina Stark für Ihre Assistenz bei einigen Versuchen.

……. Katja Morgenthal, Mirko Glinski, Karsten Oelkers, Markus Rott, Nicolas Schauer und vielen anderen für jede Menge Spaß bei und außerhalb der Arbeit.

……. Waltraud Schulze und allen Mitarbeitern der AG Weckwerth für Ihre Diskussionsbereitschaft und Hilfe.

……. Daniel Karcher und Dietrich Köster für die Versorgung mit *C. reinhardtii* Kulturen und interessante Gespräche.

Publications:

Wolschin, F., Wienkoop, S. & Weckwerth, W. Enrichment of phosphorylated proteins and peptides from complex mixtures using metal oxide/hydroxide affinity chromatography (MOAC). *Proteomics* **5**, 4389-4397 (2005).

Wolschin, F. & Weckwerth, W. Combining metal oxide affinity chromatography (MOAC) and selective mass spectrometry for robust identification of in vivo protein phosphorylation sites. *Plant Methods* **1**, 9 (2005).

Wolschin, F. & Weckwerth, W. Methionine oxidation in peptides - a source for false positive phosphopeptide identification in neutral loss driven MS^3. *Rapid Commun Mass Spectrom* **20**. 2516-2518 (2006).

Wolschin, F., Lehmann, U., Glinski, M. & Weckwerth, W. An integrated strategy for identification and relative quantification of site-specific protein phosphorylation using liquid chromatography coupled to MS^2/MS^3. *Rapid Commun Mass Spectrom* **in press** (2005).

Morgenthal, K., Wienkoop, S., Wolschin, F. & Weckwerth, W. Integrative profiling of metabolites and proteins. In Metabolomics: Methods and Protocols 57-75 (Humana Press, Totowa, NJ; 2006).

Zusammenfassung:

Reversible Proteinphosphorylierung ist für viele Mechanismen des Lebens von zentraler Bedeutung. Die geringe Abundanz von Phosphoproteinen erschwert jedoch deren Analyse. Um dieses Problem zu umgehen wurde eine neue Methode für die Anreicherung von Phosphoproteinen aus komplexen Proteinmischungen entwickelt. Sie basiert auf der Affinität der Phosphatgruppe zu Aluminiumhydroxid und wurde MOAC genannt (Metal Oxide Affinity Chromatography). Die erfolgreiche Anreicherung von Phosphoproteinen aus hochkomplexen Proteinmischungen wurde durch die Trennung von phosphorylierten und nichtphosphorylierten Standardproteinen, die Detektion mittels eines phosphatspezifischen Fluoreszenzfarbstoffes, ICP-MS (Inductively Coupled Plasma – Mass Spectrometry) und die Verwendung von Antikörpern bestätigt.

Um Phosphopeptide und Proteinphosphorylierungsstellen zu identifizieren, wurden Proteine mit einer Endoproteinase verdaut und per LC/MS^n analysiert. Die Bestimmung von Proteinphosphorylierungsstellen wurde durch die Verwendung des charakteristischen Neutralverlustes von Phosphorsäure, wie er während der Fragmentierung in einer Ionenfalle bei Peptiden die an Serin- oder Threoninresten phosphoryliert sind, erreicht. Es wurde allerdings beobachtet, dass andere posttranslationale Modifikationen wie Methioninoxidation eine Phosphorylierung während solcher Analysen imitieren können. Daher wurden Kriterien definiert um bei diesem Problem Abhilfe zu schaffen und zu einer verlässlicheren Bestimmung von Phosphorylierungsstellen zu gelangen.

Die Kombination von MOAC und des Neutralverlust-abhängigen Ansatzes erwies sich als sehr nützlich für die Untersuchung der Proteinphosphorylierung bei Pflanzen. In einer der ersten ungerichteten Analysen im Berich Pflanzenproteomics wurde dieser kombinierte Ansatz verwendet um Phosphoproteine von *A. thaliana* und *C. reinhardtii* anzureichern und über 40 Phosphopeptide, 27 Phosphorylierungsstellen und über 30 Phosphoproteinen zu identifizieren. Zusätzlich wurden über 300 putative Phosphoproteine identifiziert. In einer zielgerichteten Analyse wurde die spekulative *in vivo* Phosphorylierungsstelle des metabolischen Schlüsselenzyms Phosphoenolpyruvatcarboxylase in *A. thaliana* bestätigt.

Während die Detektion der Phosphorylierung und die Bestimmung von Phosphorylierungsstellen sehr wichtig ist um einen ersten Eindruck davon zu erlangen, ob ein Protein durch Phosphorylierung beeinflusst werden kann, liefert die Quantifizierung der Phosphorylierung detailliertere Informationen über Proteinregulation durch Phosphorylierung. Unter Verwendung der ICP-MS war es möglich den Phosphorylierungsgrad des gesamten Proteoms für verschiedene Entwicklungsstadien von *A. thaliana* zu ermitteln. Diese Werte wurden mit dem Phosphorylierungsgrad von *C. reinhardtii* verglichen. Die Daten zeigen, dass Proteinphosphorylierung am häufigsten in sich schnell teilendem Gewebe und am wenigsten häufig in ruhenden Samen vorkommt.

Um die Phosphorylierung von metabolischen Schlüsselenzymen unter unterschiedlichen physiologischen Bedingungen zu untersuchen, wurde eine robuste Methode auf einer linearen Ionenfalle entwickelt, die der relativen Quantifizierung von Änderungen in der Proteinphosphorylierung, inklusive der Phosphorylierungsstöchiometrie dient. Diese Methode wurde verwendet um einen spekulativen Zusammenhang zwischen der Temperatur und der Phosphorylierung der Sucrose-phosphatsynthase (SPS) *in vitro* zu untersuchen. Die Resultate deuten darauf hin, dass die Temperatur einen ausgeprägten Effekt auf die SPS Kinase und damit die Phosphorylierung der SPS hat.

Summary:

Reversible protein phosphorylation is of key importance for several mechanisms of life. However, the low abundance of phosphorylated proteins hinders investigations following this direction. To circumvent this problem a new method for the enrichment of phosphorylated proteins out of complex protein mixtures was developed. This method relies on the affinity of the phosphate group towards aluminum hydroxide and was termed MOAC (Metal Oxide Affinity Chromatography). Successful phosphoprotein enrichment even from highly complex protein mixtures was confirmed by separation of phosphorylated from non-phosphorylated standard proteins, detection using a phosphate-specific fluorescent stain, ICP-MS (Inductively Coupled Plasma – Mass Spectrometry), and antibodies.

To identify phosphopeptides and protein phosphorylation sites proteins were digested with an endoproteinase and subjected to LC/MS^n analysis. The determination of protein phosphorylation sites was achieved by making use of the distinct neutral loss of phosphoric acid from peptides phosphorylated on serine or threonine during fragmentation in an ion trap mass spectrometer. However, it was observed that also other posttranslational modifications such as methionine oxidation can mimic phosphorylation when using such data-dependent neutral loss scans. Remedies were defined to circumvent this problem leading to more reliable phosphorylation site identification.

The combination of MOAC and a neutral loss driven MS^3 approach showed to be highly useful for investigations on plant protein phosphorylation. In one of the first non-targeted phosphoproteomics approaches in plant science this combined approach was applied to enrich phosphoproteins of *A. thaliana* and of *C. reinhardtii* leading to the identification of over 40 phosphopeptides, 27 phosphorylation sites, and over 30 phosphoproteins. In addition, over 300 putative phosphoproteins were identified. In a targeted analysis the speculative *in vivo* phosphorylation site of the metabolic key enzyme phosphoenolpyruvate carboxylase in *A. thaliana* could be confirmed.

While the detection of phosphorylation and the determination of phosphorylation sites are very important to get a first impression whether or not a protein can be influenced by phosphorylation, quantification of phosphorylation delivers more detailed information on protein regulation by phosphorylation. Using an approach involving ICP-MS it was possible to monitor the overall protein phosphorylation degree for different developmental stages of *A. thaliana*. These values were compared to the degree of phosphorylation of proteins in *C. reinhardtii*. This comparison shows that protein phosphorylation is most abundant in rapidly dividing tissue and lowest in dormant seeds.

To monitor phosphorylation of key metabolic enzymes under different physiological conditions, a robust method for relative quantification of changes in protein phosphorylation on a linear ion trap was also developed. It was used to investigate a speculative connection between temperature and phosphorylation of sucrose-phosphate synthase (SPS) and to determine the stoichiometry of phosphorylation for the *in vitro* phosphorylated enzyme. The results indicate that temperature has a profound effect on the phosphorylation level of SPS by influencing the activity or abundance of the kinase responsible for SPS phosphorylation.

Chapter I: Introduction

I.1. *Arabidopsis thaliana* and *Chlamydomonas reinhardtii* – two model organisms of the green plant kingdom

A. thaliana and *C. reinhardtii* are two major model organisms that can be used to investigate different aspects of the green plant kingdom. While *A. thaliana* is multicellular, obligate photoautotroph, immotile, and diploid; *C. reinhardtii* is unicellular, can grow either photoautotrophically, mixotrophically, or heterotrophically, it is motile, and haploid [1]. Both organisms do not have direct economic impact but *A. thaliana* is a close relative of crop plants like rapeseed and cabbage and can thus serve as a general model of agronomic importance. The genome sequence of *A. thaliana* published in 2000 was the first plant genome to be sequenced [2] and this achievement has been greatly facilitated modern plant research ever since. No publication on the genome sequence of *C. reinhardtii* is available but the sequencing project is almost completed and drafts of this sequence can be used for research purposes.

The genome sizes of both organisms are comparable with about 100 Mb for *C. reinhardtii* [3] and about 125 Mb for *A. thaliana* [2], which is quite small considering that other plants reach 1000 Mb and more [4]. This small size facilitated genome sequencing and was one of the reasons for choosing these organisms for research. In addition, both organisms can be grown easily in large numbers, and genetic tools are available to manipulate both members of this dynamic duo making them ideal candidates for basic research.

Compared to *A. thaliana*, research on the proteome of *C. reinhardtii* is clearly underrepresented. This is mainly because of a bigger research community working with *A. thaliana* and because of the comparably well annotated *A. thaliana* genome. However, investigating aspects of the proteomes or subproteomes like the phosphoproteome (the complement of phosphorylated proteins) of these two organisms holds the promise of getting a deeper insight into the differences and similarities between these two members of the green plant kingdom.

There are many varieties of *C. reinhardtii* and *A. thaliana*, which are used for research purposes. Of those, *C. reinhardtii* strain *cc125*, which constitutes one of the standard strains, and *A. thaliana* ecotype Columbia, which was the one used for the genome sequencing project, were used in this work.

I.2. Plant proteomics or the modern version of plant protein biochemistry

The word protein was originally introduced by the Swedish chemist Berzelius who derived it from the Greek word *protas* (of primary importance) [5] and was quickly adopted by biochemical researchers working on these important biomolecules. The research area of proteomics is in principle the modern version of classical protein biochemistry. However, it usually deals with a larger number of proteins and employs methods compatible with large scale analysis like mass spectrometry and several forms of separation techniques. The term was coined in 1994 by Marc Wilkins who defined proteomics as "the study of proteins, how they're modified, when and where they're expressed, how they're involved in metabolic pathways and how they interact with one another." It is generally acknowledged today that large scale protein analysis can complement traditional biochemical studies like enzymatic essays and add valuable information about the organism under investigation (e.g. [6]). Proteomics can be regarded as the successor of genomics, the study of the genome, and the predecessor of metabolomics, the study of the metabolome. Taken together, these disciplines form parts of systems biology, an area of research which tries to gather as much information as possible about any organism and its interactions with the environment.

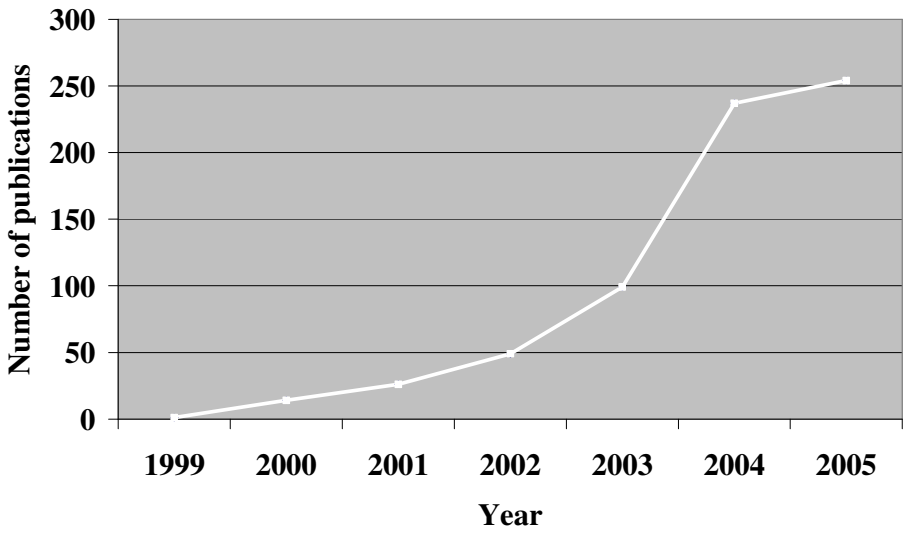


Fig. 1: The number of publications returned from the NCBI server when searching for "plant proteomics".

The interest in plant proteomics has been growing remarkably over the past years. A database search at the NCBI (national center for biotechnology information) with the term plant proteomics returned a single publication in 1999 and 254 in 2005 (figure 1).

There are several reasons why the study of the proteome is important. Proteins are involved in diverse processes including stability, energy conversion, reproduction, and communication, just to name a few and constitute the central part of metabolism. They are the mediators between RNA and metabolites and altering the level and/or activation state of proteins has a direct impact on metabolism, development, and signal transduction of the organism. The number of transcripts is only a very rough estimate of the protein amount present in a cell or organism since more than one protein can be made from one transcript and transcripts do not necessarily lead to the production of proteins. Thus, measuring proteins and their modifications is a far more accurate method when trying to get a deeper understanding of metabolism and its regulation.

I.2.1. The impact of mass spectrometry on protein research

The introduction of mass spectrometry into protein research in the late 1980s was the initial boost for proteomics. However, it was not until the 1990s that this technique became available to the broader range of researchers necessary to define it as a major application in protein biochemistry [7]. Most applications nowadays rely on either ESI (electrospray ionisation) or MALDI (matrix assisted laser desorption) as ionisation techniques (see figure 2). For biochemical research these ionisation methods are most often coupled to a TOF (time of flight) analyser or an ion trap. A breakthrough for the application of ESI in protein research was made by Fenn and colleagues who presented the identification of proteins and polypeptides by using ESI in 1988 [8]. For their important contributions to the analysis of biomolecules by mass spectrometry Fenn and Tanaka were awarded with the nobel prize for chemistry in 2002. In ion trap mass spectrometers analytes enter the ion transfer tube and are accelerated and focussed using quadrupoles, octopoles (assemblies of four or eight metal rods with alternating voltage), and lenses. They are then introduced into the trap. In the trap a collision gas (usually helium) is responsible for trapping of the molecular ions. In addition, this gas is employed as reagent gas for collision-induced reagent experiments.

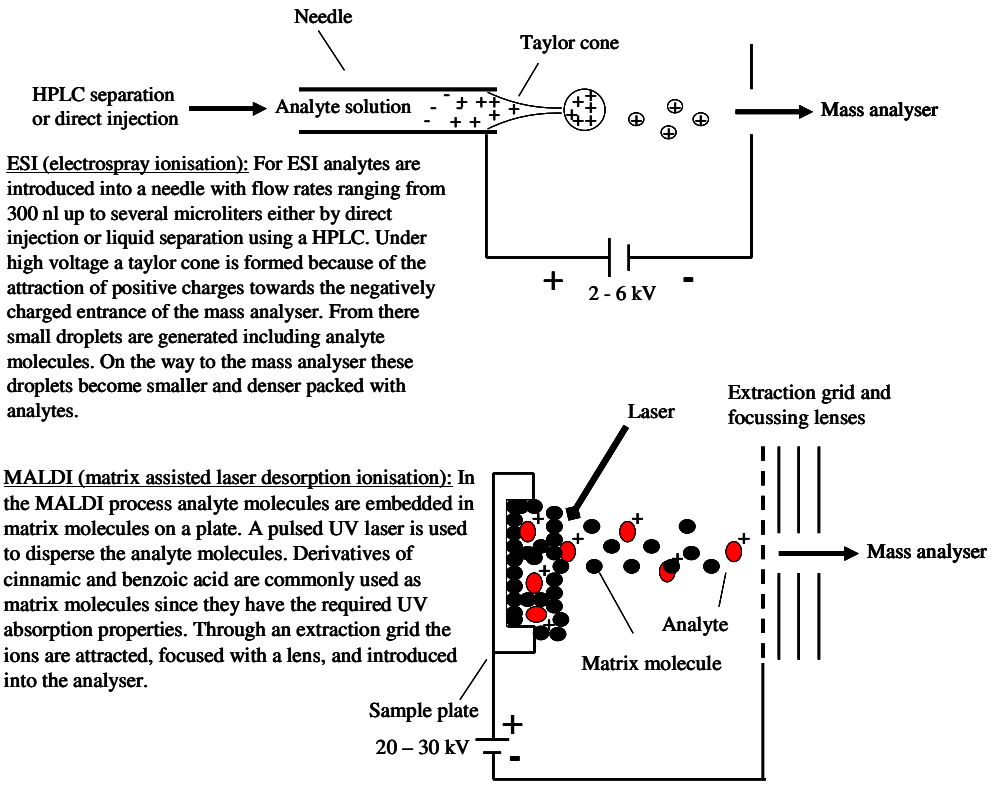


Fig. 2: ***The two prominent ionisation techniques in biological mass spectrometry: ESI and MALDI. See [9] for a comprehensive review on the ESI process.***

Major contributions to the development of MALDI and its application to the identification of peptides and proteins came from Karas, Hillenkamp, and Tanaka in the 1980s. This technique is most commonly used in combination with a time of flight analyser (TOF analyser). In a TOF analyser molecular ions are separated according to their size with the smaller molecular ions reaching the detector later. When the analyser is equipped with a reflector, ions can be fragmented using a technique called PSD (post-source decay). The TOF analyser has a high mass accuracy but unlike an ion trap can not be used to perform several consecutive fragmentation steps.

There are two main approaches to the identification of proteins: the “Top-down” and the “Bottom-up” approach [10-12]. The “Top-down” approach involves the investigation of intact proteins or huge (larger than 10 kDa) peptides. It can, because of the complexity of protein structure, only be performed using high resolution instruments like the Fourier transform ion-cyclotron mass spectrometer or an Orbitrap. Since these instruments measure protein masses with very high accuracy they can also be used to identify protein modifications. This is done by comparing the obtained mass data to the expected primary sequence of the protein. However, for an accurate measurement of proteins from a complex

mixture it is necessary to perform protein separation in advance. The high initial costs and the excessive maintenance still prevent the widespread use of these instruments despite their impressive performance.

More widespread is the “bottom-up” approach for which proteins are first degraded to small peptides using an endoprotease like trypsin. Those peptides are then separated by HPLC (high performance liquid chromatography) and analysed in a mass spectrometer, e.g. an ion trap or a MALDI-TOF. The online coupling of peptide separation by liquid chromatography to mass spectrometry is often referred to as LC-MS^n. The first stage or MS records the molecular ion without fragmentation whereas all further stages rely on the fragmentation of the entire molecule (MS^2) or of molecule fragments (all remaining stages). Subsequently, protein sequences are deduced from identified peptides matching to the proteins primary sequence. The bottom-up approach has the advantages that peptides can easily be separated in-line with the mass spectrometer and that it can be performed with more affordable instruments.

The identification of peptides can be performed in three ways: either complete *de novo* sequencing, database searching, or partial *de novo* sequencing followed by database searching. While initial experiments were still laborious and relied on manual *de novo* sequencing, during which the peptide sequence was deduced from peptide fragment masses by hand, nowadays the information of sequenced genomes and sophisticated database search algorithms are increasingly used to speed up the identification process [13-16]. For protein identification by automated database searches, mass spectra are matched against a database that stores protein or gene sequences of the organism under investigation [17]. Both approaches have their merits and drawbacks.

Due to the rather low mass accuracy and low resolution, *de novo* sequencing can not easily be performed with mass spectra generated by ion traps, which are the traditional “workhorse” mass spectrometers in proteomics. Additionally, this technique, when not performed on an instrument with high mass accuracy, fails to be able to distinguish between amino acids like lysine and glutamine and between specific amino acid combinations. On the other hand, database searching is not a stringent identification method, but rather a homology search. If the sequence to be searched is not in the database (because of a posttranslational modification, a splice variant, a mutation, or some error that occurred during the sequencing process) no correct match will be obtained.

The first experiments in plant proteomics aimed at identifying as many proteins as

possible in one sample. However, recent studies focus increasingly on the characterisation of protein modifications and on a combination of identification and quantification.

I.2.2. Quantitative proteomics

Quantitative proteomics tries to quantify as many proteins in a sample as possible. It is used to differentiate between healthy or wild type and disease, treated, or mutant forms of an organism. The approaches used for quantitative proteomics may be divided into *in vivo* and *in vitro* labelling approaches, stable isotope dilution techniques and label free approaches. However, they all involve two-dimensional gel electrophoresis and/or mass spectrometry to reveal the quantitative differences between samples. For the mass spectrometry based approaches proteins are usually digested with a protease and the resulting peptides are separated via affinity chromatography.

Quantification can then be achieved by integration of peptide ion peaks or by simply counting the sum of all spectra of different peptide ions and of their repetitive measurements belonging to the same protein (spectrum count). Traditionally the laborious peak integration is used, however, the easier to perform spectral count emerges as a good alternative [18, 19]. Since many of the methods available for protein quantification have not been applied to plants most of the examples discussed below deal with their application on non-plant organisms.

In vivo labelling approaches

These approaches rely on metabolic labelling of proteins. Radioactive or stable isotope versions of naturally occurring metabolites are introduced into an organism e.g. by feeding it with ^{35}S methionine [20], ^{15}N containing compounds [21] or with deuterated or ^{13}C labelled amino acids (stable isotope labelling by amino acids in cell cultures; SILAC [22, 23]). Proteins of the sample of interest are mixed with the control sample. Quantification can be achieved by 2d gel electrophoresis and imaging in the case radioactive labelling was employed but is most often performed by digesting the proteins followed by mass spectrometry based peptide analysis. The peptides of the marked sample exhibit a mass shift compared to the control sample making it feasible to easily identify corresponding peptide pairs with mass spectrometry. Usually, peptide ion peaks are then integrated and ratios for every peptide pair are determined.

In vitro labelling approaches

In contrast to the *in vivo* approaches proteins are derivatised after extraction. There is a large amount of different gel-free in vitro approaches. Some of them rely on the labelling of

proteins (e.g. ICAT [24], ICPL [25], or the two dimensional gel based DIGE [26, 27]) while others introduce the label on the peptide level (e.g. ^{18}O labelling [28] or iTRAQ [29]) . In general, labelling on the protein level is preferable since the samples can then be mixed at an earlier stage compared to peptide labelling and thus error sources resulting from different sample treatment can be minimised. However, after labelling and sample mixture, the downstream process is in principle the same as in the *in vivo* approach.

Stable isotope dilution techniques

In contrast to the approaches discussed so far, the approach relying on stable isotope marked standard peptides does not require the sample itself to be modified. Instead, peptides of the protein(s) of interest are synthesized as ^{18}O, deuterated, or ^{13}C marked versions of the natural peptides and spiked into the sample (e.g. [30]). The comparison of the spiked peptide to the naturally occurring version following mass spectrometry based analysis delivers the quantitative information. Since known concentrations are spiked absolute quantification is feasible with this approach.

Label-free approaches

A cost effective and comparably simple alternative to the above mentioned methods are label-free approaches. These rely on repetitive measurements of the samples of interest without the introduction of any label. Differences in samples can be visualised either using 2d electrophoresis and/or by mass spectrometry. In mass spectrometry based methods standards may be added to the sample to correct for run-to-run inconsistencies like retention time shifts. Alternatively, endogenous peptides that do not change in abundance from sample to sample can be used for standardisation. Peptide peaks or spectrum counts are compared between samples and the repetitive measurements deliver average values with acceptable standard deviations [31-35]. A variant of such an approach can be applied to quantify protein modifications e.g. phosphorylation. It makes use of unmodified peptides derived from the protein of interest for standardisation and served to determine the degree of phosphorylation for sucrose-phosphate synthase incubated *in vitro* with different plant extracts (discussed in detail in chapter IV).

I.2.3. Targeted and non-targeted approaches

With the availability of broad scale methodologies two different kinds of approaches became feasible: the hypothesis-driven or targeted approach and the non-targeted approach. The hypothesis-driven approach focuses on a well defined target and is usually restricted to a

small number of analytes. For example a hypothesis is made under which circumstances a certain protein is phosphorylated. This hypothesis is then tested and confirmed or discarded. Contrarily, for the non-targeted approach no clear target is defined and the number of analytes is usually higher than in the non-targeted approach. The aim of such an approach would be for example to identify all phosphorylation sites in a given sample.

Both approaches have their advantages and pitfalls. The hypothesis-driven approach seeks to gather additional information about a specific set of proteins and is therefore lacking the possibility to identify new and completely unexpected members of the system under investigation. On the other hand it is an approach that focuses on a well defined question and thus has the capacity for gaining important detailed information on this very issue. The non-targeted approaches often lack important background information and thus the results may be difficult to interpret. However, it can reveal unexpected coherences and interactions, which may provide the basis for hypothesis-driven approaches [36].

This means that both approaches need to be combined to learn most about a specific biological question. The non-targeted approach delivers the raw data and the hypothesis-driven approaches provide the fine tuning. I found both approaches very useful in gaining insight into the different aspects of phosphorylation and its regulation. While the major focus of my work relied on a non-targeted approach to gain insight into the phosphoproteome of *A.thaliana* and *C.reinhardtii*, I chose a hypothesis-driven approach to investigate particular phosphorylation events of established biological importance, such as the phosphorylation of phosphoenolpyruvate carboxylase and of sucrose-phosphate synthase in *A. thaliana*.

I.3. Post-translational protein modifications in plants

The term 'protein post-translational modification' describes the covalent attachment of chemical groups to a protein or protein processing by cleavage. These modifications increase the diversity of proteins tremendously and contribute to the complexity of proteomics. There are over 300 types of modifications known [37]. However, only some of them have been reported to play crucial roles for protein function in general [38] and especially in plants [39]. Ubiquitination (the addition of ubiquitin) is traditionally associated with targeting a protein for degradation. Recent findings, however, suggest a more widespread role of this modification including the response to abiotic and biotic stress [40, 41]. The addition of GPI (glycosylphosphatidylinositol) anchors is responsible for anchoring proteins in the cellular membrane and may play a role in various processes including cell signalling and recognition

[42, 43]. Recently, protein S-nitrosylation was reported for *A. thaliana* after NO treatment suggesting that this protein modification functions in plants in analogy to the animal system where it forms part of a NO signalling pathway [44]. Also, protein thiolation occurs in plants. It is increased under oxidizing conditions and seems to be involved in stress responses [45]. Protein acetylation and methylation were shown to influence the activity of plant histones [46] and methylation plays a role in the assembly of rubisco [47].

The examples discussed so far show that a variety of cellular functions rely on protein modifications. However, of all known posttranslational protein modifications phosphorylation is the most extensively studied one. It influences all layers of cellular organisation and regulation and is the main research topic of this thesis.

I.3.1. Protein phosphorylation

Since the breakthrough discoveries of reversible protein phosphorylation performed by kinases and phosphatases (Edmond H. Fischer and Edwin G. Krebs 1992 Nobel Prize in physiology or medicine [48]), research on protein phosphorylation in general and on the dynamic interplay between kinases and phosphatases has been gaining more and more interest. It has extended from initial studies describing single phosphorylation/dephosphorylation events to the complex regulations involving multisite phosphorylation in signalling cascades, which are accepted as major regulatory principle of life.

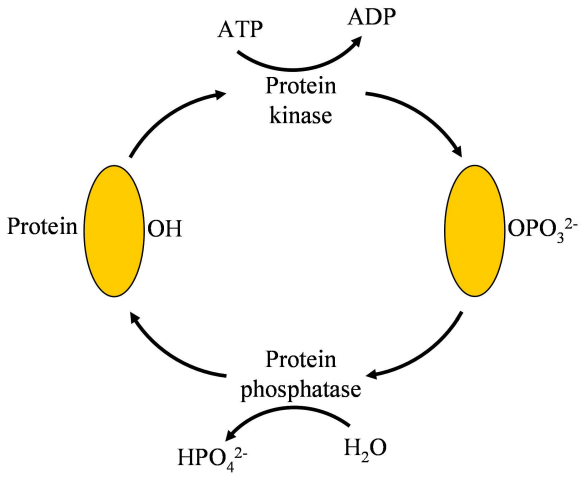


Fig. 3: Dynamic interplay between kinases and phosphatases on hydroxyl groups in reversible protein phosphorylation.

In earlier years research was restricted to only a few model systems since the methods for broader scale analysis were lacking. However, because of the importance of protein phosphorylation new methods for the analysis of this remarkable modification have been developed continuously. With the ability to enrich and sensitively detect and quantify phosphorylated proteins and peptides, as well as to pinpoint phosphorylation sites using mass spectrometry, the methods for a more thorough investigation of protein phosphorylation and their dynamics in general are now available. Many vital processes like cell trafficking, cell signalling, cell cycle control, enzyme

activation/deactivation, cation binding, degradation, apoptosis/defence, and others were found to be directly connected to protein phosphorylation [49, 50]. It can occur on several different amino acids including the well studied serine, threonine, tyrosine, histidine, and aspartic acid phosphorylation. Phosphorylation of the hydroyamino acids serine, threonine, and tyrosine occurs via the hydroxyl groups and the resulting amino acids are reffered to as *O*-phosphates or *O*-phosphomonoesters. Histidine phosphorylation leads to the formation of *N*-phosphates or phosphoramidates and is predominantely found in signal transduction systems of bacteria but also of eukaryotes [51]. These systems often rely on a histidine / aspartic acid phosphorylation relay that includes the transfer of the phosphate group from histidine to aspartic acid under the formation of acylphosphate. Little is known about phosphorylations located on lysine and arginine (*N*-phosphate), glutamic acid (acylphosphate), and cysteine (*S*-phosphate or *S*-phosphothioester) [52-54]. In eukaryotes phosphoserine, -threonine, and -tyrosine are estimated to occur at a relative abundance of ~90 % to ~10 % to ~0.05 %, respectively [55] and phosphohistidine was proposed to be 10 to 100-fold more abundant than phosphotyrosine [54, 56]. However, detailed information about the relative abundance of the different kinds of protein phosphorylations is difficult to obtain due to the chemical characteristics of these compounds. Some of these phosphoaminoacids like phosphohistidine and phospholysine are acid-labile but stable under basic conditions while for others like phosphoserine and threonine the contrary is true. Acylphosphates are generally highly reactive and labile in most forms of extreme conditions like acid or alkali [53].

Compared to single site phosphorylation, multisite phosphorylation enables the cell to develop and maintain complex regulatory pathways contributing to the different levels of hierarchal organisation. Most of the work has so far focused on mammals and bacteria and investigation of protein phosphorylation in plants has been considerably less thorough. Nevertheless, the examples described below cover a wide range of biological processes involving protein phosphorylation emphasizing the importance of this posttranslational modification also in plant biology.

Signalling/protein-protein interaction:

Phosphorylation has been observed in a wide variety of plant proteins and is described to play a role in different aspects of protein regulation. An important process, which depends on phosphorylation in plants, is phytochrome signalling. Phytochromes are proteins mediating biochemical processes in response to light. They are produced in an inactivated form and are activated upon light adsorption followed by a conformational change. Signals are then

transduced either indirectly or directly by binding of the activated phytochrome to transcriptional activators. Three phosphorylation sites have been identified on phytochrome using different methods including radioactive labelling, Edman degradation, mutation analysis and mass spectrometry [57-60]. Phosphorylation of the activated form is described to attenuate light signalling processes while dephosphorylation leads to enhanced photoresponsiveness [61]. Phosphorylation is reported to inhibit protein-protein interactions between phytochrome and its putative transducer proteins NDPK2 (nucleoside diphosphate kinase 2) and PIF3 (a basic transcription factor) [62] thereby interrupting the signal transduction process.

Other signal transduction processes are also directly linked to phosphorylation. Brassinosteroids, which regulate plant growth and development, are reported to lead to multiple phosphorylations of Brassinosteroid-insensitive1 (BRI1) and BRI-associated receptor kinase1 (BAK1) as was suggested by IMAC coupled to LC/MS/MS and partly supported by mutation analysis [63]. Phosphorylation of these proteins seems to be important for their kinase activity and might be linked to heterodimerization.

Symbiosis:

A special case of signal transduction is the one involved in interspecies contact. The establishment of symbiosis with nitrogen fixing Rhizobia and phosphate aquiring arbuscular mycorrhyzal fungi was shown to depend on multisite phosphorylation by *Yoshida et al.* [64]. By mutation analysis and mass spectrometry they determined phosphorylation sites in SYMRK (symbiosis receptor kinase), which is a RLK (receptor like kinase) crucial for symbiosis in plants [64, 65]. Activation of this kinase depends on the phosphorylation of at least two phosphorylation sites and good indications exist for the involvement of a third site.

Photosynthesis:

Phosphorylation also plays a major role in photosynthetic and associated processes. The thylakoid soluble protein of 9 kDa (TSP9) was characterized as a phosphoprotein by Carlberg *et al* [66]. They identified three phosphorylated threonines using mass spectrometry and showed that redox dependent phosphorylation of these sites leads to release of the protein from the membrane. Many other proteins involved in photosynthesis were also shown to be phosphorylated but for most of them only single phosphorylation sites are known [67, 68].

However, Vener *et al.* described a light dependent phosphorylation of PsbH (Photosystem II reaction center H protein) at two threonines and phosphorylation of PsbH was reported to increase its stability [67, 69].

Cell-to-cell trafficking:

Selective phosphorylation of the so-called tobacco mosaic virus movement protein (TMV-MP) also influences the spreading of tobacco mosaic virus in some tobacco species. Phosphorylation sites were identified by replacing putative phosphorylation sites in recombinant movement protein with alanine and glycine and subsequent determination of the absence or presence of phosphorylation using radioactive ATP *in vitro* [70]. Trutnyeva *et al.* then replaced the identified serine and threonine phosphorylation sites in a sequential manner with negatively charged amino acids and compared virus spreading for each mutant [71]. While the phosphorylation of TMV-MP at one specific serine seems to stimulate cell-to-cell movement of the virus, phosphorylation of two other sites has the opposite effect. The authors conclude that this multisite phosphorylation event may proceed in a sequential manner: first, phosphorylation leads to a spreading of the virus (a benefit for the virus) but further phosphorylation events then restrict the harmful infection (a benefit for the host plant).

Growth control:

Phosphorylation of ribosomal protein S6 at several amino acids at its carboxyterminal end is well known to be dependent on insulin/IGF (insulin-like growth factor) signalling in animals [72]. It is required for the activity of S6 leading to the translation of 5' TOP mRNAs (mRNAs which contain an oligopyrimidine tract at their 5' transcriptional start site) coding for proteins of the protein synthetic apparatus. In plants, multisite phosphorylation of specific sites in S6 in analogy to the animal system has been reported for maize using ^{32}P labelling and Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) [73]. Here, following phosphoprotein enrichment and mass spectrometry based peptide analysis I could identify an S6 phosphopeptide in *A. thaliana* belonging to the c-terminal end of the protein (see chapter III). Phosphorylation of *A. thaliana* S6 ribosomal protein was also reported by Chang et al. who subjected enriched ribosomes to beta elimination and mass spectrometry [74]. Furthermore, studies exist describing aspects of a signal transduction pathway in plants similar to the animal pathway leading from membrane perception of mitogens to S6 phosphorylation and growth control [73, 75-79]. Thus it seems that this important pathway, including multisite phosphorylation on specific residues at the c-terminus of S6, is indeed conserved in plants.

Enzymatic activity:

Of high importance and well described is the impact of protein phosphorylation on enzymatic activity. A classic example is pyruvate dehydrogenase (PDH), which forms part of the pyruvate dehydrogenase complex (PDC). This key enzyme, which converts pyruvate to acetyl-CoA and produces NADH has been shown by mutagenesis approaches and radioactive labelling to be phosphorylated in mammals on multiple serines. Phosphorylation leads to deactivation of the enzyme and might play a role in the formation of the PDC(for review see [71]). In plants, however, the situation is more complex since they posses not only a mitochondrial PDC (mt PDC) but also a plastidic PDC. No phosphorylation sites are reported for plastidic PDC so far, but two sites (one homologous to mammalian site 1 and a second serine one upstream of the mammalian site 2) have been confirmed in pea by mass spectrometry [80]. Phosphorylation of plant PDH leads to its inactivation as in the case of the mammalian PDH [81]. Future studies are likely to reveal the exact regulation of plant mt PDH by multisite phosphorylation.

Bykova *et al.* found multisite phosphorylation of formate dehydrogenase in potato tubers using a combination of ^{32}P *in vitro* labelling, IMAC (ion metal affinity chromatography) and mass spectrometry. They showed that phosphorylation led to deactivation of the enzyme and phosphorylation was inhibited by the addition of NAD^+, formate, and pyruvate in a manner similar to pyruvate dehydrogenase [82]. However, no detailed studies about the functional influence of individual phosphorylation were performed.

Phosphoenolpyruvate carboxylase (PEPC) is also recognized as a metabolic key enzyme. PEPCs catalyze the formation of oxalacetate from phosphoenolpyruvate (PEP) and bicarbonate (HCO_3^-). In this reaction Mg^{2+} and ATP are used as cofactors and inorganic phosphate (Pi) is formed and the product is readily converted to malate by malate dehydrogenase. Up to date, PEPCs have only been identified in photosynthetic organisms including green plants, algae, and bacteria. In C4 and CAM (crassulaceaen acid metabolism) plants PEPCs can serve the efficient prefixation of carbon dioxide thus accounting for the deficiencies of ribulose 1,5-bisphosphate carboxylase/oxygenase (rubisco) while in C3 plants the major task is the formation of oxalacetate and malate to replenish the tricarboxylic acid cycle in the so-called anaplerotic reactions.

The activity of some PEPCs is highly dependent on phosphorylation of a specific serine located in a conserved part of the N-terminus with the recognition motif (E/D)(K/R)XYS*IDAQLR [83]. For C3 plants less information is available compared to C4

and CAM plants but regulation by phosphorylation is also documented [84-86]. For *A. thaliana* it was reported that phosphorylation of this enzyme is light dependent and adverses the effects of inhibitory metabolites including malate [87]. However, while phosphorylation of an N-terminal serine was suggested in this study no *in vivo* phosphorylation sites of *A. thaliana* PEPCs was experimentally determined. In fact, to my knowledge a PEPC phosphorylation site, has never experimentally been pinpointed *in vivo* in any C3 plant. Therefore, I investigated whether PEPC is indeed phosphorylated at the putative site in *A. thaliana* (chapter III).

One of the most extensively studied examples for plant protein phosphorylation is sucrose-phosphate synthase (SPS). This metabolic key enzyme catalyzes the formation of sucrose-phosphate from fructose-6-phosphate and UDP-glucose. Three phosphorylation sites have been described for SPS in spinach and tobacco using a wide variety of methods including radioactive labelling, site directed mutagenesis, and mass spectrometry. Phosphorylation clearly influences the activity of this protein. One phosphorylation site is located in a 14-3-3 binding domain; another is responsible for light/dark modulation and a third site has been reported to be involved in osmotic stress response [88-91]. Interestingly, phosphorylation of the light/dark modulation site is responsible for deactivation of the enzyme while phosphorylation of the osmotic stress response site leads to the opposite effect.

Very recently, using a mass spectrometry based approach; a temperature dependency on the phosphorylation of this important enzyme was suggested (see chapter IV).

Aim of this thesis

The above examples give an overview on the variety of functions protein phosphorylation can have in plant cell biology and underline the importance of this remarkable posttranslational modification. Since the methods available at the beginning of this thesis were not sufficient to efficiently investigate plant protein phosphorylation in complex mixtures it was the aim of this thesis to develop and refine methods for non-targeted as well as targeted analysis of serine/threonine protein phosphorylation in *A. thaliana* and *C. reinhardtii.* These novel methods should be applied to solve biological questions on the regulation of plant primary metabolism. More specifically, a major focus was put on the development of a robust and affordable technique capable of efficiently enriching for phosphoproteins out of complex plant protein mixtures, and on a novel strategy relying on mass spectrometry to quantify changes in phosphorylation of metabolic key enzymes in a

robust manner. In addition, a prominent mass spectrometry based method widely used to determine phosphorylation sites (neutral loss driven MS^3) should be evaluated with respect to its robustness and applied to determine novel phosphorylation sites in *A. thaliana* and *C. reinhardtii*.

The following chapters intend to familiarize the reader with the merits and pitfalls of the most prominent approaches used to study protein phosphorylation and to explain how each of these methods can be used to gain insight into the complex world of this important covalent modification. Contributions of the author to the relatively new field of plant phosphoproteomics are presented and summarised in the final paragraphs of each chapter.

Chapter II: Phosphoprotein & phosphopeptide enrichment

Due to the low abundance of phosphorylated proteins, estimations reach from 10 to 30 % of all proteins phosphorylated at any given time [92, 93], the enrichment of phosphorylated proteins or peptides is of general interest for researchers working on phosphorylation. Indeed, the substoichiometry of phosphorylated protein species (i.e. the low number of phosphorylated proteins compared to their non-phosphorylated counterparts) in complex mixtures is a major analytical challenge in protein phosphorylation analysis [94] and often provides the basis for research on *in vivo* phosphorylation. Because of its high importance much work has been dedicated to the development and improvement of methods for phosphopeptide and to a lesser extent for phosphoprotein enrichment in the last years. Strategies for phosphopeptide and phosphoprotein enrichment are appearing continuously and many of them had not been published when this work was started. In general there are two alternatives for the enrichment, namely the enrichment of phosphorylated peptides or of intact phosphoproteins.

Phosphopeptide enrichment has the advantage that the resulting peptide mixture is less complex compared to a peptide mixture representing a total phosphoprotein digest following phosphoprotein enrichment. However, when enriching for phosphorylated peptides three main disadvantages become obvious. Firstly, the number of molecules on the peptide level is necessarily much higher than on the protein level. Secondly, and even more importantly, the identification of the respective protein often relies on just one phosphorylated peptide following phosphopeptide enrichment. Finally, studies following the peptide enrichment step are restricted to sequence analysis and no further experiments on the structure or behaviour of the proteins are possible.

In general it is acknowledged that protein identifications by tandem mass spectrometry should be based on more than one peptide [95]. This is especially true for studies where peptide identifications are achieved using database search since this is more a homology search than a stringent identification process. The correct identification of phosphorylated peptides via mass spectrometry is often not as straightforward as is the identification of unmodified peptides [96]. Therefore, it seems astonishing that, compared to general proteomics studies, less stringent criteria are sometimes applied in phosphoproteomic studies.

One should assume that for a robust identification of phosphoproteins at least one unmodified peptide of the respective protein in addition to the phosphorylated peptide is necessary. Alternatively, the protein might be identified by other means than mass spectrometry (e.g. recognition by an antibody) to support the identification. However, because many researchers try to identify as many phosphoproteins as possible in a given sample this procedure is not followed in most large scale phosphoproteomic studies, and identification is often based on a single phosphopeptide per protein.

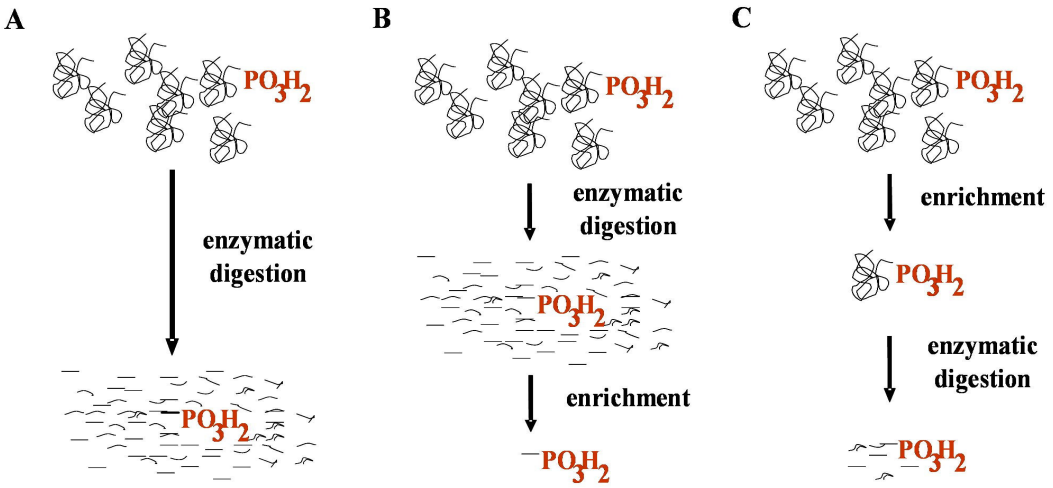


Fig. 4: ***Enrichment of phosphorylated proteins (C) leads to a compromise between reduction of sample complexity and reliable protein identification when compared to phosphopeptide enrichment (B) or no enrichment (A).***

Compared to phosphopeptide enrichment, phosphoprotein enrichment has to deal with different hurdles. First of all, peptides are far more soluble than proteins. Therefore, a much higher number of solvents can be used during the enrichment procedure. The acceptable pH range for peptides spans from near 0 to 14, but beta elimination of the phosphate group from serine and threonine can occur at pH higher than 12. For most proteins the pH at which they are soluble and stable is in the range from pH 6 to 12. In addition, the relative amount of phosphate is usually much higher in a peptide and the phosphate residues are often better exposed compared to a whole protein. Those characteristics make phosphate affinity purification far easier for peptides than for proteins.

However, enrichment of whole proteins is a compromise between no enrichment, which leads to a reliable identification of only very few phosphoproteins, and phosphopeptide enrichment that results in a less reliable identification of many phosphoprotein (figure 4). Because of the general lack of appropriate phosphoprotein enrichment methods in the plant

field at the beginning of this thesis a major focus was to establish a robust and cost effective method for the enrichment of phosphorylated proteins from plant tissue.

II.1. Antibodies and capture molecules occurring in nature

Antibodies directed against phosphoserine, -threonine [97] , and –tyrosine [98] have been used to generally enrich for plant proteins phosphorylated at the respective residues in immunoprecipitation experiments. It should be noted that the use of global antibodies directed against phosphoserine or phosphothreonine is usually associated with lower specificity compared to the anti-phosphotyrosine antibodies. This might be due to the better accessibility of the phosphate group located on tyrosine or because of the lower complexity of the structural nature of the threonine and serine antigens compared to tyrosine. Also, the production of antibodies still needs to be performed using animals like mice or rabbits – a questionable approach when effective alternatives to immunoaffinity enrichment can be used. Besides, antibodies remaining in the sample can complicate downstream analysis. However, following phosphorylation site identification one can also apply specific antibodies for a more focussed approach. These antibodies, directed against regions surrounding a specifically phosphorylated residue, are usually quite selective and can be effectively be used to enrich for a certain protein of interest. The drawbacks of this approach are that the exact residue has to be known and that the generation of these antibodies is expensive and time consuming.

In addition to antibodies there are several different protein domains designed by nature, which are capable of binding phosphorylated proteins and peptides. While SH2 (Src Homology 2) and PTB (phosphotyrosine binding) domains recognise predominantly tyrosine phosphorylation, 14-3-3 proteins as well as WW domains bind to proteins phosphorylated on serine and threonine, and FHA (forkhead associated) domains show a preference towards threonine phosphorylation [99]. In principle all of these domains or the respective proteins can be used for the enrichment of different sorts of phosphorylated peptides. In fact, immobilised 14-3-3 proteins have been used for the purification of serine/threonine phosphorylated plant proteins [100, 101]. As described, each of these domains/proteins only recognizes a subset of the total pool of phosphorylated proteins and experimental setups relying on these domains are therefore not as widely applicable as methods relying on more general affinity mechanisms. A combination of a more general affinity chromatography followed by different more restrictive approaches could however, further facilitate the analysis of the phosphoproteome.

II.2. Strategies relying on chemical derivatisation

Besides the non-modifying purification procedures described so far, methods exist that make use of chemical derivatisation strategies for the enrichment of phosphorylated proteins and peptides. Two main approaches have been described. The first one consists of beta elimination followed by Michael addition, where the phosphate group of phosphoserine or phosphothreonine (phosphotyrosine can not be derivatised using this method) is replaced with a nucleophile that is suitable for selective enrichment [102-108]. The alternative is a strategy involving temporary carbodiimide coupling of the phosphate residues of p-serine/p-threonine/p-tyrosine to a solid phase, followed by washing steps to remove non-phosphorylated species [109, 110].

However, these techniques have their limitations. Due to the involved chemical processes, unwanted side reactions may occur. This is known for the approach involving beta elimination, which is carried out at basic conditions (pH 12-14). These reaction conditions can lead to the replacement of O-glycosides [111, 112], sulfonated residues [113], and even of hydroxyl groups, which are located on serine and threonine [114]. Several other artificial protein modifications are known to occur under highly basic conditions [115] and it is hard to see how all of these reactions can be excluded while preserving replacement of phosphorylated residues.

The carbodiimide approach has the distinct advantage of including phosphotyrosine in the analysis and that no side reactions are reported so far. The initial approach [109] involved several modification steps which increases the likelihood of sample loss and modification in each step. The refined method [110] involves methylesterification, which has also its drawbacks as described above (IMAC section). Nevertheless, the method seems to be promising and might find its way into plant phosphorylation research.

II.3. Immobilised Metal Affinity Chromatography (IMAC)

The most widespread method used for phosphopeptide enrichment is IMAC (immobilised metal affinity chromatography). This technique was initially developed by Porath *et al.* [116] and was originally used to separate all kinds of different proteins. It mainly relies on the attraction of a negatively charged amino acid residue to a a positively charged metal that is immobilised on a metal chelator matrix like iminodiacetic acid. Iron and gallium

are the most widely used metals for the enrichment of phosphorylated species. Since a phosphate group has a stronger net negative charge than any other amino acid residue, phosphorylated proteins/peptides are better retained on the matrix than their non-phosphorylated counterparts. IMAC has been used to separate plant derived phosphorylated proteins [117] as well as phosphopeptides [118, 119] from non-phosphorylated ones. However, there are some difficulties when applying IMAC. Histidine and aspartate phosphorylations are not accessible by classic IMAC enrichment since they are acid-labile and IMAC loading is usually conducted under acidic conditions. IMAC has in fact been used to enrich for acidic proteins and is prone to unspecific binding of proteins/peptides rich in glutamic and aspartic acid residues [120, 121]. To circumvent the unspecific binding problem peptides can be methylesterified thus converting glutamic and aspartic acid residues into their non-charged methyl esters (figure 5).

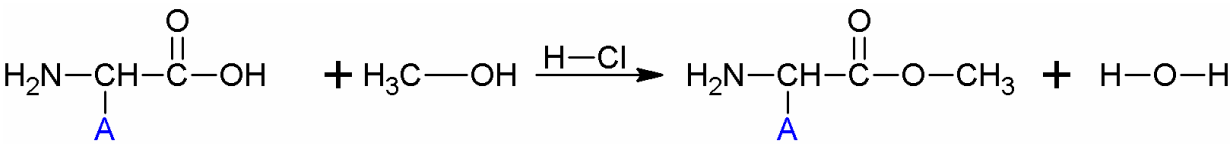


Fig. 5: ***Methylesterification on carboxylic residues of amino acids. A: any amino acid side chain.***

Unfortunately, this reaction is usually conducted under harsh conditions (exclusion of water at pH 0-1) and not free of side reactions such as deamination of glutamines and asparagines [122] making this method of methylesterification especially unsuitable for complex protein mixtures. However, it seems that the success of IMAC and methyl esterification also depends on the kind of protein sample and IMAC resin used [123].

Since IMAC has been used for the enrichment of phosphorylated proteins [124] and peptides it can in principle be applied to couple the advantages of both strategies. However, probably due to experimental difficulties, there is only limited literature describing a double enrichment both on the protein and peptide level. A recent description comes from Collins *et al.* who published a detailed protocol for this sort of sequential purification [125, 126] based on IMAC, and an analogous approach is described by Huang et al. [127]. There is no study describing this or a similar experimental setup for plant proteins. Nevertheless, recent advances in method refinement (i.e. [128]) might help to overcome some of the aforementioned difficulties of IMAC.

II.4. Metal Oxide Affinity Chromatography (MOAC)

Other inorganic approaches make use of metal oxides and hydroxides. This includes the recently developed methods based on zirconia and titania and the one I developed, which is mainly based on aluminum hydroxide. Since, oxides and hydroxides are closely related and in fact all of these compounds belong to the oxide class of minerals (see e.g. [129]), the general term metal oxide affinity chromatography (MOAC) may be used for these methods in analogy to IMAC.

II.4.1. Titania and zirconia

Recently, it was shown that metal oxides can also be used for the enrichment of phosphorylated peptides. In these studies affinity chromatography based on titania (titanium dioxide) [130-139] and zirconia (zirconium oxide) [140] was used. Online coupling of a titania precolumn and an anion exchange [133, 134] or reversed phase column [131, 135] in an HPLC (High Pressure Liquid Chromatography) setup has been shown to be useful in the selective analysis of phosphorylated peptides derived from proteolytic digests. Identification of the phosphopeptides was achieved by monitoring the UV trace [133, 134] or by using a mass spectrometer [131, 135]. Similarly, nanoparticles composed of Fe3O4/TiO2 core shell particles were used to specifically isolate and detect phosphopeptides [130]. Unspecific binding, which is also reported when using titania [131, 135], can be reduced by methylesterification [135] and the use of appropriate incubation buffers [132]. However, the use of a special incubation buffer is preferable because of the side reactions occurring during methylesterification (see above). While all of the studies named so far focussed on the analysis of standard proteins or animal tissue a very recent application comes from the plant field. In this study the authors used titania to enrich phosphopeptides of spinach stroma membranes and identified some new .phosphorylation sites in photosynthesis related proteins [141].

However, most of the many different enrichment methods described in the last paragraphs including recent applications (published in 2006), focus entirely on phosphopeptide enrichment. None of them reports the general enrichment and identification of phosphoproteins out of crude plant extract. Therefore, I developed a method relying on the affinity of phosphate residues towards aluminum hydroxide which is capable of filling the gap in methods for phosphoprotein enrichment out of complex plant protein mixtures.

II.4.2. Aluminum oxide and hydroxide

Aluminum hydroxide is a well known and widely used adjuvant in medicine. Adjuvants are additives that enhance the effectiveness of medical treatment by potentiating the immune response and functioning as a carrier for antigens [142, 143]. Because of this wide application, protein adsorption to aluminum hydroxide has been studied in considerable detail and some of those studies also investigate the binding behaviour of phosphorylated proteins to aluminum hydroxide [144-147]. In addition, aluminum oxide, a close relative of aluminum hydroxide, has been reported to exhibit a high and selective attraction to phosphorylated biomolecules [148]. The conclusion of these studies is that aluminum hydroxide and oxide both have a high affinity for phosphorylated biomolecules and that the affinity towards phosphate is considerably higher than for sulfate or nitrate. These characteristics make them good candidates for the enrichment of phosphorylated proteins and peptides out of complex mixtures. However, it is also clear from some of those studies that non-phosphorylated proteins also exhibit a considerable affinity to those matrices even if they are not phosphorylated.

Nevertheless, after careful optimisation a buffer recipe was developed that suppresses the binding of non-phosphorylated proteins while preserving the affinity for phosphorylated proteins. Figure 6 displays the separation of the two phosphorylated proteins ovalbumin and pepsin from the unphosphorylated protein bsa (bovine serum albumin).

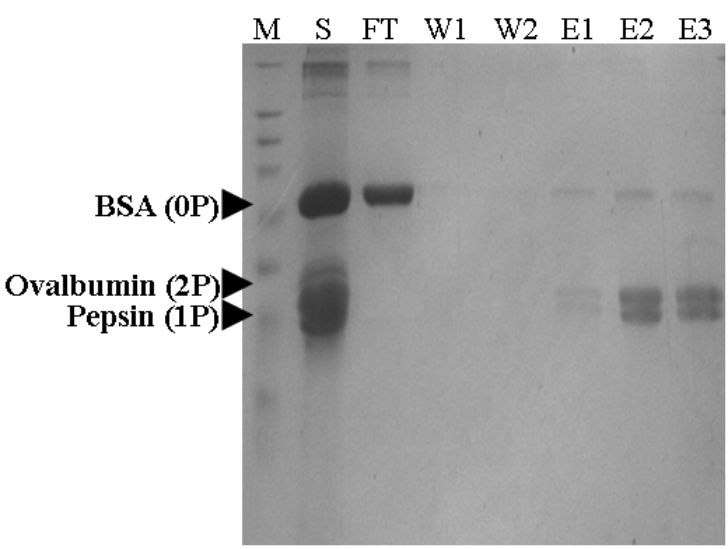


Fig. 6: Separation of the two phosphoproteins ovalbumin and pepsin from the non-phosphorylated protein bsa (bovine serum albumin). M: marker; S: sample (mixture of bsa, ovalbumin, and pepsin); FT: flow through (unbound sample); W1,2: wash; E1,2,3: eluate (bound and eluted sample).

The enrichment procedure is usually performed under denaturing conditions to decrease the threat of dephosphorylation and degradation events because of the action of endogenous proteases and phosphatases. Proteins are extracted using a mix of buffer, phenol, sodium fluoride (which serves as a phosphatase inhibitor), and beta-mercaptoethanol. After

precipitation proteins are resuspended in a special incubation buffer containing urea, subjected to affinity chromatography, and eluted using pyrophosphate (see appendix). So far, phosphoproteins from different sources including *C. reinhardtii*, *A. thaliana* cell cultures, seeds and leaves have been enriched (see this chapter and the relevant publications in Proteomics and Plant Methods).

To get a first glance at the phosphoproteomes of *A. thaliana* and *C. reinhardtii* subproteomes generated by the developed phosphoprotein enrichment were compared. Proteins of *A.thaliana* seeds, leaves, and cell culture as well as of heterotrophically grown *C. reinhardtii* were processed using the general extraction, enrichment, and protein separation procedure described in the appendix at the end of this thesis. Percentages of shared proteins were calculated by the formula: number of shared proteins / number of proteins in the sample with fewer proteins*100.

Table I: Overlap of proteins from different sources enriched by aluminum hydroxide:

Protein Overlap	***At seeds / At leaves***	***At cc / At leaves***	***At cc / At seeds***	***At cc / Cr***	***At seeds / Cr***	***Cr / At leaves***	***At cc / At cc***
Identified proteins	***181 / 83***	***234 / 83***	***234 / 181***	***234 / 137***	***181 / 137***	***137 / 83***	***98 / 75***
Shared proteins	***20***	***23***	***57***	***25***	***12***	***23***	***54***
% of shared proteins	***24***	***28***	***31***	***18***	***9***	***28***	***72***

***At: A. thaliana*, *Cr: C. reinhardtii*, cc: cell culture**.

Because of stringent filtering criteria (minimum three peptides per protein with Xcorrs 2.0, 2.5, 3.5 for a singly, doubly, or triply charged peptide, respectively) only the most abundant proteins were monitored. The minimum of three instead of two peptides per protein (compare introduction) was set because spectra were not evaluated manually. Overlap of *A. thaliana* and *C. reinhardtii* proteins was determined by searching the *C. reinhardtii* derived spectra against a *C. reinhardtii* database and subsequently identifying the *A. thaliana* homologues to these proteins by BLAST search. Interestingly, only a small percentage of proteins is shared between samples reflecting the high specialisation of the different tissues and organisms with respect to the abundant proteins (table I). However, even the overlap between two samples of the same type generated by shotgun analysis was below 100 %. This phenomenon is well known even for samples, which are not enriched in any way [19]. In fact, the percentage of overlap between samples before enrichment was almost the same as for

samples after enrichment (70%). Even though it can not be excluded that repeated analysis of biological replicates of all samples might lead to a different view on the data, repeated random sampling of some gel bands revealed consistent results. Phosphorylation sites for some but not all of the proteins were identified (see chapter two). While it is not possible to tell which of the proteins are definitively phosphorylated *in vivo* without the identification of phosphorylation sites, it is clear that subproteomes highly enriched for phosphorylated proteins are compared reflecting the specialisation of every organism and tissue type under investigation.

The protein categories that contribute most remarkably to the overlap are protein synthesis and photosynthesis which may mean that phosphorylation of proteins from these categories is especially dominant throughout the different samples, and is conserved across the tissue types and plant species under investigation.

Limitations of the developed method

As any method, the developed method for phosphoprotein enrichment is limited to some extent. The amount of protein needed for successful enrichment is high, mainly because of considerable sample loss during the enrichment procedure. This means that one needs approximately 1.5 mg protein sample to obtain 20 to 50 µg of protein in the enriched fraction (the amount depends on the initial phosphorylation degree and on sample complexity). Enrichment on the peptide level using aluminum hydroxide suffers to a bigger extent from this loss and even more material is needed. This can be explained by the stronger retention phosphopeptides have on the matrix as compared to whole proteins. Also, it was observed that phosphopeptide enrichment using aluminum hydroxide resulted in samples that could not easily be analysed using electrospray ionisation. A possible explanation for this phenomenon is that matrix particles were present in the samples and interfered with the electrospray process, and peptide purification with C18 material did not eliminate these complications. However, the problem was not observed when enriching for whole proteins probably because of the extensive cleaning procedures involving protein precipitation after enrichment. Because of enriching for phosphorylated proteins instead of peptides the number of identified phosphopeptides is small in comparison to phosphopeptide enrichment. To further investigate this problem, a double enrichment approach with aluminum hydroxide and titainia was employed and compared to a single enrichment procedure. However, it was found that the phosphopeptide enrichment with titania was not as specific as expected (see chapter III).

A general problem common to all enrichment methods is that the non-phosphorylated counterparts are usually discarded. However, this makes it impossible to gain the ultimate information on protein phosphorylation: the stoichiometry of phosphorylation (i.e. the ratio of a phosphorylated protein to its non-phosphorylated cognate). While many studies focus on the determination of phosphorylation sites and some on the comparative quantification of protein phosphorylation (i.e.: in which of the samples can we find more phosphorylation?) there are only limited reports on phosphorylation stoichiometry. Only analysis of phosphorylation stoichiometry delivers precise information on regulatory processes since these do not depend on the overall amount of phosphorylated protein but on the ratio of phosphorylated versus non-phosphorylated protein. In principle, one can analyse the flow-through and all washing steps of any enrichment procedure to determine this stoichiometry. This however, is usually not done because of technical limitations (e.g. a part of the sample is often retained by the matrix even after extensive washing and elution steps) and because of the time such an approach takes. On the other hand, as described before, the identification of phosphorylation sites is greatly hindered without enrichment. Consequently, enrichment is the first and very important step, which provides a solid basis for further phosphorylation research.

Appendix chapter II

Wolschin, F., Wienkoop, S. & Weckwerth, W. Enrichment of phosphorylated proteins and peptides from complex mixtures using metal oxide/hydroxide affinity chromatography (MOAC). ***Proteomics*** 5**, 4389-4397 (2005).**

REGULAR ARTICLE

Enrichment of phosphorylated proteins and peptides from complex mixtures using metal oxide/hydroxide affinity chromatography (MOAC)

Florian Wolschin, Stefanie Wienkoop and Wolfram Weckwerth

Max Planck Institute of Molecular Plant Physiology, Potsdam, Germany

A novel method termed metal oxide affinity chromatography (MOAC) of enriching for phosphorylated proteins and peptides based on the affinity of the phosphate group for $Al(OH)_3$ is presented here. When compared to commercial phosphoprotein-enrichment kits, this method is more selective, more cost effective and easily applicable to method optimization. The use of glutamic and aspartic acid in the loading buffer significantly enhances selectivity. Standard protein mixtures and complex *Arabidopsis thaliana* leaf protein extracts were tested for efficacy of enrichment. The method can be applied to proteins extracted using either mild or denaturing conditions. The same $Al(OH)_3$ material is suitable for the enrichment of phosphopeptides out of a tryptic digest of α-casein. Peptide phosphorylation was revealed by beta-elimination of phosphate groups. Enrichment and *in vivo* phosphorylation of *A. thaliana* leaf proteins were confirmed with Pro-Q diamond stain. Several of the phosphoprotein candidates that were identified by MS are known to be phosphorylated *in vivo* in other plant species.

Received: December 13, 2004
Revised: February 17, 2005
Accepted: March 24, 2005

Keywords:
Beta-elimination / IMAC / MALDI / Nano-ESI / Protein phosphorylation

1 Introduction

Phosphorylation is one of the most prominent post-translational protein modifications in living cells and its investigation is of key interest in the field of proteomics. Examples of regulation by phosphorylation abound: increasing or decreasing the biological activity of enzymes, helping proteins to travel between subcellular compartments, allowing interactions between proteins to occur, and labeling proteins for degradation. However, systematic knowledge about *in vivo* protein phosphorylation is lacking for most organisms. Since only 10–30% [1, 2] of cellular proteins are thought to be phosphorylated, enrichment facilitates their analysis. A number of strategies have been applied to separate phosphorylated proteins and peptides from the non-phosphorylated. One such strategy involves the use of antibodies specific for phosphorylated amino acids. This method is well established for proteins phosphorylated on tyrosine [3, 4] and has been used recently for the isolation of proteins phosphorylated on serine and threonine residues [5]. The technique, however, is limited due to the availability, high costs, and specificity of the antibodies.

Another commonly used strategy is IMAC. It relies on the affinity of the phosphate group to metal ions immobilized on a matrix such as agarose *via* acidic compounds like iminodiacetic acid (IDA) or nitrilotriacetic acid (NTA) [6]. IMAC is based largely on ionic interactions, so consequently peptides rich in glutamic and aspartic acid residues are often co-purified. Other amino acid residues such as cysteine and histidine may also interact with the IMAC material making selectivity for the phosphate group difficult [7]. This disadvantage can be partly overcome by converting carboxyl residues to methyl esters [8], but the process of methyl esterification is quite harsh and may lead to unwanted peptide modification [9].

Correspondence: Dr. Wolfram Weckwerth, Max Planck Institute of Molecular Plant Physiology, 14424 Potsdam, Germany
E-mail: weckwerth@mpimp-golm.mpg.de
Fax: +49 331 567 8134

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Using iron-loaded IDA coupled to agarose, Andersson and Porath [10] were able to separate three forms of phosphoovalbumin differing only in the extent of phosphorylation. However, no studies on nonspecific binding were performed. A similar procedure is now commercially available from BD Biosciences.

In another study, binding and elution of phosphorylated compounds (phosphoamino acids, phosphopeptides, phosphoproteins, and nucleotides) at high NaCl concentrations were investigated on an IDA column loaded with aluminum ions. The authors found significant discrimination towards the binding of phosphorylated compounds when applying 1 M NaCl [11], but no biological samples were tested.

An alternative to the BD™ kit is a kit developed by Qiagen, but no information is provided on details of the procedure, for instance on buffer compositions. Thus, optimization for complex biological samples like plant tissue is difficult or even impossible. Additionally, the commercial kits are quite expensive.

Other techniques for the enrichment of phosphorylated proteins consist of approaches for the chemical derivatization of phosphoproteins by beta-elimination and Michael's addition of different linkers with or without subsequent digestion and affinity purification of the derivative compounds [12, 13]. So far, these methods have only been demonstrated for case studies, and raise some important problems. First, the replacement works only for serine and threonine phosphorylated amino acids, leaving residues phosphorylated on tyrosine or other residues unaltered. Second and more important, side reactions may occur under the conditions of the replacement. Proteins and peptides may be degraded and hydroxide moieties on serine and threonine can also be replaced in some cases [14, 15]. In addition, glycosidic residues linked to serine or threonine are replaced, leading to indistinguishable detection of phosphorylation and glycosylation or other modifications on serine and threonine residues. One possibility is to remove the glycosidic residues (via glycosidase) or the glycosylated proteins/peptides (*via* lectin affinity chromatography) prior to the derivatization step, which should be carried out under the mildest conditions possible [16]. Another option is the combination of phosphoprotein/peptide enrichment and subsequent beta-elimination followed by Michael's addition [17, 18].

However, a major advantage of the beta-elimination is the exchange of the labile phosphate group with various stabile marker molecules giving better ionization efficiency in MS, and thus facilitating phosphorylation site identification and relative quantification [19–21].

Recently titania was used for affinity chromatography of phosphopeptides [22–25]. This technique seems to be a promising alternative to IMAC. Nevertheless, nonspecific binding is observed and the technique has to be developed further.

In the present study, a novel $Al(OH)_3$-based technique for the enrichment of phosphoproteins and phosphopeptides from sub-stoichiometric mixtures is described circumventing many of the aforementioned problems. Most important, due to the accessibility of the $Al(OH)_3$ material in large amounts, the procedure can easily be optimized for all kinds of biological samples. Here, we demonstrate the selective and sensitive enrichment of phosphoproteins from a complex protein sample from *Arabidopsis thaliana* leaf, thereby identifying candidates for *in vivo* phosphorylation. Furthermore, the same material is shown to be suitable for the enrichment of phosphopeptides.

2 Materials and methods

2.1 Chemicals

Imidazole, and sodium chloride were from Merck (Darmstadt, Germany). CHAPS was from Roth and Triton X-100 from Serva. $Al(OH)_3$ and all remaining chemicals were obtained from Sigma (München, Germany).

2.2 Model proteins used for method validation

Model proteins were selected based on the characteristics of their primary sequence. p*I* and molecular weight were calculated using online software available at http://scansite.mit.edu/calc_mw_pi.html. Sequences were derived from PubMed (http://www.ncbi.nlm.nih.gov/entrez/) or expasy (http://www.expasy.org/). Glucose oxidase was obtained from Roche, all other proteins were from Sigma.

2.3 Batch protein loading to $Al(OH)_3$

The binding of different model proteins to $Al(OH)_3$ was examined using a batch procedure. Different amounts of $Al(OH)_3$ were washed twice with an incubation buffer consisting of 30 mM MES, 20 mM imidazole, 0.2 M potassium aspartate, 0.2 M sodium glutamate and 0.25% CHAPS. In experiments with denatured proteins, 8 M urea was included in the incubation buffer. The pH of the buffer was adjusted to 6.1. Protein solutions (in incubation buffer) were added and incubated for 30 min. The matrix was washed with five to six volumes (in respect to sample volume) of incubation buffer, and elution was performed using a phosphate or pyrophosphate buffer (see below and results section).

The protein concentration in the fractions was determined using the Bradford assay [26] with ovalbumin as a standard.

2.4 Binding and elution of standard proteins

Of the three standard proteins, pepsin (1 phosphate), ovalbumin (0–2 phosphates), and casein (8 phosphates), 1 mL each of a 1 mg/mL solution was incubated for 30 min at room temperature with 80 mg $Al(OH)_3$. Bound protein was eluted in steps with different concentrations of a potassium phosphate buffer, pH 7.2. Incubation time for each step in this experiment was 1 min.

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

2.5 Comparison of different methods for the enrichment of phosphorylated proteins

For the comparison of the $Al(OH)_3$ matrix with the BD Biosciences phosphoprotein enrichment kit, 2.5 mL of a 1 mg/mL solution of a mixture of standard proteins was used for each experiment. Of the non-phosphorylated proteins conalbumin, lysozyme, and glucose oxidase, 700 μg of each was used, and of the phosphorylated proteins α-casein and ovalbumin, 200 μg of each was present in the mixture, resulting in a stoichiometry of about 16% phosphorylated *versus* non-phosphorylated proteins. The BD kit was used according to the manual with slight modifications as follows: incubation time was raised to 30 min (instead of 20 min) and the last 1 mL of the washing steps was saved for analysis.

The Qiagen column was used according to the manual, loading 25 mL sample diluted in the supplied loading buffer to give a protein concentration of 0.1 mg/mL.

For MOAC, 80 mg $Al(OH)_3$ matrix were washed twice with a total of 3.6 mL incubation buffer before loading the sample. The loading fraction (2.5 mL) was added and incubated for 30 min at 4°C on a rotator. The mixture was centrifuged for 1 min at 14 000 rpm and the supernatant was discarded. Subsequently, the matrix was washed with six volumes of incubation buffer (with respect to the sample volume), and once with 0.9 mL incubation buffer, and the supernatant was retained for analysis. After two more washing steps with 1.8 mL 10 mM HEPES buffer, pH 7.0 to remove excess salt, bound proteins were eluted by adding four times 0.9 mL 0.5 M potassium phosphate buffer, pH 6.0. For each elution step the matrix was incubated for about 10 min on a rotator at room temperature. Protein was precipitated with methanol/chloroform according to Wessel and Flügge [27] and fractions were analyzed by SDS-PAGE [28] after dissolving pellets in 1× SDS sample buffer (45 mM Tris, pH 6.9, 10% glycerin, 1% SDS, 0.01% bromophenol blue, and 50 mM DTT). Eluate and wash fractions were concentrated ten times compared to the sample and flow through before loading.

2.6 MOAC enrichment of phosphoproteins from a *A. thaliana* leaf protein sample

A. thaliana leaf proteins were extracted by adding a mixture of three volumes buffer-saturated phenol (15 mL) and one volume 50 mM HEPES-KOH, pH 7.2 containing 1% 2-mercaptoethanol, 40% sucrose, and 40 mM NaF (5 mL) to 2 g leaf material ground in liquid nitrogen. After mixing for 20 min at 4°C, protein was precipitated from the phenol phase (upper phase) with five volumes ice-cold acetone for 1 h at −20°C and dissolved in 1.5 mL incubation buffer (see above) containing 8 M urea. This buffer was also used for washing the matrix before use and in the washing steps (five steps with 1.6 mL and one step with 0.8 mL). Subsequently, 1.5 mL of a 0.5 mg/mL protein solution was used for incubation with 80 mg matrix. Elution was performed by incubation of the matrix with 800 μL 100 mM potassium pyrophosphate buffer containing 8 M urea, pH 9.0 for 20 min. Proteins were precipitated with methanol/chloroform prior to gel loading; 6 μg of the sample, the flow through, or the eluate in 1 × SDS buffer were loaded onto the gel. Staining of phosphorylated proteins was performed using the Pro-Q diamond stain (Invitrogen, Karlsruhe, Germany) following the manufacturer's instructions, and using the de-staining solution containing ACN. Phosphorylated proteins were visualized using a Dark Reader™ (Clare Chemicals). The same gel was subsequently stained with CBB after washing three times with water.

2.7 Nano-LC/MS/MS analysis of enriched phosphoproteins

The in-gel tryptic digest of phosphoproteins was analyzed using a monolithic RP nano column (30 cm × 100 μm) [29] with an LC-MS system comprising a Surveyor HPLC instrument with flow splitter and an LCQ Deca Xplus IT (Thermo Finnigan, San Jose). The tryptic digest was first loaded onto a peptide trap (Bischoff Chromatography, Leonberg, Germany) with a flow rate of 2 μL/min, and subsequently washed for 5 min with 100% solvent A (0.1% formic acid in water). After washing, peptides were eluted and separated using a flow rate of 200–300 nL/ min and a 0–80% gradient of solvent B (0.1% formic acid in ACN) for 90 min. The voltage was applied directly to the analyte solution using a T-piece. Eluting peptides were continuously analyzed by selecting the three most abundant signals of a survey scan (mass range m/z 500–2000) for sequential MS/MS fragmentation. The MS/MS spectra were searched against an *Arabidopsis* database ((http://www.arabidopsis.org/) using Turbosequest implemented in Bioworks 3.1 (Thermo Finnigan). Matches were filtered with Xcorr versus charge filter of Bioworks 3.1 with minimum Xcorrs of 2.0, 2.0, and 3.3 for singly, doubly, and triply charged fully tryptic peptides, a minimum Δcn of 0.1, and a minimum of two peptides per locus. Using DTASelect [30] , all redundancy of scans, peptides and proteins were removed. Additionally, spectra were evaluated manually using a DTASelect Graphical User Interface (GUI).

2.8 Enrichment of phosphorylated peptides

Bovine α-casein was dissolved in 25 mM NH_4HCO_3 (pH 8.2, 10% ACN, 5 mM $CaCl_2$) to a concentration of 1 mg/mL and digested with immobilized trypsin at 37°C overnight. An aliquot of 200 μL was dried in a rotary vacuum evaporator and redissolved in 500 μL of the incubation buffer described above but lacking 0.5 M NaCl. An $Al(OH)_3$ slurry, 75 μL, was added to this re-suspension and the mixture incubated for 30 min on a rotator at 4°C. After centrifugation (12 000 rpm, 2 min), the supernatant was discarded and the matrix was washed five times with 0.5 mL incubation buffer without

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

NaCl. Elution was performed in two steps by adding 75 μL sodium pyrophosphate, pH 8.3 (60 mM and 250 mM, respectively) and incubating the mixture at room temperature on a rotator for 10 min. After centrifugation (12 000 rpm, 2 min) the supernatant was saved for analysis and the gel slurry was discarded.

2.9 Beta-elimination of phosphate groups

Enriched fractions of α-casein were subjected to beta-elimination by mixing 10 μL of the eluate of the procedure described above with 90 μL saturated $Ba(OH)_2$ solution (0.15 M) for 1 h at 37°C. Then, 10-μL fractions were desalted with C18 zip-tips (Millipore) according to the manufacturer's instructions. Peptides were acidified by adding 5 μL 5% TFA to 20 μL of each fraction (1% TFA final concentration). Elution was performed from each zip tip with three times 3 μL H_2O/ACN, 50:50 containing 0.1% TFA.

2.10 MALDI-TOF analysis of peptides

From each desalted sample, 5 μL was mixed with 5 μL 2',4'6'-trihydroxyacetophenone (THAP) matrix (10 mg/mL), dissolved in 50:50 H_2O/ACN containing 0.1% TFA, and mixed at a ratio of 9:1 with ammonium citrate (50 mg/mL in H_2O); 1 μL of the mixture was used for analysis.

Spectra were acquired using a voyager Pro system using the following instrument settings: accelerating voltage 20 kV; grid voltage 83%; guide wire voltage 0.003; delay time 250 ns. Samples were measured in negative reflector mode.

3 Results

3.1 Optimization of the loading procedure using phosphoproteins/non-phosphoproteins with different potentials for specific/unspecific binding

To test if the binding of proteins to different matrices is primarily due to the phosphorylated residues and not to other residues, standard proteins with different characteristics concerning their phosphorylation state, their amino acid composition and their p*I* were chosen (Table 1).

Proteins were selected with different degrees of phosphorylation (pepsin, ovalbumin, α-casein) and non-phosphorylated proteins with p*I*s varying from about 5 (glucose oxidase) to over 9 (lysozyme). Glucose oxidase from *Aspergillus niger* was chosen since its p*I* (4.2) is lower than that of ovalbumin (a phosphoprotein with a p*I* of 5.19–5.06). Thus, glucose oxidase is prone to bind to a metal affinity column although it is not phosphorylated. Conalbumin and bovine serum albumin have a high number of glutamic acid, aspartic acid, and cysteine residues. In addition, the number of histidines is higher than in any other standard protein. Conalbumin is a protein that binds iron in nature and so may bind to IMAC columns loaded with iron. Since it is thought that ligand exchange may play a role in absorption of the proteins to the matrix, the number of serine and threonine residues was also examined. Conalbumin with 82 threonine plus serine residues, and glucose oxidase with 87, exhibit the highest numbers of hydroxyl residues and they are not phosphorylated.

Unspecific binding to the material should be recognized using this set of model proteins. This is most important in the context of comparing other enrichment kits such as the BD and the Qiagen kit.

Table 1. Selected model proteins and their characteristics[a)]

Protein	α-Casein	Conalbumin	Glucose oxidase	Ovalbumin	Lysozyme	Pepsin	BSA
Accession no.	CAA42516	CAA68468	AAF59929	AAB59956	AAL69327	P00791	NP_851335
Molecular weight	25146	77786	65770	43031	16248	41425	69326
p*I*	4.48	6.85	4.98*	5.19–5.06	9.36	3.97	5.82
No. of phosphates	8	0	0	0–2	0	1	0
No. of amino acids	214	705	605	386	147	386	607
Cysteine	1	**32**	3	6	9	7	**35**
Tryptophan	2	**10**	10	3	6	6	3
Histidine	5	12	**20**	7	1	3	**17**
Methionine	6	**12**	12	17	3	5	5
Glutamic acid	7	**45**	30	33	2	15	**59**
Aspartic acid	25	**46**	38	14	7	30	**40**
Serine	16	**49**	41	38	11	**48**	32
Threonine	6	**38**	41	15	7	27	35

a) Prominent properties are shown in bold.

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Different loading buffers were tested for their inhibiton of unspecific binding of the various non-phosphorylated amino acid residues to the $Al(OH)_3$ matrix. High salt concentrations (0.5 M NaCl) were also used to prevent ionic interactions but did not seem to have an effect on specificity. A pH of 6.1 and the addition of imidazole (a histidine analogue) inhibits the binding of histidines to the matrix. The addition of glutamic acid and aspartic acid salts suppresses the binding of glutamic acid- and aspartic acid-rich proteins to the matrix (see also Fig. 1). The optimized loading buffer system consists of glutamate, aspartate, imidazole, and CHAPS at pH 6.1 (see experimental section). This loading buffer was used for all subsequent analyses.

3.2 Elution profile of different phosphoproteins from $Al(OH)_3$

The three standard proteins, pepsin (1 phosphate), ovalbumin (0–2 phosphates), and casein (8 phosphates), were incubated separately with $Al(OH)_3$ and bound proteins were eluted in steps with different concentrations of a potassium phosphate buffer pH 7.2. Results are summarized in Table 2.

Table 2. Elution behavior of model phosphoproteins using MOAC[a)]

KP buffer	Pepsin	Ovalbumin	α-Casein
125 mM	++	+	−
250 mM	+	+	−
500 mM	−	++	+
1000 mM	−	−	++

a) ++, highest protein content of all eluates of the respective protein; +, protein eluted from the matrix; −, no protein detected in the eluate. KP-buffer, potassium phosphate buffer, pH 7.2

A correlation between the degree of phosphorylation and the strength of binding to the matrix was observed. α-casein eluted later and/or at higher phosphate concentrations in the elution buffer than ovalbumin and pepsin (Table 2).

Further investigations showed that the retention of phosphorylated proteins also depends on the amount of matrix used, the time used for elution, and the pH of the elution buffer. When using less matrix and an elution buffer of pH 9.0, more protein can be eluted at a lower elution buffer molarity. In addition, pyrophosphate buffers are more effective than phosphate buffers. This becomes especially important when trying to elute highly phosphorylated proteins. Using 90 mM pyrophosphate as elution buffer, the recovery of bound casein was about 75%. Recovery of the other bound phosphoproteins was nearly 100%.

3.3 Comparison of MOAC with commercial kits

A mixture of five standard proteins (conalbumin, glucose oxidase, ovalbumin, casein, and lysozyme, see Table 1) was used to compare the capabilities for phosphoprotein enrichment of the BD Biosciences phosphoprotein enrichment kit, the phosphoprotein purification kit from Qiagen and MOAC with $Al(OH)_3$ as the matrix. The phosphorylated proteins α-casein and ovalbumin represented about 16% of total protein (a concentration thought to represent the phosphorylation stoichiometry in living cells). To each matrix [BD, Qiagen or $Al(OH)_3$], 2.5 mg of the mixture was added.

The enrichment of phosphoproteins is shown using the BD kit in Fig. 1A. α-Casein, ovalbumin, conalbumin, and small amounts of glucose oxidase appeared in the eluate fraction. All applied proteins were present in the flow through, albeit to extents not directly proportional to their application levels. While α-casein was clearly visible, the ovalbumin level was low. Proteins found in the flow-through

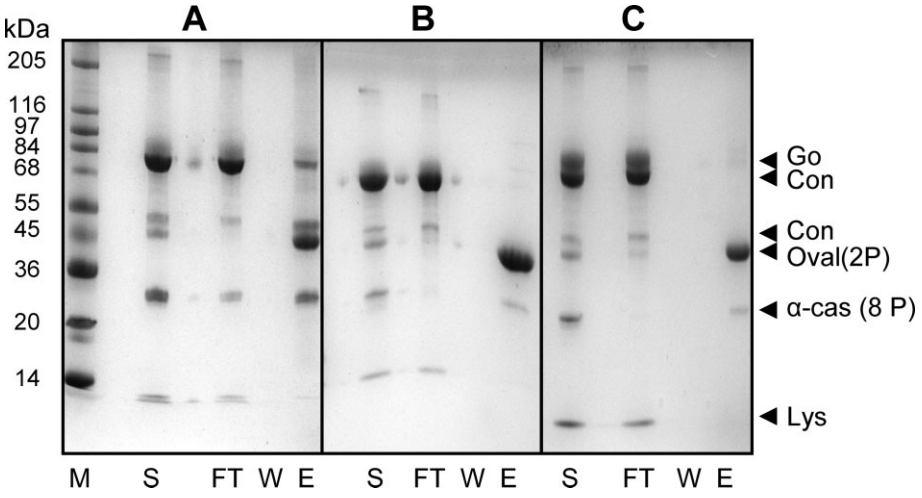


Figure 1. A mixture of five different model proteins was applied to different matrices to reveal differences in the selectivity for enrichment of phosphoproteins. (A) BD Kit; (B) Qiagen Kit; (C) MOAC [$Al(OH)_3$]. M, marker; S, protein mix; FT, flow through; W, last wash; E, eluate. GO, glucose oxidase; Con, Conalbumin; Oval, ovalbumin (2 phosphates); α-cas, α-casein (8 phosphates); Lys, lysozyme. Eluate and washing fractions are concentrated ten times.

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

were glucose oxidase, conalbumin, and lysozyme. A substantial enrichment of phosphorylated proteins was seen; however, non-phosphorylated proteins were also clearly visible in the eluate fraction. The phosphorylated α-casein was not completely retained by the BD matrix. The reason for that is not known since the binding capacity of the BD column was not exceeded. The Qiagen kit in Fig. 1B shows nearly no phosphorylated protein in the flow through, and almost exclusively phosphorylated proteins in the eluate. The amount of ovalbumin in the eluate clearly exceeded the amount of casein. The MOAC procedure is shown in Fig. 1C. In the flow through no α-casein and only a small fraction of ovalbumin were detected. The non-phosphorylated proteins were almost exclusively found in this fraction. Only minimal amounts of non-phosphorylated protein were be found in the eluate. The performance of MOAC with respect to the enrichment of phosphorylated proteins is comparable to the Qiagen kit.

3.4 Enrichment of *in vivo* phosphoproteins from an *A. thaliana* leaf protein sample

To minimize nonspecific binding due to protein-protein complexes or DNA/RNA-protein complexes, we established a protocol to enrich phosphoproteins out of a mixture of denatured proteins. Proteins were separated from DNA/RNA by phenol extraction [31] and incubated with $Al(OH)_3$. The loading fraction, the flow through, and eluate were submitted to SDS-PAGE and stained first with Pro-Q, which selectively stains phosphorylated proteins [32, 33], followed by CBB staining. While no phosphorylated proteins could be detected in the sample and the flow through (10 μg total protein loaded), the same protein amount of an eluate fraction gave a strong positive phosphoprotein signal in the gel with discrete protein bands (see Fig. 2A). CBB staining of the same gel shows significantly different protein pattern between loading/flow through fraction and eluate, indicating the specificity of the method for phosphoproteins out of a complex sample. A further important proof is that the band pattern observed in the Pro-Q stain of the eluate matches the protein bands observed after CBB stain thus inferring minimal binding of non-phosphorylated proteins to the matrix (Fig. 2B).

Several major bands were cut out of the gel, digested with trypsin, and analyzed by LC/MS/MS on an LCQ ion trap. Analysis with Bioworks software/Turbosequest™ and DTA-Select [30] resulted in the identification of several proteins, most of which are known from the literature to be phosphorylated in other plant species (see Table 3). The molecular

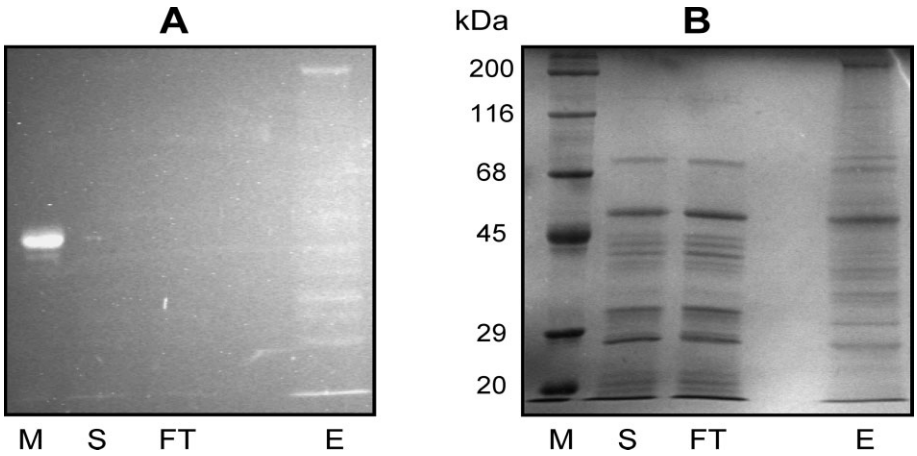


Figure 2. MOAC enrichment of *in vivo* phosphoproteins from an *A. thaliana* leaf protein extract. Phosphorylation was confirmed by Pro-Q phosphostain [32, 33] (A) and total protein content was visualized with CBB staining on the same gel; 10 μg total protein per lane was loaded (B). The marker band appearing after staining with Pro-Q corresponds to the phosphorylated protein ovalbumin. M, marker; S, sample; FT, flow through ; E, eluate.

Table 3. Phosphoproteins identified in a MOAC fraction of *A. thaliana* leaf protein extract

Protein	Plant species	Literature	*in vivo/in vitro* phosphorylation
Acidic phosphoprotein P2 (60S ribosomal subunit)	*Zea mays*	Bailey-Serres *et al.* [37]	*in vivo*
Ribosomal protein S6	*Zea mays*	Bailey-Serres *et al.* [37]	*in vivo*
ssU rubisco	*Cicer Arietinum*	Aggarwal *et al.* [38]	*in vivo*
lsU rubisco	*Spinacia oleracea*	Guitton *et al.* [39]	*in vitro*
Glyceraldehyde-3-phosphate dehydrogenase	*Spinacia oleracea*	Guitton *et al.* [39]	*in vitro*
Sucrose phosphate synthase	*Spinacia oleracea*	Winter & Huber [40]	*in vivo*
Rubisco activase	*Oryza sativa*	Komatsu *et al.* [41]	*in vitro*
Glutamine synthetase 2	*Nicotiana tabaccum*	Riedel *et al.* [42]	*in vivo*

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

weights of the identified proteins corresponded well with their mobility in the SDS-PAGE. An updated list can be requested from the corresponding author.

3.5 Enrichment of phosphorylated peptides derived from α-casein

Bovine α-casein was digested with trypsin, and peptides were enriched using MOAC. Phosphorylation was verified by mass and beta-elimination (Fig. 3). The matrix used was described recently as enhancing the signals of phosphopeptides in MALDI experiments [34]. After addition of $Ba(OH)_2$, we observed an immediate precipitation. This is probably due to an insoluble barium phosphate complex but because of the excess of added $Ba(OH)_2$ it did not interfere with beta-elimination.

Nearly all of the peaks found in the eluate fraction (Fig. 3B) could be assigned by mass to derive from phosphorylated peptides of either S1-casein or S2-casein. Enriched phosphopeptides include multiply phosphorylated peptides (up to 5 phosphorylation sites see Table 4). Peptides 1, 2, 3, and 14 were not identified. However, peptides 2, and 3 were confirmed by beta-elimination to be phosphorylated and peptide 1 is probably the dephosphorylated form of peptide 2. Interestingly, the product of beta-elimination was stable enough under the chosen conditions to enable analysis even without further Michael's addition. Most of the peaks observed after beta-elimination correspond to peaks found in the eluate minus one or multiples of phosphoric acid, thus further verifying successful enrichment. That some peptides were not identified after beta-elimination could be due to unwanted side reactions.

4 Discussion

This study shows that $Al(OH)_3$, in combination with a special incubation/washing buffer, can be used to selectively enrich for phosphorylated proteins and peptides from complex samples. The concept of using an adequate buffer system is essential to obtain high selectivity. This method does not rely on a chemical modification, and thus theoretically allows purification of native proteins as well as of denatured proteins. When using proteins in their native form, we observed nonspecific binding (data not shown). This is probably due mainly to the association of non-phosphorylated proteins with phosphorylated proteins or with phosphate-containing biomolecules such as DNA and RNA. The purification of denatured protein minimizes nonspecific binding of unphosphorylated proteins but also minimizes protein yield because of the inaccessibility of some phosphate groups in denatured proteins. When compared to commercial kits for enriching phosphorylated proteins, the advantage of the developed method is obvious: high selectivity at low cost. In the case of the BD Biosciences kit, it can even be said that MOAC seems

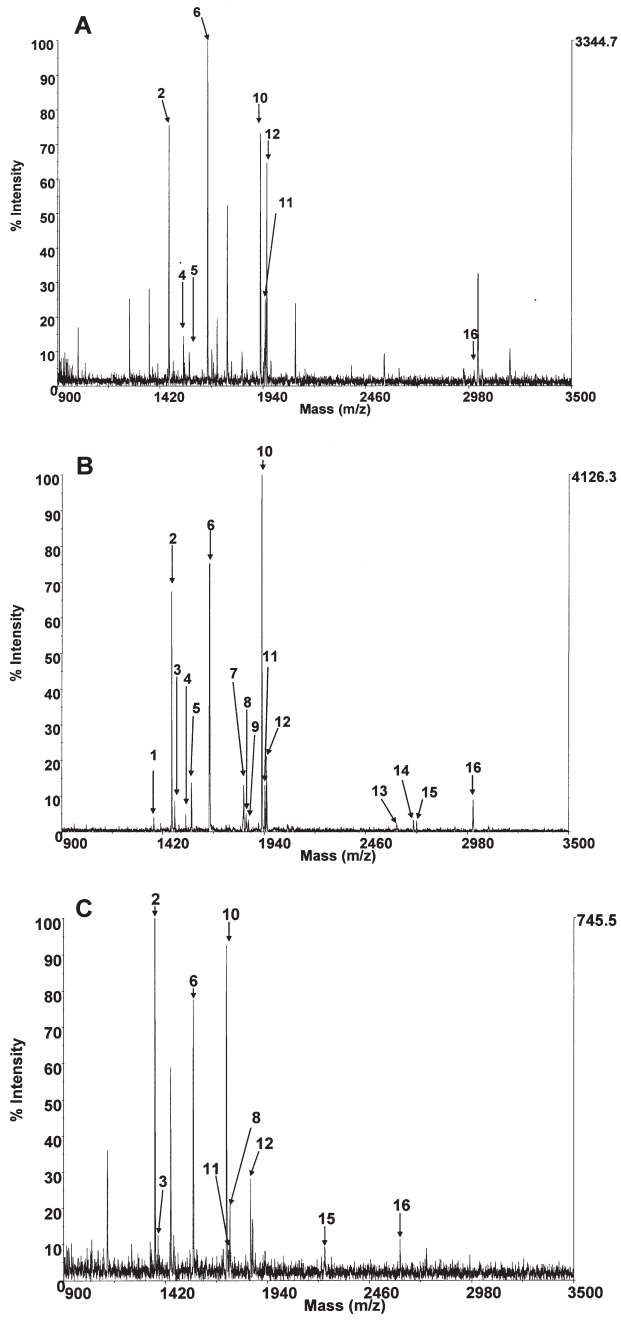


Figure 3. MOAC enrichment of phosphopeptides from a tryptic digest of α-casein. (A) MALDI-fingerprint of a tryptic digest of α-casein. (B) Eluate generated by MOAC from the same sample. (C) Beta-eliminated fraction from (B). Numbers indicate peaks found in the eluate and after beta-elimination. Unlabelled peaks in the fractions correspond to identified non-phosphorylated peptides and to unidentified peptides.

to be more selective than the commercial kit under the conditions described here. In addition, it can be applied to a wide range of samples, protein amounts, and to phosphorylated peptides. The main disadvantage is that highly phosphorylated proteins are quite strongly retained by the matrix and

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 4. Observed phosphopeptides of α-casein[a)]

Peptide sequence	No. in spectrum	Protein	Number of phosphorylation sites	Calculated $[M-H]^-$	Observed $[M-H]^-$	Observed mass after beta-elimination	Loss
DIGpSEpSTEDQAMEDIK	10	S1	2	1926.8	1926.7	1730.7	196
–[b)]	7	S1	1	–	1834.0[a]	–	–
DIGpSEpSTEDQAM*EDIK	11	S1	2	1942.8	1942.7	1746.7	196
–[b)]	8	S1	1	–	1846.6[a]	1748.7	97.9
VPQLEIVPNpSAEER	6	S1	1	1659.8	1659.7	1561.7	98
–[b)]	5	S1	0	–	1568.0[a]	–	–
YKVPQLEIVPNpSAEER	12	S1	1	1951.1	1950.9	1852.9	98
–[b)]	9	S1	0	–	1859.5[a]	–	–
EQLpSTpSEENSK	4	S2	2	1537.6	1538.3	1342.6	195.7
NANEEEYSIGpSpS-pSEEEpSAEVATEEVK	16	S2	4	3007.6	3007.0	2615.1	392.9
QMEAEpSIpSpS-pSEEIVPNpSVEQK	15	S1	5	2720.4	2720.2	2230.4	489.8
NTMEHVpSpS-pSEESIIpSQETYK	13	S2	4	2618.4	2618.7	–	–
–[b)]	2	–	1	–	1465.6	1367.6	98
–[b)]	1	–	0	–	1373.9[a]	–	–
–[b)]	3	–	–	–	1481.5	1383.6	97.9
–[b)]	14	–	–	–	2703.6	–	–

a) Molecular weight and the sequences of the peptides found in the eluate and after beta elimination are given (see spectra in Fig. 3B, C). *Indicates oxidation; p, phosphorylation. Some unstable phosphorylated peptides can be found in the phosphorylated, unphosphorylated, or dephosphorylated form in the eluate.

b) Putative phosphopeptides exhibit losses of only 92–96 Da probably due to metastable fragmentation [43].

quantitative elution becomes difficult. This problem, however, can be overcome at least partly by using the appropriate elution buffer. We tried several phosphate-based buffers, and a pyrophosphate buffer with a high pH (around 9) seems to be the most effective choice. MOAC also works with Al_2O_3 but not as well as with $Al(OH)_3$ (data not shown), which may have a higher affinity for phosphorylated compounds because of the structural differences or due to the electron saturation. The mechanism of the binding of phosphorylated biomolecules to Al_2O_3 has been described previously and seems to be a form of ligand exchange mechanism [35].

We suggest that release of phosphorylated compounds from the matrix during MOAC experiments follows a similar mechanism:

$$Al_2O_3H_2O_3P\text{-Protein} + H_2O_3P\text{-X} \leftrightarrow H_2O_3P\text{-Protein} + Al_2O_3H_2O_3P\text{-X} \quad (1)$$

$$Al(OH)_2\ H_2O_3P\text{-Protein} + H_2O_3P\text{-X} \leftrightarrow H_2O_3P\text{-Protein} + Al(OH)_2\ H_2O_3P\text{-X} \quad (2)$$

The usefulness of the method was shown not only with standard proteins but also with a complex leaf cell extract from *A. thaliana*. This enrichment for phosphoproteins makes further analysis much easier. We were able to identify many proteins known to be phosphorylated in different species. Interestingly, one of the proteins was a sucrose-phosphate synthase isoform, which is known to be phosphorylated in spinach. In a recent study, we used several synthetic peptides of the putative phosphorylation sites within the sucrose-phosphate synthase protein family for the analysis of kinase activities [19]. Based on these studies the different phospho-sites exhibited different behavior, indicating a variable multitsite phosphorylation dynamic of the protein. However, the protein isoforms are known to be expressed in hardly detectable amounts in *A. thaliana*. Therefore, the enrichment of this protein is an essential requirement for further investigations on the *in vivo* phosphorylation.

We evaluated the enrichment of phosphorylated peptides from a tryptic digest of α-casein using MOAC. It was possible to capture and release phosphopeptides with one up to five phosphorylation sites. The coverage of almost all known phosphorylation sites in α-casein in parallel with the absence of non-phosphorylated peptides demonstrates the selectivity and specificity of the MOAC procedure even for the enrichment of phosphopeptides.

Taken together with the results for columns packed with titania published by other groups [22–25, 36], we conclude that after optimization, affinity of phosphorylated peptides and proteins to metal oxides and metal hydroxides can be used for purification. For a schematic overview of the strategy see Fig. 4.

In addition, the conditions developed for specific binding in our experiments might help to improve IMAC, which still shows specificity problems. In the future, the enrichment of

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

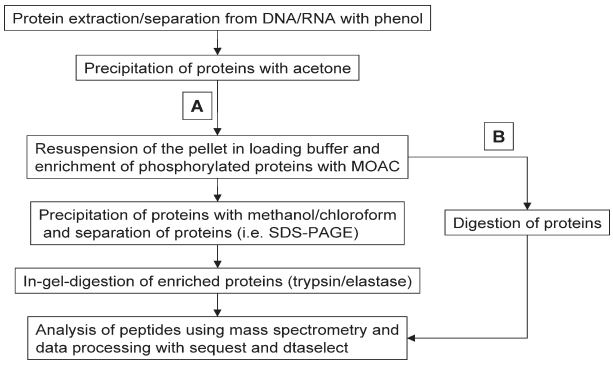


Figure 4. Strategies for the analysis of phosphoproteins in complex samples using the MOAC procedure. (A) Scheme of the method presented here with protein separation on SDS-PAGE. (B) Scheme for the direct analysis of enriched phosphoproteins.

phosphorylated proteins and peptides will be combined to identify phosphorylation sites and at the same time improve reliability of identification through higher sequence coverage.

We thank Megan McKenzie for revising the manuscript. We thank the Max Planck Society for financial support and Wolf D. Lehmann, Mark Stitt and Mirko Glinski for valuable discussion.

5 References

[1] Alberts, B., Bray, D., Lewis, K., Raff, M. *et al.*, *Molecular Biology of the Cell*, Garland Publishing, New York 1994, 195–222.

[2] Mann, M., Ong, S. E., Gronborg, M., Steen, H., Jensen, O. N., Pandey, A., *Trends Biotechnol.* 2002, *20*, 261–268.

[3] Pandey, A., Podtelejnikov, A. V., Blagoev, B., Bustelo, X. R., Mann, M., Lodish, H. F., *Proc. Natl. Acad. Sci. USA* 2000, *97*, 179–184.

[4] Steen, H., Kuster, B., Fernandez, M., Pandey, A., Mann, M., *J. Biol. Chem.* 2002, *277*, 1031–1039.

[5] Gronborg, M., Kristiansen, T. Z., Stensballe, A., Andersen, J. S., Ohara, O., Mann, M., Jensen, O. N., Pandey, A., *Mol. Cell. Proteomics* 2002, *1*, 517–527.

[6] Ueda, E. K., Gout, P. W., Morganti, L., *J. Chromatogr. A* 2003, *988*, 1–23.

[7] Porath, J., Carlsson, J., Olsson, I., Belfrage, G., *Nature* 1975, *258*, 598–599.

[8] Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., White, F. M., *Nat. Biotechnol.* 2002, *20*, 301–305.

[9] He, T., Alving, K., Feild, B., Norton, J., Joseloff, E. G., Patterson, S. D., Domon, B., *J. Am. Soc. Mass Spectrom.* 2004, *15*, 363–373.

[10] Andersson, L., Porath, J., *Anal. Biochem.* 1986, *154*, 250–254.

[11] Andersson, L., *J. Chromatogr.* 1991, *539*, 327–334.

[12] Oda, Y., Nagasu, T., Chait, B. T., *Nat. Biotechnol.* 2001, *19*, 379–382.

[13] Goshe, M. B., Conrads, T. P., Panisko, E. A., Angell, N. H., Veenstra, T. D., Smith, R. D., *Anal. Chem.* 2001, *73*, 2578–2586.

[14] McLachlin, D. T., Chait, B. T., *Anal. Chem.* 2003, *75*, 6826–6836.

[15] Whitaker, J. R., Feeney, R. E., *Crit. Rev. Food Sci. Nutr.* 1983, *19*, 173–212.

[16] Wells, L., Vosseller, K., Cole, R. N., Cronshaw, J. M., Matunis, M. J., Hart, G. W., *Mol. Cell. Proteomics* 2002, *1*, 791–804.

[17] Ahn, Y. H., Park, E. J., Cho, K., Kim, J. Y., Ha, S. H., Ryu, S. H., Yoo, J. S., *Rapid Commun. Mass Spectrom.* 2004, *18*, 2495–2501.

[18] Thompson, A., Schafer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Hamon, C., *Anal. Chem.* 2003, *75*, 1895–1904.

[19] Glinski, M., Romeis, T., Witte, C., Wienkoop, S., Weckwerth, W., *Rapid Commun. Mass Spectrom.* 2003, *17*, 1579–1584.

[20] Weckwerth, W., Willmitzer, L., Fiehn, O., *Rapid Commun. Mass Spectrom.* 2000, *14*, 1677–1681.

[21] Salih, E., *Mass Spectrom. Rev.* 2004, Epub ahead of print.

[22] Sano, A., Nakamura, H., *Anal. Sci.* 2004, *20*, 565–566.

[23] Kimura, Y., Shibasaki, S., Morisato, K., Ishizuka, N., Minakuchi, H., Nakanishi, K., Matsuo, M. *et al.*, *Anal. Biochem.* 2004, *326*, 262–266.

[24] Pinkse, M. W. H., Uitto, P. M., Hilhorst, M. J., Ooms, B., Heck, A. J. R., *Anal. Chem.* 2004, *76*, 3935–3943.

[25] Miyazaki, S., Morisato, K., Ishizuka, N., Minakuchi, H., Shintani, Y., Furuno, M., Nakanishi, K., *J. Chromatogr. A* 2004, *1043*, 19–25.

[26] Bradford, M. M., *Anal. Biochem.* 1976, *72*, 248–254.

[27] Wessel, D., Flugge, U. I., *Anal. Biochem.* 1984, *138*, 141–143.

[28] Laemmli, U. K., *Nature* 1970, *227*, 680–685.

[29] Wienkoop, S., Glinski, M., Tanaka, N., Tolstikov, V., Fiehn, O., Weckwerth, W., *Rapid Commun. Mass Spectrom.* 2004, *18*, 643–650.

[30] Tabb, D. L., McDonald, W. H., Yates, J. R., *J. Proteome Res.* 2002, *1*, 21–26.

[31] Weckwerth, W., Wenzel, K., Fiehn, O., *Proteomics* 2004, *4*, 78–83.

[32] Schulenberg, B., Goodman, T. N., Aggeler, R., Capaldi, R. A., Patton, W. F., *Electrophoresis* 2004, *25*, 2526–2532.

[33] Steinberg, T. H., Agnew, B. J., Gee, K. R., Leung, W. Y., Goodman, T., Schulenberg, B., Hendrickson, J. *et al.*, *Proteomics* 2003, *3*, 1128–1144.

[34] Yang, X., Wu, H., Kobayashi, T., Solaro, R. J., van Breemen, R. B., *Anal. Chem.* 2004, *76*, 1532–1536.

[35] Colettipreviero, M. A., Previero, A., *Anal. Biochem.* 1989, *180*, 1–10.

[36] Sano, A., Nakamura, H., *Anal. Sci.* 2004, *20*, 861–864.

[37] Bailey-Serres, J., Vangala, S., Szick, K., Lee, C. H., *Plant Physiol.* 1997, *114*, 1293–1305.

[38] Aggarwal, K. K., Saluja, D., Sachar, R. C., *Phytochemistry* 1993, *34*, 329–335.

[39] Guitton, C., Mache, R., *Eur. J. Biochem.* 1987, *166*, 249–254.

[40] Winter, H., Huber, S. C., *Crit. Rev. Biochem. Mol. Biol.* 2000, *35*, 253–289.

[41] Komatsu, S., Masuda, T., Hirano, H., *FEBS Lett.* 1996, *384*, 167–171.

[42] Riedel, J., Tischner, R., Mack, G., *Planta* 2001, *213*, 396–401.

[43] Schnolzer, M., Lehmann, W. D., *Int. J. Mass Spectrom.* 1997, *169*, 263.

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Chapter III: Detection of protein phosphorylation & determination of phosphorylation sites

In this chapter techniques for the general detection of protein phosphorylation and for phosphorylation site pinpointing are described and discussed. All techniques have been applied to plant material. The general detection techniques comprise: labelling with radioactive ^{32}P, labelling with a fluorescent probe, detection using anti-phosphoaminoacid-antibodies, phosphatase treatment in combination with mass shift analysis, and finally mass spectrometry based methods alone. Likewise, there is a variety of methods that can be used for protein phosphorylation site pinpointing. Site-specific protein phosphorylation analysis was classically performed using ^{32}P labelling or chemical derivatisation in combination with Edman sequencing. The onset of biological mass spectrometry led to the interpretation of specific fragmentation patterns characteristic for phosphorylated peptides using different instrumental setups.

III.1. Antibodies

General phospho-specific and phosphorylation site-specific antibodies can, in concert with an appropriate secondary antibody or directly coupled to an enzyme/fluorescent stain, be used to detect phosphorylation. This method has been successfully applied to detect tyrosine, serine, and threonine phosphorylation in plants [149-152]. However, the drawbacks of antibody based approaches mentioned above (chapter II) have to be considered.

III.2. Strategies relying on chemical derivatisation

Peptides phosphorylated on serine and threonine can undergo beta elimination during fragmentation in a mass spectrometer. This neutral loss of phosphoric acid often leads to incomplete fragmentation of the peptide in mass spectrometry based analyses and complicates the interpretation of spectra. To circumvent this problem the phosphate group can be replaced with a more stable residue resulting in more informative fragmentation spectra (figure 7).

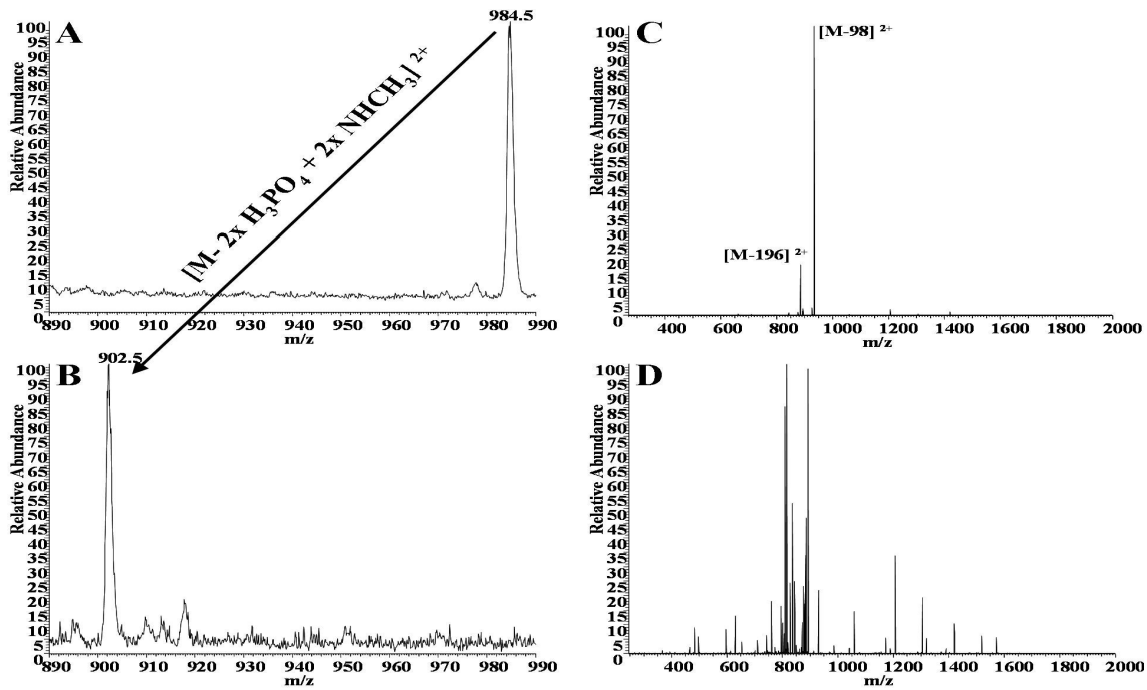


Fig. 7: ***Derivatisation of a doubly phosphorylated peptide by beta elimination and consecutive addition of methylamine. A, B: MS trace of the original peptide and the peptide derivatised with methylamine, respectively. C, D: MS^2 spectra of A and B. The derivatised peptide delivers a spectrum with more sequence information compared to the original peptide that displays almost exclusively the loss of one and two molecules of phosphoric acid ($[M-98]^{2+}$ and $[M-196]^{2+}$).***

To this end beta elimination of H_3PO_4 followed by Michael addition can be used for peptides phosphorylated on serine and threonine, always considering the disadvantages named before (chapter II). This technique has indeed been used to map phosphorylation sites in plant proteins [59, 74, 153]. However, because of the possible side reactions it is most useful to study known protein phosphorylation sites with high phosphorylation stoichiometry in greater detail.

III.3. Radioactive labelling

For a long time labelling with radioactive ^{32}P has been the method of choice due to its sensitivity and selectivity. Usually, cell cultures or parts of plants are fed with ^{32}P-orthophosphate *in vivo*. Alternatively, radioactive γ-ATP is applied in *in vitro* labelling experiments using purified enzymes and/or substrates (i.e. [85]). These procedures are followed by the extraction of proteins and detection of phosphorylation using scintillation counting and/or phosphor-imaging.

Not only is ^{32}P labelling used to detect phosphorylation but also to identify the peptide and ultimately the amino acid, which is phosphorylated. This can be achieved by Edman sequencing and mass spectrometry as discussed below. Even though ^{32}P labelling is still the most sensitive approach for the detection of phosphorylation it has two major drawbacks. Firstly, there are handling and waste disposal complications since ^{32}P is a radioactive compound and stringent safety rules apply for the use of it. The second and even more concerning issue is that the incorporation of radioactive phosphorus may considerably alter the *in vivo* state of the cells compared to non-radioactive phosphorus as has been shown in studies involving mammalian cells [154, 155].

III.4. Phosphatase treatment

The treatment of phosphorylated proteins/peptides with a phosphatase results in dephosphorylation accompanied by a mass shift of 80 Da. This approach has been described for different plant phosphoproteins and is a good means to detect and confirm phosphorylation if sample amount is not a limiting factor (one analysis before and one after phosphatase treatment is performed). For example, phosphorylation sites identified in proteins of the moss *Physcomitrella patens* were confirmed using phosphatase treatment and mass spectrometry [156]. Detection of phosphorylation of the bifunctional enzyme 6-Phosphofructo-2-kinase/fructose-2,6-bisphosphatase 2-phosphatase from *A. thaliana* was supported by alkaline treatment of the whole complex followed by the occurrence of a mass shift after SDS-PAGE [157]. However, phosphatase treatment in combination with mass shift measurements is only a general proof of phosphorylation.

III.5. Edman Sequencing

This technique developed by Pehr Edman led to a major breakthrough in biotechnology. For the first time, the amino acid sequence of a protein could be elucidated. Edman sequencing relies on the sequential degradation of the amino acids at the N-terminus of a polypeptide chain, which is coupled to a solid phase. The reagent phenylisothiocyanate (PTC) is added to the coupled peptide and the modified N-terminal amino acid can be selectively detached under strongly acidic conditions. The resulting amino acid derivatives are analysed using HPLC. By comparing the elution behaviour of the sample and standard amino acid derivatives stepwise peptide sequencing can be performed. If a phosphopeptide sample

was marked with ^{32}P phosphorylation site, pinpointing can be achieved by concominant measurement of ^{32}P activity (e.g. [158]). The major drawbacks are that the proteins to be sequenced have to be purified almost to homogeneity in advance and that it is time consuming.

III.6. Prediction programs / phosphorylation site databases

There are several phosphorylation site prediction programs available, which rely on different algorithms to elucidate the probability of phosphorylation on specific amino acids for any given protein. However, because of the enormous complexity of the cellular proteome, prediction programs still tend to lead to false positive results and their output has to be handled with care [159].

Also, several databases have been constructed, which contain data on experimentally verified phosphorylation sites (see appendix of the thesis for a collection of database links). There is even a major one for data repository on phosphorylation sites in plants called PlantsP, which however, does not contain much information on *in vivo* sites at the moment. As these databases are growing they are becoming more and more useful for the researcher and they will probably play an important role in the future for serving as encyclopaedias of phosphorylation sites and for the development of more sophisticated prediction programs.

III.7. Mutation analysis

Putative or experimentally identified protein phosphorylation sites can be mutated to confirm phosphorylation and/or to study its biological impact. To this end, mutants are generated with a single amino acid substitution at the phosphorylation site and compared to the wildtype form [160]. This technique is very useful for functional studies of experimentally established phosphorylation sites. However, due to its time consuming and laborious nature it is not generally employed in the search for unknown phosphorylation sites.

III.8. Dye technology

Dyes have been recently developed, which selectively stain phosphorylated proteins and peptides no matter if they are phosphorylated on serine, threonine, or tyrosine [161-163].

These dyes rely on the recognition principle of inorganic phosphate receptors developed by chemists [164-166]. The structure of the commercially available Pro-Q stain (Molecular Probes) is still a company secret but it is likely to be similar to the published structures. This dye seems to be very selective when a detailed protocol is stringently followed. In figure 8 the enrichment of phosphorylated proteins from *A. thaliana* cell cultures by aluminum hydroxide is shown. Phosphoproteins were enriched, subjected to SDS-PAGE and selectively stained by Pro-Q. Total protein content was visualised by coomassie staining.

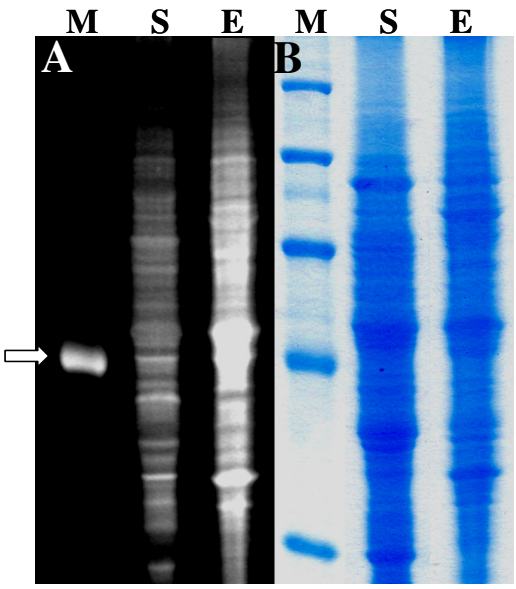


Fig. 8: Enrichment of phosphorylated proteins from A. thaliana cell cultures using aluminum hydroxide was followed by SDS-PAGE and selective staining of phosphorylated proteins. A: Staining of phosphorylated proteins with Pro-Q. B: Total protein staining with Coomassie. M: marker; E: eluate (enriched fraction). The arrow indicates the only phosphorylated protein in this marker (ovalbumin).

The only phosphorylated protein in the marker gives a positive signal using Pro-Q stain, while other marker proteins fail to be detected. In addition, the phosphoprotein fraction delivers clearly a stronger signal in the Pro-Q staining than the sample before enrichment, even though the protein amount is higher in the original sample. However, it should be noted that O-sulfonation can also contribute to the staining, albeit apparently to a lesser extent [163]. A big advantage of these stains is their compatibility with enzymatic in-gel digestion and mass spectrometry. A protein of interest can be subjected to gel electrophoresis, preliminarily examined for its phosphorylation, digested, and finally analysed by mass spectrometry. The disadvantages are the failure to differentiate between the three phosphorylated residues and the lower sensitivity when compared to ^{32}P labelling and antibody based approaches. However, fluorescent staining was found to greatly facilitate phosphoprotein analysis.

III.9. Fragmentation techniques in biological mass spectrometry

Depending on the instrument type used there is a variety of different fragmentation signatures typical for phosphopeptides, which can be used for the detection of phosphorylation and/or for the determination of phosphorylation sites. Every fragmentation technique leads to different fragmentation patterns, some of them leaving the phosphorylation intact, while others evoke the detachment of the phosphate group during the fragmentation process. Since mass spectrometry based methods rely on the determination of mass to charge ratios (m/z) the similarity of the mass shifts caused by phosphate (+ HPO_3) and sulfate (+SO_3) ester formation (the difference is only 0.009 mass units) is of concern. Only high precision and very expensive instruments like an FTICR (Fourier Transform Ion Cyclotron Resonance) mass spectrometer are able to distinguish between these two compounds following the m/z values.

However, it was reported that CID (Collision-Induced Dissociation) fragmentation of peptides phosphorylated at serine or threonine residues leads to a loss of about 98 Da (corresponding to H_3PO_4), while the sulphated counterparts lose 80 Da (corresponding to HSO_3^-) [113]. This special behaviour of phosphopeptides can be used to determine if a peptide is phosphorylated on a serine/threonine or not by searching for a loss of 98 Da (H_3PO_4) in the positive or for 79 Da (PO_3^-) in the negative mode [167-169]. Phosphorylation on tyrosine residues is usually more stable, but a loss of 80 Da (HPO_3) in the positive mode or presence of the immonium ion of phosphotyrosine (216.043 m/z) is a good indicator of tyrosine phosphorylation. However, one should be aware of the fact that peptides sulphated at tyrosine might also loose 80 Da (SO_3^-) in the positive mode [113, 170] and that the nominal mass of 216 Da can also be generated by other ions [171].

CID and PSD (Post-Source Decay) are the most common fragmentation techniques. The remaining techniques, however, promise to be important in future research projects involving plant phosphoproteomics.

III.9.1. PSD (Post-Source Decay)

PSD is a fragmentation technique typical for MALDI-ToF-MS instruments (Matrix Assisted Laser Desorption Ionisation-Time of Flight- Mass Spectrometry). It relies on metastable fragmentation processes proceeding in the flight tube when the instrument is used

in the reflector mode. However, this kind of fragmentation often results in spectra, which are hard to interpret and loss of phosphoric acid and related compounds from phosphopeptides is also commonly observed complicating the identification of phosphorylation sites.

III.9.2. ECD (Electron Capture Dissociation) and ETD (Electron Transfer Dissociation)

ECD and ETD are similar fragmentation techniques, which have the big advantage, that posttranslational modifications such as phosphorylation stay intact during the fragmentation process because of a milder fragmentation procedure [172-175]. ECD was introduced by McLafferty and coworkers and relies on the exposure of multiply protonated peptides to electrons in an FTICR mass spectrometer [176]. The capture of electrons is associated with the fragmentation of the peptide mostly into c and z type ions. Unfortunately, ECD is only possible on the very expensive FTICR instruments and not on more affordable mass spectrometers like ion traps since the molecules have to be held in a special energy stage, which is not easily achieved in ion traps. To overcome this limitation, Syka et al. developed a special interface for ion trap mass spectrometers that facilitates a process which is comparable to ECD [173]. This process is called ETD and also involves the transfer of electrons to protonated peptides. In this technique radical anionic species of polyaromatic hydrocarbons like fluoranthene are introduced into the trap and transfer electrons to multiply protonated peptides. This transfer leads to an internal fragmentation process believed to be equivalent to the one when using ECD. A slightly modified process can be used to sequence larger peptides of more than 20 amino acids and even proteins [172].

Unfortunately, these techniques are still not widely distributed since suitable instrumentation is very expensive (ECD) or only rarely available (ETD). However, they hold great promise for further research on protein phosphorylation.

III.9.3. CID (Collision-Induced Dissociation) / MS^3 fragmentation

CID is the most commonly applied fragmentation technique and applies to instrument types such as ion traps and triple quadrupole mass spectrometers. Peptide precursors are protonated on the peptide backbone, recorded (MS), and fragmented using a collision gas like helium or argon (MS/MS or MS^2). It is in this fragmentation step that the characteristic loss of 98 for phosphoric acid can be observed.

These fragments can then be fragmented further to yield even more detailed and reliable information of the peptide sequence (MS/MS/MS or MS^3; figure 9).

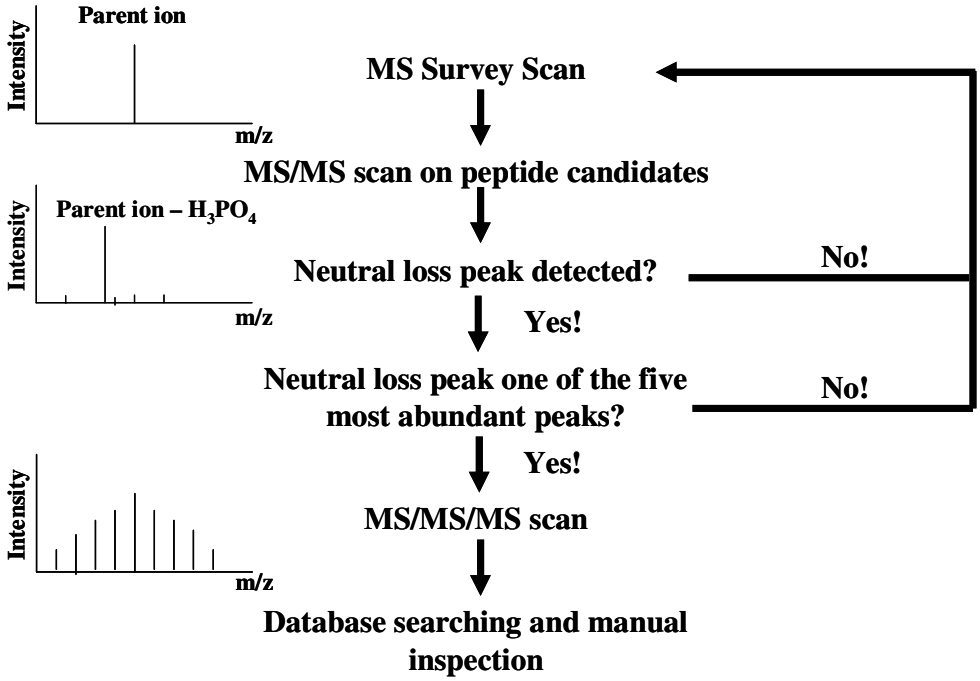


Fig. 9: ***Outline of a neutral loss-based CID dependent method used to analyse phosphopeptides.***

This second fragmentation step is especially useful, when phosphopeptides are to be analysed since CID often leads to a dominant loss of phosphoric acid from the peptide precursor in the first fragmentation step, which hinders fragmentation of the peptide backbone compared to unmodified peptides and complicates MS/MS analysis. The MS^3 step helps to overcome this problem by fragmenting the resulting dominant peak and the site of phosphorylation can, in the case of phosphoserine and phosphothreonine, be assigned indirectly by the identification of the respective dehydro-residues found in the MS^3 spectrum.

III.9.4. Combining phosphoprotein enrichment and the detection of phosphorylation sites

The non-targeted approach

By a combination of phosphoprotein enrichment using the developed method based on aluminum hydroxide and of the neutral loss driven MS^3 approach described above it was possible to identify phophopeptides and determine phosphorylation sites in over 30 proteins from *C. reinhardtii* and *A. thaliana* leaves, cell culture, and seeds (see table II).

Table II: Identified phosphopeptides

Accession nr.	Protein Description	Phosphopeptide sequence
	A. thaliana	
At5g1110	Sucrose-phosphate synthase	R.ISpSVDVFENWFAQHK.E (in vitro!)
At5g52300	low-temperature-responsive 65 kD protein (LTI65) / desiccation-responsive protein 29B (RD29B)	R.MKVTDEpSPDQKSR.Q K.ESDINKNpSPARFGGESK.A K.LPLSGGGpSGVKETQQGEEK.G K.LGYTGENGGGQSEpSPVKDETPR.S K.GAVTSWLGGKPKpSPR.S R.EAHQEPLNpTPVSLLSATEDVTR.T
At4g11740	Ara4 interacting protein, putative	R.SGpSPEEEHASINPAER.G
At3g12960	expressed protein similar to seed maturation protein from [Glycine max]	R.DIKDIKGTRTDDpSPR.-
At1g01100	60S acidic ribosomal protein P1 (RPP1A)	K.KKDEPAEEpSDGDLGFGLFD.-
At2g27710	60S acidic ribosomal protein P2 (RPP2B)	K.KEEKEEpSDDDMGFSLFE.-
At4g25890	60S acidic ribosomal protein P3 (RPP3A)	K.KKEEpSEEEEGDFGFDLFG.
At3g11250	60S acidic ribosomal protein P0 (RPP0C)	K.KEEpSDEEDYEGGFGLFDEE.-
At3g09200	60S acidic ribosomal protein P0 (RPP0B)	K.VEEKEEpSDEEDYGGDFGLFDEE.-
At2g40170	Em-like protein	R.KEQLGTEGpYQQMGR.K
At5g40420	glycine-rich protein / oleosin	R.HFQFQpSPYEGGR.G
At1g07985	Expressed protein	R.KLVDKVVGSSSPTNIHpSK.S
At5g50600	short-chain dehydrogenase/reductase (SDR) family protein similar to sterol-binding dehydrogenase steroleosin from [Sesamum indicum]	R.STLYPESIRTPEIKpSD.- R.ELGpSPNVVTVHADVSKPDDCRR.I
At1g29350	expressed protein	R.SGpSTHFSSTDSGNFQGK.S R.ELDFQYpSPFSAQQSMQSR.T
At4g25580	stress-responsive protein-related contains weak similarity to Low-temperature-induced 65 kDa protein	R.RGAPTLTPHNTPVSLLpSATEDVTR.T R.RGAPTLpTPHNpTPVSLLSATEDVTR.T R.GAPTLpTPHNTPVSLLSATEDVTR.T K.VTDPIGKpSTGEIGAASTVSAIGR.L
At4g31700	40s ribosomal protein S6 (RPS6A)	R.SRLpSSAAAKPSVTA.-
At5g65410	Zinc finger homeobox	R.ERpSEDPoMETSSAESGGGIR.K
At3g28920	Zinc finger	R.SoMDoMTPKpSPEPESETPTR.I
At2g37340	Splicing factor RSZ33	R.DRpSPVLDDEGSPK.I
At2g28490	Cupin family protein	K.GpSGpSpSECEDSYNIYDKK.D
At2g21490	Dehydrin family protein	K.FGpSGKHKDEQTPATATTTGPATTDQPHEK.K
At5g57655	Xylose isomerase family protein	K.AMEVpTHYLGGENYVFWGGR.E

At3g44110	*DNAJ heat shock*	*R.GEDVVHPLKVSLEDVYLGpTMK.K*
At1g13020	*EIF4B5*	*R.KADTEVSEpTPTAVK.T*
At1g53310	*phosphoenolpyruvate carboxylase*	*K.oMApSIDVHLR.Q*
At5g61780	*tudor domain-containing protein / nuclease family protein*	*R.TGIWEYGDIQpSDDEDNVPVRKPGRG.-*
C. reinhardtii		
163724	*10 kDa photosystem II polypeptide, similar to At1g79040, chloroplast-targeted*	*K.VGLNpSIEDPVVK.Q*
155812	*Uncharacterized conserved protein, contains BSD domain*	*K.ATADpSDDEGESAKPASTSLGAAAAASGK.E*
153417	*No annotation*	*R.KLESAApTVAER.Q* *K.VAVAPPSRPGpSGK.V* *R.SGpSAKVAVAPSR.Q*
166149	*RubisCO activase, chloroplast precursor*	*R.SIDAGVDApSDDQQDITR.G*
159642	*major component of the salt-soluble outer vegetative cell wall*	*R.ATASGGAALApSVENGSSAPSAGK.D*

All sequences are tryptic. pT, pS, pY: phosphothreonine,-serine,and –tyrosine; oM: oxidized methionine. .-: end of sequence. Cleavage sites are marked with a dot. Green: highly reliable (identification and site determination); Blue: most probably true (exact site unsecure); red: supported by some observations (found in enriched fraction, high Xcorrs, phosphorylation suggested by the annotation program but no MS^3 spectrum available due to the lack of a dominant neutral loss, no ions for phosphorylated residue(s)). At least one additional peptide was found for every protein, most hits are supported by more than three additional peptides.

To avoid annotation of false positives, stringent filtering criteria (Xcorrs: 2.0, 2.5, 3.5 for singly, doubly, and triply charged peptides, respectively. At least two peptides identified for every protein) were applied and all spectra were inspected manually. A high stringency is of crucial importance for mass spectrometry based phosphopeptide analysis and is explained in detail in the pitfalls section. Many of the proteins for which phosphorylation sites were found in *A. thaliana* seeds could be confirmed to be seed specific and are therefore interesting targets for studies on seed germination (see publication at the end of this chapter). On the other hand, a whole set of phosphorylated ribosomal proteins, which are not expressed in a tissue specific manner could be identified. These proteins belong to the group of 60 s ribosomal acidic proteins and are phosphorylated at their C-termini. Many of these proteins have been reported to be phosphorylated at homologues sites in other organisms so their phosphorylation seems to be a general mechanism and probably plays a role in protein biosynthesis. This non-targeted approach delivered important insights into plant protein phosphorylation and was among the very first in the field of plant phosphoproteomics.

To further investigate the possibility of using this methodological setup for a more hypothesis-driven approach I chose to examine the putative phosphorylation of phosphoenolpyruvate carboxylase and malate dehydrogenase (see next paragraph).

It was also investigated if a double enrichment procedure (phosphoprotein enrichment followed by phosphopeptide enrichment) was more effective than phosphoprotein enrichment alone.

Therefore, phosphoproteins of *C. reinhardtii* or *A. thaliana* cell cultures were enriched using aluminum hydroxide and applied to SDS-PAGE. Either, the proteins (about 25 µg) were separated for 1.5 h at 120 volts, 10 lanes were cut out of the gel, digested separately and peptides analysed as explained in the appendix of this thesis, or proteins (about 50 µg) were only desalted by SDS-PAGE (run for 10 min at 120 volts), digested as mixture, and phosphopeptides were enriched by titania performed by Dr. Ralf Krüger (DKFZ, Heidelberg). Peptides were analysed as described in this chapter and in the appendix at the end of this thesis. The results show that neither the single nor the double enrichment procedure resulted in the identification of a large number of phosphopeptides (table III). This can be explained partly by the fact that the amount of sample in these experiments was not very high. However, it could be observed that the vast majority of peptides (over 90%) identified after the double enrichment procedure were not phosphopeptides indicating that specific enrichment of

phosphopeptides out of such complex mixtures as the ones under investigation is difficult to achieve.

Table III: Phosphopeptides identified after single phosphoprotein enrichment and after combined phosphoprotein & phosphopeptide enrichment

A. thaliana cell culture				
Accession nr.	***Protein***	***Sequences***	***Protein enrichment***	***Protein & peptide enrichment***
At5g61780.1	***tudor domain-containing protein / nuclease family protein***	***R.TGIWEYGDIQpSDDEDNVPVRKPGRG.-***	**+**	**-**
At1g53310.1	***phosphoenolpyruvate carboxylase***	***K.oMApSIDVHLR.Q***	**+**	**+**
At3g09200.1	***60S acidic ribosomal protein P0 (RPP0B)***	***K.VEEKEEpSDEEDYGGDFGLFDEE.-***	**+**	**+**
At4g25890.1	***60S acidic ribosomal protein P3 (RPP3A)***	***KKEEpSEEEEGDFGFDLFG.***	**+**	**+**
At3g11250.1	***60S acidic ribosomal protein P0 (RPP0C)***	***K.KEEpSDEEDYEGGFGLFDEE.-***	**-**	**+**
C. reinhardtii				
JGI accession nr.	***Protein***	***Sequences***	***Protein enrichment***	***Protein & peptide enrichment***
163724	***10 kDa photosystem II polypeptide, similar to At1g79040, chloroplast-targeted***	***K.VGLNpSIEDPVVK.Q***	**-**	**+**
155812	***Uncharacterized conserved protein, contains BSD domain***	***K.ATADpSDDEGESAKPASTSLGAAAAASGK.E***	**-**	**+**
153417	***Putative protein***	***R.KLESAApTVAER.Q*** ***K.VAVAPPSRPGpSGK.V*** ***R.SGpSAKVAVAPSR.Q***	**+** **+** **+**	**+** **+** **-**
166149	***RubisCO activase, chloroplast precursor***	***R.SIDAGVDApSDDQQDITR.G***	**-**	**+**
159642	***major component of the salt-soluble outer vegetative cell wall***	***R.ATASGGAALApSVENGSSAPSAGK.D***	**-**	**+**
167609	***Proline rich region protein***	***K.GTGFSALFATPK.R*** ***K.TGSGFSFFGTGK.V***	**+** **+**	**-** **-**

-: end of sequence. All sequences are tryptic. Cutting sites are marked with a dot.

At the state used, the combined enrichment resulted in a comparable number of identified phosphopeptides in a shorter period of time since ten analyses (ten gel pieces, ten extractions, ten LC/MS runs) were necessary for the phosphoprotein samples while only one analysis (one gel piece, one extraction, one (however quite long) LC/MS run) was needed for

the doubly enriched sample. However, to obtain maximal information for the doubly enriched sample (maximal sequence information) one should analyse the flow through as well, which doubles the time for analysis and the time needed for the enrichment procedure has also be taken into account. The improvement of the phosphopeptide enrichment procedure used or the application of alternative methods might overcome the unspecifity problem.

However, the reader might be disappointed by the rather low number of reported phosphopeptides using the non-targeted approach. This is partly because much of the work was dedicated to time consuming method development and considerably less time could be dedicated to the generation of inventory lists but can not be explained by this issue alone. In the plant field there are two publications which can serve as excellent sources of comparison. Both were published very recently and claim the identification of hundreds of phosphoproteins from *A. thaliana* and *C. reinhardtii,* respectively. In fact, there are diverse reasons for this discrepancy but the most obvious one is that these studies operate with material that is already enriched in phosphorylation before IMAC based phosphopeptide enrichment. This is either by using tissue naturally rich in phosphorylation i.e. membranes [119] or by treating the tissue *in vivo* with phosphatase inhibitors long before protein extraction to artificially elevate the level of phosphorylation [177]. While the former is an attractive approach to obtain more data on protein phosphorylation the latter one opens up a few questions. Does the phosphorylation one measures after incubating for a long time with phosphatase inhibitors reflect the *in vivo* situation (i.e. does the addition of phosphatase inhibitors before harvest only increase naturally occurring phosphorylation) or might it be that artificial phosphorylation is introduced by this treatment leading to the identification of false positives? Additionally, the criteria accepted to be reliable for successful identification vary between studies making a direct comparison of the numbers of identified phosphopeptides difficult. As will be discussed in this chapter, the neutral loss driven MS^3 approach for the identification of phosphopeptides, which is used in many recent phosphopeptide studies has not only its merits can also lead to the identification of false positives if not applied with care. Also, the initial criteria used for filtering data obtained after database analysis often vary to a large extent between studies and some researchers define a phosphopeptide with a missed cleavage compared to the same peptide without missed cleavage as an additional hit to enlarge their list of identified phosphopeptides, others count those two peptides as one phosphopeptide. Since missed cleavages occur quite frequently this issue should be considered when comparing lists of peptides.

It is indeed difficult to determine the best criteria and it is a source of ongoing controversy between researchers that apply proteomics [95, 178, 179]. Solutions to this problem are a matter of ongoing research but until now manual validation and the application of stringent filtering criteria are often the only way to obtain high quality data.

Investigating phosphorylation of phosphoenolpyruvate carboxylase and malate dehydrogenase

Phosphoenolpyruvate carboxylase (PEPC) and malate dehydrogenase (MDH) are two key enzymes of central plant metabolism. Even though there are several isoenzymes with diverse function the general idea is that the function of these two enzymes is connected. The oxaloacetate produced by PEP carboxylase can be transported to the mitochondrion where it is converted to malate by mitochondrial malate dehydrogenase [180-182]. Oxaloacetate as well as malate participates in the anaplerotic reactions that replenish the tricarboxylic acid cycle. As stated before, phosphorylation plays an important role in the regulation of several key enzymes in plant metabolism. One of the most thoroughly investigated enzymes in this regard is PEPC. On the other hand there is no report known to the author on phosphorylation of plant malate dehydrogenase and thus these two enzymes were selected for a hypothesis-driven approach.

In *A. thaliana* four genes are known to code for different PEPC isoforms. While one isoform does not show the phosphorylation consensus motif (E/D)(K/R)XYS*IDAQLR (PEPC4; At1g68750) two of them display this motif exactly (PEPC2; At2g42600 and PEPC3; At3g14940) and one (PEPC1; At1g53310) bears only a slight variation of this motif in which alanine is replaced by valine and glutamine by histidine (figure 10).

```
PEPC1          MANRKLEKMASIDVHLRQLVPGKVSEDDKLVEYDALLLDRFLDILQDLHG-------EDL 53
PEPC3          MAGRNIEKMASIDAQLRQLVPAKVSEDDKLVEYDALLLDRFLDILQDLHG-------EDL 53
PEPC2          MAARNLEKMASIDAQLRLLAPGKVSEDDKLIEYDALLLDRFLDILQDLHG-------EDV 53
PEPC4          --------MTDTTDDIAEEISFQSFEDD-----CKLLGSLFHDVLQREVGNPFMEKVERI 47
```

Fig. 10: ***Sequence alignment of the different A. thaliana PEPC isoforms. Only the first part of the sequence is shown. The phosphorylation site is marked in bold italic and underlined.***

To this end phosphoproteins were enriched using aluminum hydroxide and separated by SDS gel electrophoresis. Protein bands corresponding to the size of a PEPC subunit (about 100 kD) or of malate dehydrogenase (about 36 kD) were cut out of the gel and proteins were trypsinized. Peptides were submitted to nano-HPLC and directly eluted into a linear ion trap mass spectrometer. Both enzymes were reliably identified after this procedure and the

evaluation of MS^3 spectra suggested the finding of phosphorylation sites for both of them. However, after careful validation, only the site of the PEPC could be maintained. These two examples are well suited to illustrate the possibilities but also the pitfalls one encounters when applying the suggested procedure.

Proofs for phosphorylation of PEPC1 came from several directions. Initial identification was made by a MS^3 spectrum of the respective peptide, verified by the corresponding MS^2 spectrum and ultimately by comparing the experimentally obtained spectra to spectra derived from a synthetic peptide. This confirmation supports the assumption that regulation of PEPC by phosphorylation on the respective residue is a common theme in the plant kingdom and is valuable information for all researchers working on this important enzyme. At the same time this result shows that the approach I used is suitable not only for non-targeted analyses but also for the verification of a hypothesis.

For malate dehydrogenase one MS^3 spectrum also strongly suggested phosphorylation. However, when comparing the actual neutral loss stemming from the MS^2 spectrum (32 Da) to the theoretical value (49 Da) a clear discrepancy became obvious. The answer to this and related cases can be found in the following paragraph and in the relevant publication attached to this chapter.

III.9.5. Pitfalls of CID-dependent neutral loss-based phosphopeptide analysis

Besides the many advantages the CID-based method has, there are also potential pitfalls that have to be considered. The first question one might ask is whether it is true that the dominant loss of 98 Da is indeed always restricted to phosphorylated peptides. To clarify this important issue the MS^2 and MS^3 spectra were examined in detail.

There are several possibilities how losses, which are not related to phosphorylation can mimic the loss of phosphoric acid. To clarify this issue one should consider the following: a peptide is modified with an unknown compound which corresponds to a value of about 32, 49, or 98 Da (which also correspond to the loss of phosphoric acid of a triply, doubly, or singly charged ion, respectively). This unknown modification may be for example oxidation of methionine. This modification often leads to a dominant $[M-64]^{n+}$ ion (loss of methane sulfenic acid, $HSOCH_3$) [183]. If the precursor is doubly charged, this means a loss of 32 Da, which is very close to the 32.7 Da observed for a loss of phosphoric acid from a triply charged precursor. Subsequently, an MS^3 fragmentation step is triggered because the program does not incorporate the information of the charge state. The resulting MS^3 spectrum can then, because

of sequence identity or because of the enormous number of possible combinations, match to a totally different peptide than the MS^2 spectrum and can lead to the false positive identification of a phosphopeptide. In fact, this kind of behaviour was observed quite often and increasingly when not applying rigorous filtering criteria. While the original database search with the MS^3 spectrum leads to the identification of a presumably phosphorylated *A. thaliana* peptide, a modified database search including MS^2 spectra and the additional modification of minus 48 (loss of methane sulfenic acid) can reveal that indeed a methionine oxidation and not phosphorylation was responsible for triggering a MS^3 spectrum. This was also true for the mysterious phosphorylation site of malate dehydrogenase (for details see relevant publication on methionine oxidation and phosphopeptide analysis in RCM).

Consequently, the information of the MS^2 spectrum has to be considered: does the presumed loss of phosphoric acid match the charge state of the MS^2 peptide (e.g. is a loss of 49 from the precursor observed for a doubly charged ion)? Additionally, direct information about the site of phosphorylation can in some cases be gathered from the MS^2 spectrum and thus contribute to reliable phosphorylation site identification.

A highly suspicious candidate for false identifications is also sulfonation because of its similarity to phosphorylation. As stated before Medzhiradsky *et al.* report the specificity of a loss of 98 Da for a peptide phosphorylated on serine compared to the loss of 80 Da for the same peptide sequence but bearing a sulfated serine instead of the phosphorylated one [113]. However, this study only gives one example and there is no study that systematically investigates the impact of the amino acid sequence on the fragmentation of a peptide sulfated on serine and threonine.

Therefore, it was investigated whether this was true for a broader range of modified peptides. For this purpose a set of 12 peptides was selected, suitable to determine if amino acids directly neighbouring the modified amino acid can influence the fragmentation behaviour (see table IV). This set was generated by altering the sequence of a peptide from pyruvate carboxylase, which had been determined to be phosphorylated (see above).

Indeed, all of the tested peptides phosphorylated on serine or threonine displayed a dominant loss of 98 Da and all of the cognate sulfated peptides a loss of 80 Da thus confirming the assumption that the fragmentation behaviour was indeed characteristic and specific and did not depend on the amino acid residues directly neighbouring the modified amino acid.

Table IV: Peptide sequences used to investigate the fragmentation behaviour of sulfated and phosphorylated peptides:

Peptide sequence
MAS^#IDVHQLR*
MIS^#IDVHQLR*
MKS^#IDVHQLR*
MPS^#IDVHQLR*
MFS^#IDVHQLR*
MAS^#PDVHQLR*
MAS^#KDVHQLR*
MAS^#ADVHQLR*
MAS^#FDVHQLR*
MAT^#IDVHQLR*
MAY^#IDVHQLR*

*^#: sequences were investigated unmodified, sulfated, or phosphorylated on the preceding residue. Sulfated peptides were generated from unmodified peptides by the protocol described in [113].

However, some sulfated peptides in addition showed a loss of 98 Da represented by one of the three most abundant peaks (figure 11).

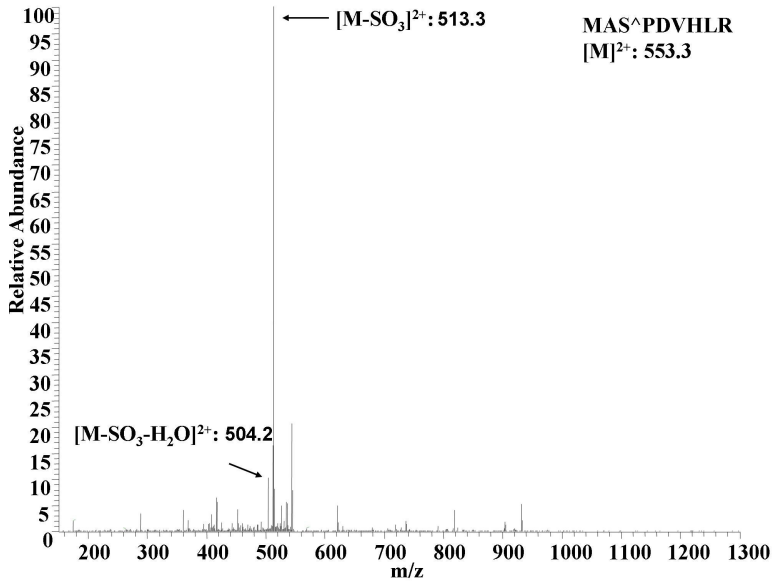


Fig. 11: MS² spectrum of the sulfated peptide MAS^PDVHLR. ^denotes sulfation.

The observed loss of 98 Da probably stems from the additional loss of water from the peptide and not from the loss of sulphuric acid. Since it can be one of the three to five most abundant peaks in the MS^2 fragmentation step it can trigger MS^3 fragmentation automatically and leads to a false positive hit. However, manual inspection of the MS^2 spectra can easily reveal this pitfall and is therefore strongly recommended.

Appendix chapter III

Wolschin, F. & Weckwerth, W. Combining metal oxide affinity chromatography (MOAC) and selective mass spectrometry for robust identification of in vivo protein phosphorylation sites. *Plant Methods* 1, 9 (2005).

Wolschin, F. & Weckwerth, W. Methionine oxidation in peptides - a source for false positive phosphopeptide identification in neutral loss driven MS^3. *Rapid Commun Mass Spectrom* in press (2006).

Plant Methods

BioMed Central

Methodology

Open Access

Combining metal oxide affinity chromatography (MOAC) and selective mass spectrometry for robust identification of *in vivo* protein phosphorylation sites

Florian Wolschin and Wolfram Weckwerth*

Address: Max Planck Institute of Molecular Plant Physiology, 14424 Potsdam, Germany

Email: Florian Wolschin - wolschin@mpimp-golm.mpg.de; Wolfram Weckwerth* - weckwerth@mpimp-golm.mpg.de

* Corresponding author

Published: 01 November 2005

Plant Methods 2005, **1**:9 doi:10.1186/1746-4811-1-9

Received: 27 August 2005

Accepted: 01 November 2005

This article is available from: http://www.plantmethods.com/content/1/1/9

© 2005 Wolschin and PWeckwerth; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Protein phosphorylation is accepted as a major regulatory pathway in plants. More than 1000 protein kinases are predicted in the *Arabidopsis* proteome, however, only a few studies look systematically for *in vivo* protein phosphorylation sites. Owing to the low stoichiometry and low abundance of phosphorylated proteins, phosphorylation site identification using mass spectrometry imposes difficulties. Moreover, the often observed poor quality of mass spectra derived from phosphopeptides results frequently in uncertain database hits. Thus, several lines of evidence have to be combined for a precise phosphorylation site identification strategy.

Results: Here, a strategy is presented that combines enrichment of phosphoproteins using a technique termed metaloxide affinity chromatography (MOAC) and selective ion trap mass spectrometry. The complete approach involves (i) enrichment of proteins with low phosphorylation stoichiometry out of complex mixtures using MOAC, (ii) gel separation and detection of phosphorylation using specific fluorescence staining (confirmation of enrichment), (iii) identification of phosphoprotein candidates out of the SDS-PAGE using liquid chromatography coupled to mass spectrometry, and (iv) identification of phosphorylation sites of these enriched proteins using automatic detection of H_3PO_4 neutral loss peaks and data-dependent MS^3-fragmentation of the corresponding MS^2-fragment. The utility of this approach is demonstrated by the identification of phosphorylation sites in *Arabidopsis thaliana* seed proteins. Regulatory importance of the identified sites is indicated by conservation of the detected sites in gene families such as ribosomal proteins and sterol dehydrogenases. To demonstrate further the wide applicability of MOAC, phosphoproteins were enriched from *Chlamydomonas reinhardtii* cell cultures.

Conclusion: A novel phosphoprotein enrichment procedure MOAC was applied to seed proteins of *A. thaliana* and to proteins extracted from *C. reinhardtii*. Thus, the method can easily be adapted to suit the sample of interest since it is inexpensive and the components needed are widely available. Reproducibility of the approach was tested by monitoring phosphorylation sites on specific proteins from seeds and *C. reinhardtii* in duplicate experiments. The whole process is proposed as a strategy adaptable to other plant tissues providing high confidence in the identification of phosphoproteins and their corresponding phosphorylation sites.

Background

The proteome of different developmental stages of any kind of organism reflects more directly than the genome or the transcriptome the metabolic specialisation for the actual developmental state. In plants several studies on the proteome of different developmental stages have been conducted [1]. Seed dormancy plays a crucial role in the life cycle of plants and its proteome reflects the metabolic processes during this important developmental period. However, investigation on posttranslational modifications of the proteins gives an even more detailed view on the complex nature of seed metabolism. Protein phosphorylation has been widely described as a major regulatory protein posttranslational modification influencing many important processes in living cells [2-4]. Thus, measuring protein phosphorylation is essential to reveal regulatory and signal pathways. However, the study of protein phosphorylation confronts the researcher with several hurdles.

A complicating fact is that many proteins are not only phosphorylated at one site, but on multiple sites and that each modification seems to have different regulatory functions [5,6]. Therefore, detection of protein phosphorylation and identification of phosphorylation sites are needed for the understanding of protein regulation.

Traditionally, phosphorylation is detected by specific antibodies and/or by incorporating radioactive [^{32}P]orthophosphate into proteins. However, while immunolabelling is often unspecific, incorporation experiments using radioactivity might result in artificial phosphorylation events and impose waste disposal problems. Only recently has it become possible to reliably detect protein phosphorylation by resolving the proteins of interest on a gel and submitting the gel to fluorescent phosphate specific dyes followed by a staining of total protein [7-11].

The low abundance of phosphoproteins and the accurate identification of the specific phosphorylation sites, however, still impose problems.

Because of their sensitivity and resolving power, selective enrichment of phosphorylated peptides/proteins and liquid chromatography coupled to mass spectrometry (LC-MS) based methods are now widely used for phosphorylation site identification (for review see [4,12,13]).

The low abundance of phosphorylated peptides can be circumvented in part by enrichment of the phosphopeptides with IMAC (Immobilised metal affinity chromatography) after tryptic digestion of phosphoproteins. While numerous publications exist describing the enrichment of phosphopeptides from animal sources considerably less studies have been published for plant tissue (for review see [14]). These broad range studies focus for example on thylakoid proteins and plasma membrane proteins of *A. thaliana* [15,16] or on the moss *Physcomitrella patens* [17].

However, relying on just one peptide for protein identification (as is commonly done after phosphopeptide enrichment) is prone to the identification of false positives. What is more, peptides phosphorylated on serines or threonines tend to loose their phosphate group during the fragmentation process in the mass spectrometer thus further complicating correct assignment. This drawback can be circumvented by the enrichment of complete phosphoproteins since this approach leads to the identification of several peptides per protein and therefore enhances the reliability of protein identification.

Recently, we reported a novel method for the enrichment of phosphorylated proteins out of complex mixtures termed MOAC (metal oxide affinity chromatography) [18]. This method adds another tool to the repertoire of methods for the identification of phosphorylated proteins and might help to overcome some of the specificity problems associated with IMAC. In this initial study we showed that MOAC can be used to enrich phosphorylated proteins from plant leaf tissue [18].

In the present study we use MOAC for the enrichment of phosphorylated proteins from *A. thaliana* seeds and for proteins from *C. reinhardtii* cell cultures and show that the method is suitable and reproducible for different kinds of samples.

However, for the identification of the exact sites of phosphorylation further steps are necessary. As mentioned above the ionization efficiency and the quality of a phosphopeptide spectra is sometimes not good enough for reliable identification and even more difficult is the determination of the exact phosphorylation site. To circumvent these problems phosphate groups can be replaced by beta-elimination and Michael addition with more stable residues thereby increasing the ionization efficiency and improving the fragmentation behaviour [5,19-23]. Yet, these approaches sometimes result in unwanted side reactions and are difficult to perform on complex mixtures. Another promising method is the analysis of peptide sequences by electron transfer dissociation [24] but this requires sophisticated modification of the mass spectrometer not yet widely available.

The approach we employed to identify phosphorylation sites makes use of the neutral loss of H_3PO_4 during MS^2 fragmentation. The dominant neutral loss peak observed in many MS^2 spectra derived from serine/threonine phosphorylated peptides [25] is routinely broken down in an additional MS^3 step and putative phosphopeptides are

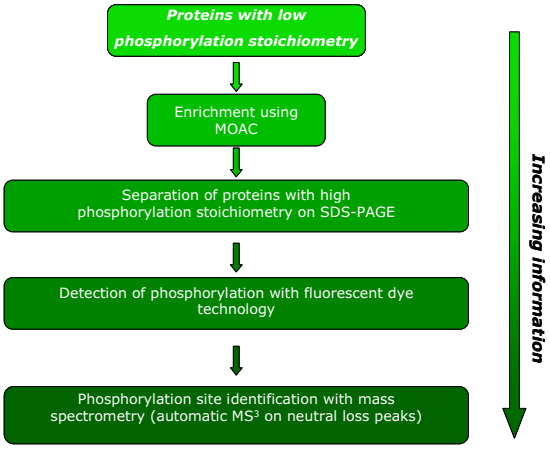


Figure 1
Proposed strategy for robust identification of serine/threonine phosphorylation in plants. In case of low stoichiometry as for the *in vivo* situation enrichment of phosphoproteins is necessary. To cope with this a novel enrichment procedure called metaloxide affinity chromatography (MOAC) is used in the strategy [18]. The whole approach is applicable to identify phosphorylation sites out of complex samples. The library of *in vivo* sites may be used for screening of protein phosphorylation dynamics using triple-quadruple mass spectrometry [6].

automatically detected in a single LC/MS run [26]. The resulting MS^3 spectra are often informative enough to identify the correct peptide and the phosphorylation site in a database search. However, considering the information of both MS^2 and MS^3 data is sometimes necessary to obtain unambiguous identification [27,28].

In the present work we combine the MOAC enrichment of phosphoproteins and selective mass spectrometry for a detailed study of protein phosphorylation. Phosphorylation site identification is demonstrated for enriched seed proteins and is achieved using data-dependent MS^2 as well as neutral-loss triggered MS^3 fragmentation on a linear ion trap mass spectrometer. The proposed strategy is suitable to identify putative phosphoprotein candidates with high sequence coverage and at the same time it allows identification of corresponding protein phosphorylation sites. To avoid false positive identification of phosphorylation sites only hits with congruent MS^2 and MS^3 spectra were considered in this study. Careful validation of the data led to the identification of 16 phosphorylation sites in nine seed proteins, some of them known to be phosphorylated also in the mammalian system such as ribosomal proteins. A comparison with microarraydata showed that these proteins are mainly seed specific.

Results

General strategy for the identification of serine/threonine phosphorylation in plants

An overview on the general strategy for the identification of serine /threonine phosphorylation in plants is shown in Figure 1. The strategy includes enrichment of phosphoproteins using MOAC, phosphorylation detection using fluorescent dye technology, and determination of the phosphorylation site with neutral loss driven MS^3. Most important, each step of the procedure gives further confidence for a robust identification of phosphoproteins and phosphorylation sites. For a detailed description see the following sections.

MOAC enrichment of phosphoproteins

Proteins were extracted with phenol and enriched for phosphoproteins using metaloxide affinity chromatography (MOAC) [18]. Proteins were separated using SDS-PAGE and visualised with a phosphate-specific stain followed by staining with coomassie R-250 (Figure 1). Similar amounts of total protein from samples taken before and after MOAC (Figure 2B) are accompanied by strong differences in the phosphostain signal indicating clear enrichment of phosphoproteins out of the complex sample (Figure 2A). Bands corresponding to enriched seed proteins with a strong signal in the phosphostain were excised, digested with trypsin, and analysed by mass spectrometry (see below).

Identification of in vivo phosphorylation sites in A. thaliana seed proteins

Following the proposed strategy (see Figure 1) we identified after careful validation 16 phosphorylation sites in 9 proteins using the combined information of MS^2 and MS^3 data. The advantage of using the combined data is exemplified by the MS^2 and MS^3 spectra in Figure 3. In the first MS^2 fragmentation step a very intense fragment ion stemming from a dominant neutral loss of phosphoric acid is seen. This spectrum alone is often not suitable for database search. On the other hand it is a distinct indication for a phosphopeptide. The dominant neutral loss fragment triggers in a second scan event an MS^3 fragmentation step. The combination of MS^2 and MS^3 leads to the clear phosphorylation site identification with increased confidence based on the observed neutral loss of phosphoric acid [27,28]. A further level of confidence is provided by the combination of protein identification (see Figure 1 and protein sequence coverage in table 1) and the corresponding protein phosphorylation site in one LC/MS analysis. This information is missing in strategies where only phosphopeptides are enriched and protein identification is based solely on the detection of one phosphopeptide. The phosphorylation sites are shown in Table 1. The sites of the two ribosomal proteins are known to be phosphorylated in mouse [29]. For some of the proteins

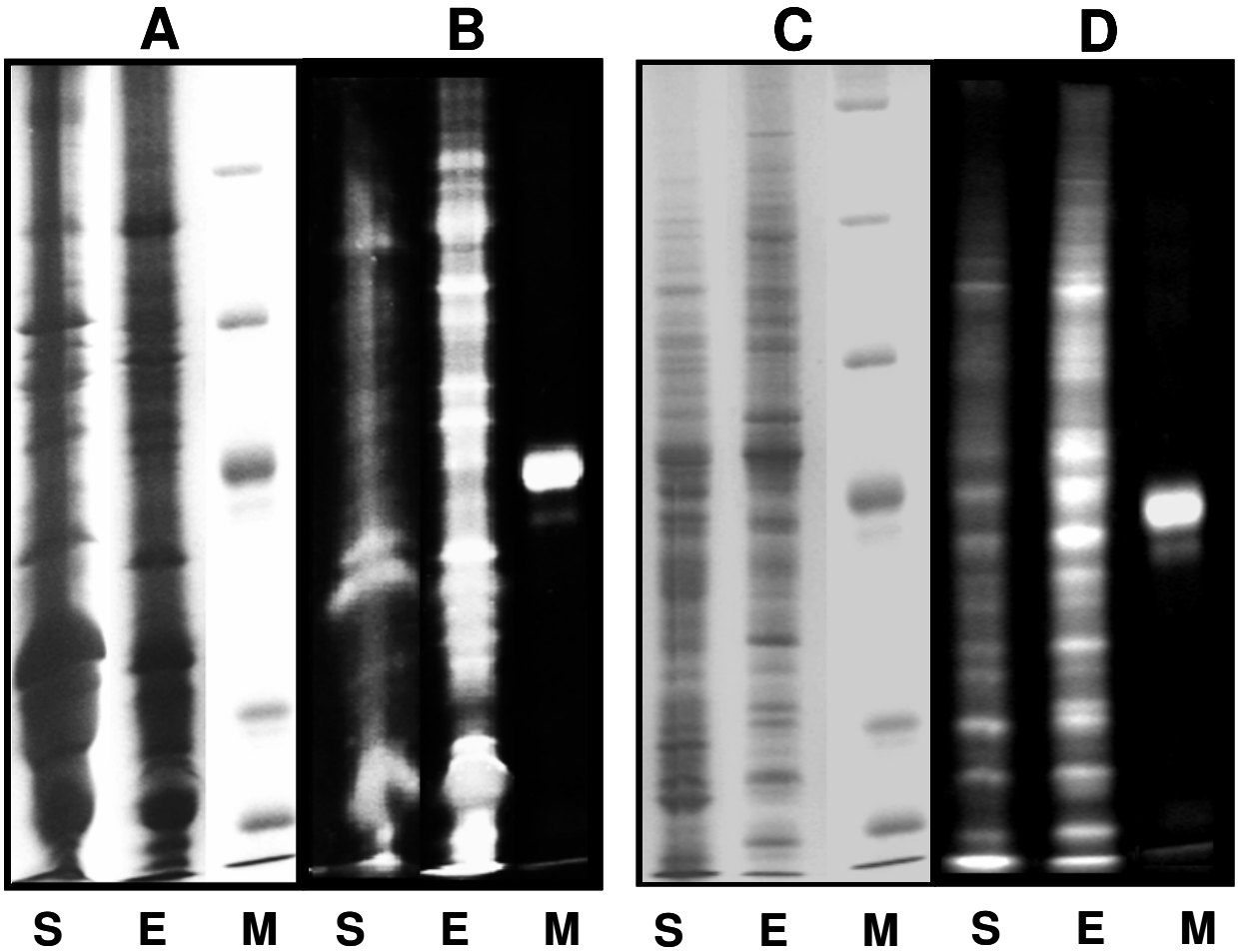


Figure 2
A. thaliana seed proteins and *C. reinhardtii* proteins before and after MOAC. M: marker S: sample (before MOAC) E: eluate (after MOAC); A: *A. thaliana* seed proteins, Coomassiestaining; B: *A. thaliana* seed proteins, Phosphostaining. C: *C. reinhardtii* proteins, Coomassiestaining; D: *C. reinhardtii* proteins, Phosphostaining. The labelled marker protein in the phosphostain is the phosphorylated protein Ovalbumin.

identified after MOAC, MS^2 data suggested phosphorylation but MS^3 data were lacking. All these ambiguous results were not included in the table. The analysis of gene expression data revealed that five of the identified proteins are apparently seed specific (table 1). Notably, phosphorylation is most probably not confunded with o-sulfonation in these experiments since sulfated peptides exhibit the characteristic loss of 80 Da instead of 98 Da during collision induced dissociation [30], which does not lead to the triggering of an MS^3 spectrum.

Reproducibility of the method

The enrichment process was repeated using A. thaliana leaf and C. reinhardtii proteins. Similar patterns were observed in SDS PAGE analysis combined with phospho- and coomassiestaining (data not shown). Two prominent bands, one of the enriched seed sample at about 65 kDa (corresponding to the mw of the protein with the highest number of phosphorylation sites identified in the first experiment), and one of the enriched C. reinhardtii sample (in duplicate) at about 56 kDa, were selected for testing the reproducibility of phosphorylation site identification. With the seed protein we tested analytical reproducibility after storage of the sample for 1 month (see table 2). Reproducibility of the whole procedure including MOAC enrichment is demonstrated by the repeated identification of phosphorylation sites in the 56 kDa C. reinhardtii phosphoprotein (table 2).

Phosphorylation of sterol dehydrogenase

One phosphorylation site was identified as serine 95 in an isoform of the A. thaliana short chain sterol dehydroge-

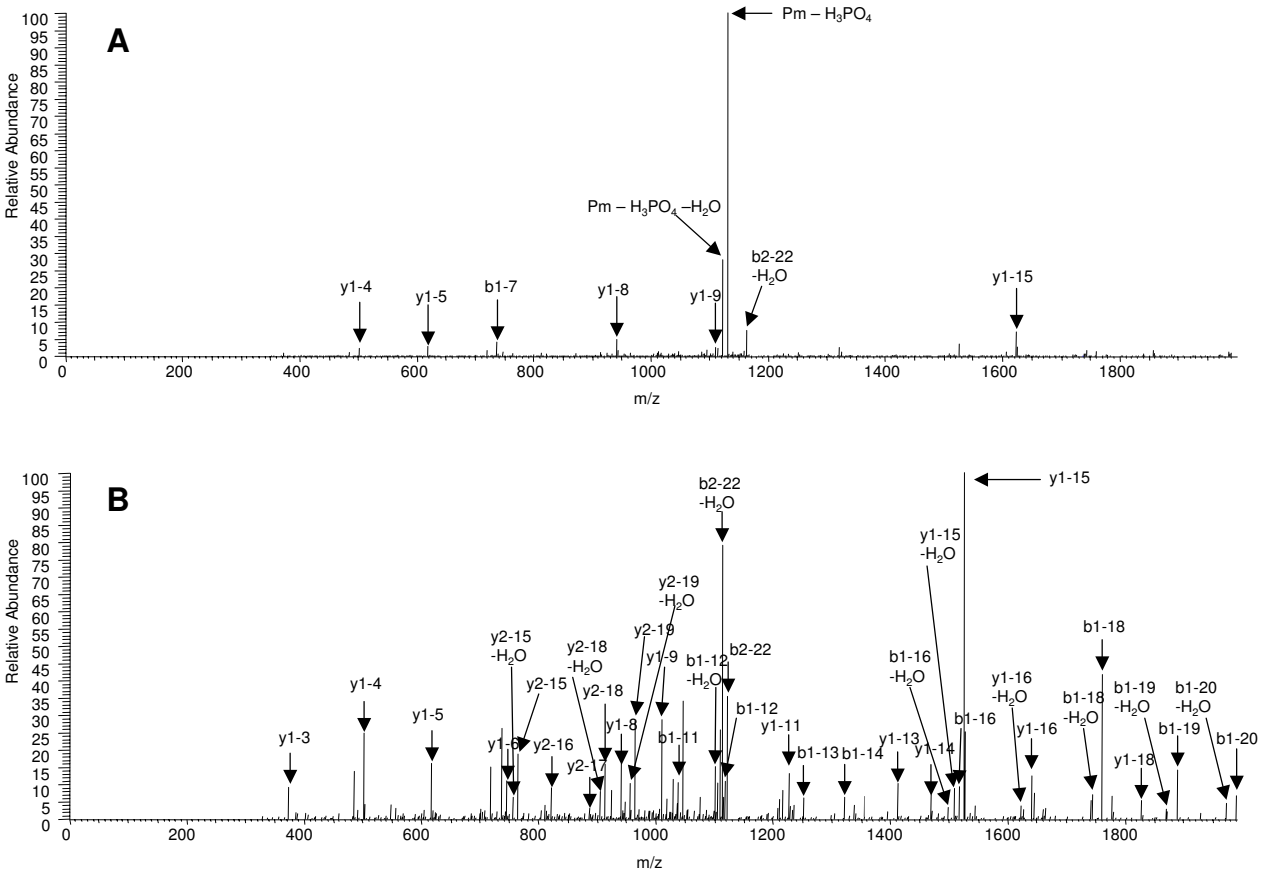


Figure 3

MS^2 spectrum and the cognate neutral loss MS^3 spectrum of the phosphopeptide LGYTGENGGGQSEpSPVKDETPR exemplifying the additional sequence information gained of an MS^2 spectrum compared to a MS^3 spectrum. A: MS^2 spectrum (Xcorr: 2.424); B: MS^3 spectrum (Xcorr: 3.248). Exact localisation of the phosphorylation site is only possible with the MS^3 spectrum in this case. Only ions above the noise signal were annotated.

nases gene family. We aligned protein sequences belonging to the same family from A. thaliana (At) and Sesamum indicum (Si) (Figure 4). Shown is the region surrounding the phosphorylation site. The region of interest is part of a proposed NADP+ binding domain [31] and displays high homology. The serine phosphorylation site is conserved in 6 out of 8 sequences. The two remaining sequences (At4 and Si1) show no serine at the site of interest and might belong to a separate group.

Discussion

The approach described in this work is suitable to reliably identify and routinely screen for serine/threonine phosphorylation in plant proteins. Several lines of evidence are integrated into the strategy thus making unambiguous identification of protein phosphorylation possible: (i) Enrichment for phosphoproteins based on the affinity of phosphate to MOAC, (ii) a specific phosphostain reveals phosphorylation of the proteins and confirms enrichment, (iii) gel separation of the proteins helps to guarantee high confidence in protein identification, and (iv) a highly selective method based on mass spectrometry specific for phosphorylation is used for site identification.

Phosphorylated proteins are enriched by MOAC, a method that can be easily adapted to suit the sample of interest since it is inexpensive and the components needed are widely available. For demonstration MOAC phosphoprotein enrichment was applied to *A. thaliana* leaf proteins [18], *A. thaliana* seed material and *Chlamydomonas reinhardtii* cell cultures (this study, Figure 2).

Table 1: Identified phosphorylation sites in seed proteins. Gene expression data were derived from a previous study [38].

Accession nr.	Description	Protein sequence coverage	Site of phosphorylation (designated as "p")	Seed specific expression
At5g52300.1	low-temperature-responsive 65 kD protein (LTI65)/desiccation-responsive protein 29B (RD29B)	65.8% (M_r 65971 Da)	MKVTDEpSPDQKSR ESDINKNpSPARFGGESK LPLSGGGpSGVKETQQGEEK LGYTGENGGGQSEpSPVKDETPR GAVTSWLGGKPKpSPR EAHQEPLNpTPVSLLSATEDVTR	+
At3g12960.1	expressed protein similar to seed maturation protein from [Glycine max]	93% (M_r 9464 Da)	DIKDIKGTRTDDpSPR.-	+
At1g01100.1	60S acidic ribosomal protein P1 (RPP1A)	81.2% (M_r 11162 Da)	KKDEPAEEpSDGDLGFGLFD.-	-
At2g27710.1	60S acidic ribosomal protein P2 (RPP2B)	87% (M_r 11444 Da)	KEEKEEpSDDDMGFSLFE.-	-
At5g40420.1	glycine-rich protein/oleosin	42.2% (M_r 21279 Da)	HFQFQpSPYEGGR	+
At1g07985.1	Expressed protein	46.5% (M_r 16461 Da)	KLVDKVVGSSSPTNIHpSK	
At5g50600.1	short-chain dehydrogenase/reductase (SDR) family protein similar to sterol-binding dehydrogenase steroleosin from [*Sesamum indicum*]	61% (M_r 39087 Da)	STLYPESIRTPEIKpSD.- ELGpSPNVVTVHADVSKPDDCRR	+
At1g29350.1	expressed protein	14.6% (M_r 90879 Da)	SGpSTHFSSTDSGNFQGK	No data
At4g25580.1	stress-responsive protein-related contains weak similarity to Low-temperature-induced 65 kDa protein	57.8 % (M_r 66520 Da)	RGAPTLTPHNTPVSLLpSATEDVTR GAPTLpTPHNTPVSLLSATEDVTR	+

The following staining with a phosphospecific fluorescent dye is a quick and easy to use method to detect protein phosphorylation and its changes on gels. However, mass spectrometry-based site identification leads to more detailed information about site specific regulation. This holds especially true for proteins phosphorylated at multiple sites [5,6]. While it is not possible with the fluorescent stain to define if a protein is singly or multiply phosphorylated and on which amino acid the phosphate moiety is located, this can be done by mass spectrometry [32]. Consequently, enriched proteins are digested and the peptides are analysed in experiments in which an additional fragmentation event (MS^3) is triggered when a peptide looses phosphoric acid during the first fragmentation step (MS^2). The combined information of MS^2 and MS^3 data is then used to obtain high quality data about the peptide sequence and its phosphorylation site (see also Figure 3A and 3B).

Since this is a proof of concept study we did not aim at identifying all phosphorylation sites present in the enriched fraction but at setting up a robust and convenient workflow for the analysis of *in vivo* protein phosphorylation in plants. The major drawback of the method in its current state is that preferably phosphorylation sites of high abundant phosphoproteins are detected and that often the protein can be assigned reliably but data on the phosphopeptides are lacking. However, this might be circumvented by separating complex mixtures using established chromatography prior to MOAC and/or by coupling a phosphopeptide enrichment step to the protein enrichment. Another problem could be dephosphorylation occurring during or prior to gel separation or during sample storage. This might also explain differences in the reproducibility test. Nevertheless, most of the tested sites could be reproducibly found in a second experiment thus reconfirming the robustness of the method. However, if the sample amount is not limiting separation and digestion after enrichment might also be performed in solution without SDS-PAGE. This, of course would lead to a missing confirmation step in the strategy.

Albeit the phosphorylation of *A. thaliana* seed proteins probably plays a crucial role in seed development and dormancy [33,34], to our knowledge this is the first time a broad approach has been used to identify phosphorylation sites in seed proteins. In more than half of the identified seed derived phosphopeptides (9 out of 16) the identified phosphorylation sites are directly neighboured by a proline. Additionally, in three peptides the sites are located adjacent to aspartic acid residues. This could be due to enhanced cleavage at proline and aspartic acid residues during the fragmentation process in the mass spectrometer which has been described before [35]. Interestingly, we did find dominant neutral loss in the respective MS^2 spectra, even though it was reported recently that proline and aspartic acid containing phosphopeptides exhibit a less pronounced neutral loss of H_3PO_4 during fragmentation in a Q-Tof mass spectrometer [36]. This apparent difference might be explained by the different instrument types used in the studies or by the special nature of the investigated phosphopeptides. Both

Table 2: Reproducibility test of phosphorylation site identification. +: detected; -: not detected.

Protein	At5g52300.1 low-temperature-responsive 65 kD protein		jgi code 153417 putative protein	
Mw [Da]	65971		56515	
Experiment	1	2	1	2
Sequence coverage [%]	65.8	67.5	29. 1	28.0
Peptide 1	GAVTSWLGGKPKpSPR	+	KLESAApTVAER	+
Peptide 2	LGYTGENGGGQSEpSPVKDETPR	+	VAVAPPSRPGpSGK	+
Peptide 3	LPLSGGGpSGVKETQQGEEK	+	SGpSAKVAVAPSR	-
Peptide 4	EAHQEPLNpTPVSLLSATEDVTR	+		
Peptide 5	ESDINKNpSPARFGGESK	-		
Peptide 6	MKVTDEpSPDQKSR	-		

serine-proline (SP) and serine-aspartate (SD,E) containing phosphorylation sites are postulated as putative kinase substrates [15]; SP is a MAP-kinase motif which coincides with the importance of MAP-kinases in cellular processes [37] (see also http://www.neb.com/nebecomm/tech_reference/protein_tools/protein_kinase_substrate_recognition.asp). A majority of the identified phosphoproteins appear to be seed specific since their expression is reported to be highly dominant if not exclusively expressed during this developmental stage (see table 1 and [38]).

Multisite phosphorylation seems to be quite common as indicated by the fact that more than one phosphorylation site was found in 3 of the 9 proteins. The sites identified on the two ribosomal proteins are the same c-terminal sites identified in mouse [29] thus adding another evidence for the conservation of phosphorylation sites throughout different species [39]. Phosphorylation of these ribosomal P-proteins at their c-terminal end has also been proposed for *Saccharomyces cerevisiaea*, *Rattus norvegicus*, *Trypanosoma cruzei*, and Zea mays [40]. It has been shown that phosphorylation of these proteins leads to accelerated degradation in yeast [41] and this might also hold true for *A. thaliana*. Interestingly, another two of the identified phosphopeptides, belonging to proteins otherwise unrelated to the ribosomal P-proteins were also found to be phosphorylated at their c-terminal end indicating either the accessibility of c-terminal phosphopeptides for enrichment, digestion and detection or a general pattern, putatively for protein degradation.

The identification of a phosphorylated protein with homology to a short chain sterol dehydrogenase (Sop2) seems to be especially interesting for seed development. These proteins are thought to play a vital role in plant signal transduction elicited by sterols [31,42] and therefore their phosphorylation/dephosphorylation might have large implications on seed development. Serine 95 in the peptide ELGpSPNVVTVHADVSKPDDCRR derived from Sop2 (which we showed to be phosphorylated in *A. thaliana*) is highly conserved in five out of six homologues in *A. thaliana* as well as in one out of two homologues in *Sesamum indicum* (see Figure 4). It is located in the $NADP^+$ binding domain of the protein [31] and might be important for enzyme specificity and selectivity.

Methods

Chemicals

All chemicals were from Sigma (München, Germany). The aluminium hydroxide was purchased as aluminium hydroxide hydrate (ordering nr. A-1577; Sigma).

Seeds

A. thaliana (ecotype Columbia) seeds were taken from an in-house seed stock.

Chlamydomonas reinhardtii culture

C. reinhardtii (wildtype CC-125) was grown for 7 days in 16/8 hour light/dark regime in liquid culture containing TAP Medium [43]. Cells were harvested in the light period and centrifuged for 10 min at 4000 rpm. The supernatant was discarded and the Pellet was used for the extraction of proteins.

Denatured protein extraction from A. thaliana seeds and C. reinhardtii

A. thaliana seed proteins and C. reinhardtii proteins were extracted by adding a mixture of three volumes buffer-saturated phenol (15 ml) and one volume 50 mM Hepes-KOH, pH 7.2 containing 1% β-mercaptoethanol, 40 % sucrose, and 40 mM NaF (5 ml) to 2 g of seed material ground in liquid nitrogen or to a pellet derived from 100 ml of C. reinhardtii culture, respectively. After mixing for 20 minutes at 4°C, protein was precipitated out of the phenol phase with five volumes ice-cold acetone over night at -20°C. The pellet was washed twice with 100% ice-cold methanol and air dried for 15 min.

```
Si2 gi|16033752| IGESLAYEYAKRGACLVLAARRERSLQEVAERARDLGSPDVVVVRADVSKAEDCRKVVDQ 118
At1 gi|18423187| IGEQLAYEYACRGACLALTARRKNRLEEVAEIARELGSPNVVTVHADVSKPDDCRRIVDD 118
At5 gi|15232779| IGEHVAYEYAKKGAYLALVARRRDRLEIVAETSRQLGSQNVIIIPGDVSNVEDCKKFIDE 118
At6 gi|15232777| IGEHVAYEYAKKGAKLALVARRKDRLEIVAETSRQLGSGDVIIIPGDVSNVEDCKKFIDE 117
At3 gi|15241203| IGEHLAYEYARRGAYLTLVARREDRLQVVADRCRKLGSPDVAVVRGDVSVIKDCKRFVQE 118
At2 gi|15241261| IGEALAYEYGKRGAYLALVDIRGEPLFHVAALAELYGSPEVLPLVADVSKLQDCERFIRA 118
At4 gi|15234888| IGEQIAYEYAKRGANLVLVARREQRLRVVSNKAKQIGANHVIIIAADVIKEDDCRRFITQ 120
Si1 gi|21311775| IGEQIAYQYAKRGANLVLVARREHRLRGISENARRLGAPNVLIMAADVVKEEECRRFINE 119
                 *** :**:*. :** *.*.  *   *  ::  ..  *: .*  : .**   .:*.:.:
```

Figure 4
Sequence alignment of different short chain sterol dehydrogenases. Si: *Sesamum indicum*; At: *Arabidopsis thaliana*. The phosphopeptide is shown in red and the site of phosphorylation in bold.

Enrichment of phosphoproteins using Metal Oxide Affinity Chromatography (MOAC)

The pellet obtained by denatured protein extraction was dissolved in 1.5 ml incubation buffer containing 30 mM MES, 0.2 M potassium glutamate, 0.2 M sodium aspartate, 0.25% Chaps, and 8 M urea with a pH of 6.1. 1.5 ml of a 0.5 mg/ml protein solution was used for 30 min incubation with 80 mg of $AlOH_3$ at 4°C ($AlOH_3$ was washed before once with the incubation buffer described above). The incubation was followed by five washing steps of 1.6 ml and one step of 0.8 ml. Finally, proteins were eluted from the matrix by incubation with 800 µl of 100 mM potassium pyrophosphate buffer pH 9.0 containing 8 M urea for 30 min at RT. Proteins were precipitated with methanol/chloroform prior to gel loading as described by Wessel & Fluegge [44] and about one half of the eluted protein fraction was loaded onto the gel.

Determination of protein content

Protein concentrations were determined via the dye-binding method of Bradford as described previously [45], using ovalbumin as a standard. Each measurement was made in triplicate and the mean values were used.

SDS-PAGE, Phosphostaining, and Coomassie staining

Pellets derived from the methanol/chloroform precipitation were dissolved in 2 × SDS sample buffer (90 mM Tris, pH 6.9, 20% Glycerin, 2% SDS, 0.02% bromophenolblue, and 100 mM DTT) and approximately 30 µg were subjected to SDS-PAGE (800 µl MOAC eluate fraction was precipitated and dissolved in 40 µl of 1 × SDS sample buffer). After the separation proteins were stained with Pro-Q diamond stain (Invitrogen, Karlsruhe, Germany) essentially following the instructions in the manual. For specificity, we found exchanging fixation solution once (after 30 min) and leaving the gels in the fixation solution over night to be crucial steps. Phosphorylation was visualised using a chemdoc station and a 550 nm filter. After visualisation the phosphostain gels were washed three times with ddH_2O and stained with coomassie.

In-gel tryptic digest

Seed protein spots exhibiting a strong signal after phosphostaining were excised and digested over night with trypsin as described before [46]. Peptides were extracted from the gel in three consecutive steps using increasing percentages of acetonitrile (5 %, 50 %, 90% each containing 1% formic acid).

Protein and phosphorylation site identification using nano LC-linear-iontrap-MS

Peptides were loaded directly onto a ProteoSpher® Micro column (0.5 mm × 15 mm) at a flow rate of 3 µl/min and separated in a 85 min gradient from 80% A (0.1% formic acid, and 2.5 % acetonitrile in water) to 100 % B (0.1% formic acid in methanol). Separation and measurements were performed with a nano-LC-pump (Agilent 1100) coupled to an LTQ ion trap (Thermoelectron) with a nano-ESI-source. The voltage was applied directly to the analyte solution using a T-piece. To identify tryptic peptides, phosphopeptides and phosphorylation sites, automatic data-dependent acquisition was performed consisting of a full scan (m/z 400–2000), a subsequent MS^2, and a neutral loss scan of 98, 49, or 32.7 (H_3PO_4 for the +1, +2, and +3 charged ions, respectively) in the five most abundant MS^2 fragments. An MS^3 scan was automatically collected on the corresponding neutral loss fragments of the MS^2 scan events. Peptides were identified by searching the spectra against an A. thaliana database http://www.arabidopsis.org/ or against the C. Chlamydomonas database version 2.0 http://genome.jgi-psf.org/chlre2/chlre2.download.html containing trypsin and keratin sequences using the Sequest algorithm (ThermoElectron, Dreieich, Germany) and filtering the results with an Xcorr of 2.0, 2.5, and 3.5 for singly, doubly, and triply charged ions, respectively. Protein hits were accepted when at least three peptides with the corresponding Xcorr criteria described above were identified. The spectra derived from phosphopeptides were verified manually (charge state and identification of MS^2 and MS^3 spectra were checked for their concordance).

Sequence alignment

Sequences were derived from NCBI http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein and alignment was performed using ClustalW http://www.ebi.ac.uk/clustalw/ with the default settings.

Authors' contributions

FW carried out the optimization of MOAC with seed material, phosphorylation site identification using MS^2 and MS^3 and participated in writing and drafting the manuscript. WW drafted the conception of this study, advised throughout the project and participated in writing and drafting the manuscript.

Acknowledgements

We thank Megan McKenzie for revising the manuscript. We thank the Max Planck Society for financial support and Mirko Glinski, Stefanie Wienkoop, and Mark Stitt for valuable discussion. Finally, we thank Daniel Karcher, Dietrich Köster, and Christian Bölling for providing *C. reinhardtii*, Regina Stark and Karsten Oelkers for technical assistance.

References

1. Agrawal GK, Yonekura M, Iwahashi Y, Iwahashi H, Rakwal R: System, trends and perspectives of proteomics in dicot plants Part II: Proteomes of the complex developmental stages. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005, 815:125-136.
2. Rubin CS, Rosen OM: **Protein phosphorylation.** *Annu Rev Biochem* 1975, **44:**831-887.
3. Ma H: **Protein phosphorylation in plants: enzymes, substrates and regulators.** *Trends Genet* 1993, **9:**228-230.
4. Glinski M, Weckwerth W: **The Role of Mass Spectrometry in Plant Biology.** *Mass Spectrometry Reviews* 2005, **in press:**.
5. Glinski M, Romeis T, Witte CP, Wienkoop S, Weckwerth W: **Stable isotope labeling of phosphopeptides for multiparallel kinase target analysis and identification of phosphorylation sites.** *Rapid Commun Mass Spectrom* 2003, **17:**1579-1584.
6. Glinski M, Weckwerth W: **Differential Multisite Phosphorylation of the Trehalose-6-phosphate Synthase Gene Family in Arabidopsis thaliana: A Mass Spectrometry-based Process for Multiparallel Peptide Library Phosphorylation Analysis.** *Mol Cell Proteomics* 2005, **4:**1614-1625.
7. Martin K, Steinberg TH, Goodman T, Schulenberg B, Kilgore JA, Gee KR, Beechem JM, Patton WF: **Strategies and solid-phase formats for the analysis of protein and peptide phosphorylation employing a novel fluorescent phosphorylation sensor dye.** *Comb Chem High Throughput Screen* 2003, **6:**331-339.
8. Schulenberg B, Goodman TN, Aggeler R, Capaldi RA, Patton WF: **Characterization of dynamic and steady-state protein phosphorylation using a fluorescent phosphoprotein gel stain and mass spectrometry.** *Electrophoresis* 2004, **25:**2526-2532.
9. Schulenberg B, Arnold B, Patton WF: **An improved mechanically durable electrophoresis gel matrix that is fully compatible with fluorescence-based protein detection technologies.** *Proteomics* 2003, **3:**1196-1205.
10. Steinberg TH, Agnew BJ, Gee KR, Leung WY, Goodman T, Schulenberg B, Hendrickson J, Beechem JM, Haugland RP, Patton WF: **Global quantitative phosphoprotein analysis using Multiplexed Proteomics technology.** *Proteomics* 2003, **3:**1128-1144.
11. Goodman T, Schulenberg B, Steinberg TH, Patton WF: **Detection of phosphoproteins on electroblot membranes using a small-molecule organic fluorophore.** *Electrophoresis* 2004, **25:**2533-2538.
12. Salih E: **Phosphoproteomics by mass spectrometry and classical protein chemistry approaches.** *Mass Spectrom Rev* 2004:DOI 10.1002/mas.20042.
13. Zeller M, Konig S: **The impact of chromatography and mass spectrometry on the analysis of protein phosphorylation sites.** *Anal Bioanal Chem* 2004, **378:**898-909.
14. Laugesen S, Bergoin A, Rossignol M: **Deciphering the plant phosphoproteome: tools and strategies for a challenging task.** *Plant Physiol Biochem* 2004, **42:**929-936.
15. Nuhse TS, Stensballe A, Jensen ON, Peck SC: **Phosphoproteomics of the Arabidopsis plasma membrane and a new phosphorylation site database.** *Plant Cell* 2004, **16:**2394-2405.
16. Aro EM, Rokka A, Vener AV: **Determination of phosphoproteins in higher plant thylakoids.** *Methods Mol Biol* 2004, **274:**271-285.
17. Heintz D, Wurtz V, High AA, Van Dorsselaer A, Reski R, Sarnighausen E: **An efficient protocol for the identification of protein phosphorylation in a seedless plant, sensitive enough to detect members of signalling cascades.** *Electrophoresis* 2004, **25:**1149-1159.
18. Wolschin F, Wienkoop S, Weckwerth W: **Enrichment of phosphorylated proteins and peptides from complex mixtures using metal oxide/hydroxide affinity chromatography (MOAC).** *Proteomics* 2005, **[Epub ahead of print]:**.
19. Weckwerth W, Willmitzer L, Fiehn O: **Comparative quantification and identification of phosphoproteins using stable isotope labeling and liquid chromatography/mass spectrometry.** *Rapid Communications in Mass Spectrometry* 2000, **14:**1677-1681.
20. Jaffe H, Veeranna, Pant HC: **Characterization of serine and threonine phosphorylation sites in beta-elimination/ethanethiol addition-modified proteins by electrospray tandem mass spectrometry and database searching.** *Biochemistry* 1998, **37:**16211-16224.
21. McLachlin DT, Chait BT: **Improved beta-elimination-based affinity purification strategy for enrichment of phosphopeptides.** *Anal Chem* 2003, **75:**6826-6836.
22. Goshe MB, Conrads TP, Panisko EA, Angell NH, Veenstra TD, Smith RD: **Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses.** *Anal Chem* 2001, **73:**2578-2586.
23. Klemm C, Schroder S, Gluckmann M, Beyermann M, Krause E: **Derivatization of phosphorylated peptides with S- and N-nucleophiles for enhanced ionization efficiency in matrix-assisted laser desorption/ionization mass spectrometry.** *Rapid Communications in Mass Spectrometry* 2004, **18:**2697-2705.
24. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF: **Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry.** *Proc Natl Acad Sci U S A* 2004, **101:**9528-9533.
25. DeGnore JP, Qin J: **Fragmentation of phosphopeptides in an ion trap mass spectrometer.** *J Am Soc Mass Spectrom* 1998, **9:**1175-1188.
26. Olsen JV, Mann M: **Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation.** *Proc Natl Acad Sci U S A* 2004, **101:**13417-13422.
27. Giorgianni F, Beranova-Giorgianni S, Desiderio DM: **Identification and characterization of phosphorylated proteins in the human pituitary.** *Proteomics* 2004, **4:**587-598.
28. Jin WH, Dai J, Zhou H, Xia QC, Zou HF, Zeng R: **Phosphoproteome analysis of mouse liver using immobilized metal affinity purification and linear ion trap mass spectrometry.** *Rapid Commun Mass Spectrom* 2004, **18:**2169-2176.
29. Shu H, Chen S, Bi Q, Mumby M, Brekken DL: **Identification of phosphoproteins and their phosphorylation sites in the WEHI-231 B lymphoma cell line.** *Mol Cell Proteomics* 2004, **3:**279-286.
30. Medzihradszky KF, Darula Z, Perlson E, Fainzilber M, Chalkley RJ, Ball H, Greenbaum D, Bogyo M, Tyson DR, Bradshaw RA, Burlingame AL: **O-sulfonation of serine and threonine: mass spectrometric detection and characterization of a new posttranslational modification in diverse proteins throughout the eukaryotes.** *Mol Cell Proteomics* 2004, **3:**429-440.
31. Lin LJ, Tzen JT: **Two distinct steroleosins are present in seed oil bodies.** *Plant Physiol Biochem* 2004, **42:**601-608.
32. Wolschin F, Lehmann U, Glinski M, Weckwerth. W: **An integrated strategy for identification and relative quantification of site-specific protein phosphorylation using liquid chromatography coupled to MS2/MS3.** *Rapid Commun Mass Spectrom* 2005, **in press:**.
33. Borisjuk L, Rolletschek H, Radchuk R, Weschke W, Wobus U, Weber H: **Seed development and differentiation: A role for metabolic regulation.** *Plant Biology* 2004, **6:**375-386.
34. Walker-Simmons MK: **Protein kinases in seeds.** *Seed Science Research* 1998, **8:**193-200.
35. Wysocki VH, Tsaprailis G, Smith LL, Breci LA: **Mobile and localized protons: a framework for understanding peptide dissociation.** *J Mass Spectrom* 2000, **35:**1399-1406.
36. Salek M, Di Bartolo V, Cittaro D, Borsotti D, Wei J, Acuto O, Rappsilber J, Lehmann WD: **Sequence tag scanning: a new explorative strategy for recognition of unexpected protein alterations by nanoelectrospray ionization-tandem mass spectrometry.** *Proteomics* 2005, **5:**667-674.
37. Samaj J, Baluska F, Hirt H: **From signal to cell polarity: mitogen-activated protein kinases as sensors and effectors of cytoskeleton dynamicity.** *J Exp Bot* 2004, **55:**189-198.
38. Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, Lohmann JU: **A gene expression map of Arabidopsis thaliana development.** *Nat Genet* 2005, **37:**501-506.
39. Weckwerth W, Selbig J: **Scoring and identifying organism-specific functional patterns and putative phosphorylation sites in protein sequences using mutual information.** *Biochem Bioph Res Co Biochem Bioph Res Co* 2003, **307:**516-521.

40. Rodriguez-Gabriel MA, Bou G, Briones E, Zambrano R, Remacha M, Ballesta JP: **Structure and function of the stalk, a putative regulatory element of the yeast ribosome. Role of stalk protein phosphorylation.** *Folia Microbiol (Praha)* 1999, **44:**153-163.
41. Nusspaumer G, Remacha M, Ballesta JP: **Phosphorylation and N-terminal region of yeast ribosomal protein P1 mediate its degradation, which is prevented by protein P2.** *Embo J* 2000, **19:**6075-6084.
42. Lin LJ, Tai SS, Peng CC, Tzen JT: **Steroleosin, a sterol-binding dehydrogenase in seed oil bodies.** *Plant Physiol* 2002, **128:**1200-1211.
43. Harris EH: **The Chlamydomonas Sourcebook.** *Academic Press, Inc, San Diego, CA, USA* 1989.
44. Wessel D, Flugge UI: **A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids.** *Anal Biochem* 1984, **138:**141-143.
45. Bradford MM: **Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding.** *Analytical Biochemistry* 1976, **72:**248-254.
46. Otto A, Thiede B, Muller EC, Scheler C, WittmannLiebold B, Jungblut P: **Identification of human myocardial proteins separated by two-dimensional electrophoresis using an effective sample preparation for mass spectrometry.** *Electrophoresis* 1996, **17:**1643-1650.

Letter to Editor

To the Editor-in-Chief Sir,

Methionine oxidation in peptides– a source for false positive phosphopeptide identification in neutral loss driven MS^3

Florian Wolschin and Wolfram Weckwerth

Affiliation:
Max Planck Institute of Molecular Plant Physiology, 14424 Potsdam, Germany
Email addresses:
wolschin@mpimp-golm.mpg.de
weckwerth@mpimp-golm.mpg.de

* Corresponding author:
Dr. Wolfram Weckwerth
Max Planck Institute of Molecular Plant Physiology, 14424 Potsdam, Germany
Email: weckwerth@mpimp-golm.mpg.de;
Phone: +49 331 567 8109
Fax: +49 331 567 8134

For phosphopeptide identification the dominant neutral loss of H_3PO_4 from serine or threonine during CID (colllisional induced dissociation) can be exploited to trigger a further MS^3 fragmentation step in ion trap mass spectrometers [1].
The detection of such a neutral loss is reported to be a specific marker for phosphorylated peptides, and the subsequent MS^3 fragmentation step on the dominant neutral loss fragment ion provides additional sequence information [2-5].
The former phosphorylation site is then found in the MS^3 spectrum of a peptide as a dehydroalanine residue in case the peptide was phosphorylated on serine or as a dehydrobutyric acid residue if the peptide was phosphorylated on threonine,

respectively (see figure 1). Both residues can be easily identified by including minus 18 Da for serine and threonine as variable modifications when using database searching programs like Sequest or Mascot.

In a recent study neutral loss from peptides containing alkylated methionine was described to mimic phosphorylation [6]. Following this line we investigated other processes with potentially misleading neutral losses. Here, it is demonstrated that methionine oxidation in peptides can mimic phosphorylation and thus lead to the automatic recording of MS^3 spectra in neutral loss driven phosphopeptide screening experiments (see figure 1). These spectra can in turn lead to the identification of false positive phosphopeptides following database searching.

Tryptic peptides of a *A. thaliana* plant protein, which was first separated by SDS-PAGE [3] were analyzed with reversed phase chromatography coupled to a LTQ mass spectrometer (Thermo, San Diego). MS spectra spectra were acquired and MS^2 triggered for the three most intense peaks. Automatic neutral loss driven MS^3 fragmentation was triggered if [M – 98]n+ ions were observed in one of the five most abundant peaks. The isolation window was set to 3 Da, collision energy to 35 (arbitrary units), and the activation time to 50 ms.

Obtained spectra were analysed by using Sequest (Thermo, San Diego) directed automatic database searching including STY (serine/threonine/tyrosine) plus 80 Da (phosphorylation), ST minus 18 Da (loss of phosphoric acid leading to a dehydroalanine residue or a dehydrobutyric acid residue in a peptide, respectively), and M (methionine) plus 16 Da (oxidation) always as variable modification. An *A. thaliana* sequence database (http://www.arabidopsis.org/) was used as the protein sequence source. For refined searches M minus 48 Da (corresponding to the loss of methane sulfenic acid) was included (see figure 1). The Xcorr threshold was set to 2.0, 2.5, and 3.5 for singly, doubly, and triply charged peptides, respectively. In many analyses the charge state of the peptide precursor in the MS^2 spectrum did not match to the observed neutral loss when assuming phosphoric acid as neutral loss. In many cases the dominant loss corresponded to a [M -64]n+ ion, which is typical for the neutral loss of methane sulfenic acid from a peptide containing oxidized methionine [7, 8]. However, the same dehydrobutyric acid residue can also be observed following neutral loss of phosphoric acid from a peptide containing a phosphothreonine residue instead of an oxidized methionine (figure 1). Since the neutral loss of methane sulfenic acid from a doubly charged precursor ion (32 Da) is similar to the neutral loss

of phosphoric acid from a triply charged phosphopeptide (32.7 Da) within the mass window of a typical ion trap experiment (2-3 Da), an MS^3 spectrum was triggered automatically. The resulting MS^3 spectra were matched to phosphopeptides in the database and thereby led to the identification of false positives, because of sequence similarity including only a methionine/threonine substitution (see figure 2) or by chance (spectra not shown). Using only neutral loss triggered MS^3 spectra led to false positive peptide identification of up to 30% of the total number of MS^3 spectra. Figure 2 shows the MS^3 spectrum and false positive identification of malate dehydrogenase At1g53240 (putative phosphopeptide and corresponding Xcorr shown in italic letters) after database search using common modification settings as described above. In figure 3 the corresponding MS^2 spectrum displays a dominant [M-64]2+ ion that triggered MS^3. Using this MS^2 spectrum for database search led to the identification of another malate dehydrogenase At3g15020. This is an isoform which differs from the first isoform only by the replacement of a threonine through a methionine in the corresponding tryptic peptide. Careful analysis of the data revealed that the observed dominant peak can be explained by the loss of methane sulfenic acid from a doubly charged precursor ion ([M-64]2+). Since loss of methane sulfenic acid leads to the formation of a dehydrobutyric acid in a peptide (see fig. 1), correct identification is only possible when comparing MS^2 and MS^3 data. However, inclusion of methionine minus 48 Da (corresponding to methionine sulfoxide minus 64) as modification in the MS^3 based database searching revealed the methionine containing peptide (see figure 2).

The *A. thaliana* proteome encodes over 1600 tryptic sequences (at least 4 amino acids long) that differ only in the replacement of methionine for threonine and are therefore putative candidates for the observed phenomenon. Furthermore, methionine oxidation is a frequently occurring process during protein preparation. It is the automated manner in which neutral loss driven MS^3 and database searching are used that gives rise to the identification of false positives. The inclusion of zoom scans helps to determine the charge state accurately. However, the authors are not aware of a software that couples the neutral loss function to a certain charge state. We conclude that false positives resulting from methionine oxidation can best be avoided by comparing precursor charge states and fragmentation of MS^2 and MS^3 spectra.

Literature

1. DeGnore J P, Qin J. *J Am Soc Mass Spectrom* 1998; **9:** 1175.
2. Giorgianni F, Beranova-Giorgianni S, Desiderio D M. *Proteomics* 2004; **4:** 587.
3. Wolschin F, Weckwerth W. *Plant Methods* 2005; **1:** 9.
4. Wolschin F, Lehmann U, Glinski M, Weckwerth W. *Rapid Commun Mass Sp* 2005; **19:** 3626.
5. Gruhler A, Olsen J V, Mohammed S, Mortensen P, Faergeman N F, Mann M, Jensen O N. *Mol Cell Proteomics* 2005.
6. Kruger R, Hung C W, Edelson-Averbukh M, Lehmann W D. *Rapid Commun Mass Sp* 2005; **19:** 1709.
7. Reid G E, Roberts K D, Kapp E A, Simpson R I. *J Proteome Res* 2004; **3:** 751.
8. Guan Z, Yates N A, Bakhtiar R. *J Am Soc Mass Spectrom* 2003; **14:** 605.

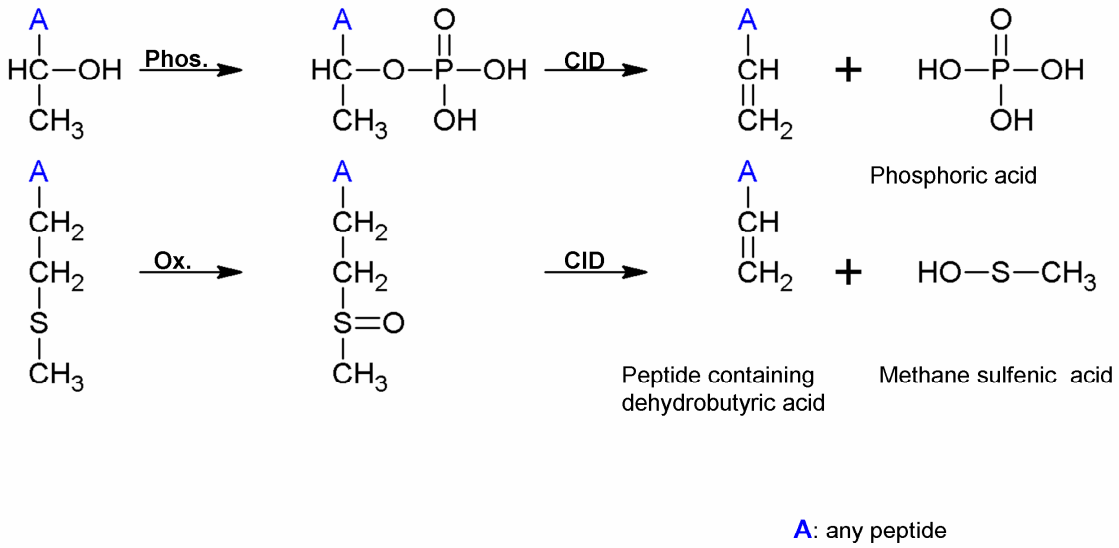
Figures:

Fig.1: Formation of a dehydrobutyric acid in a peptide from phosphorylated threonine or oxidized methionine by neutral loss in MS^2. Phos.:Phosphorylation; Ox.: Oxidation; CID.

Fig.2: Identification of peptide sequences from a MS^3 spectrum. The MS^3 spectrum was automatically triggered on a neutral loss peak and searched against an *A. thaliana* database. The identification in cursive denotes the search result with commonly implemented modifications (STY +80, ST -18, M +16). The identification in normal form denotes the search result considering loss of methane sulfenic acid (methionine -48 Da).T^: Threonine -18 Da; M#: Methionine -48 Da.

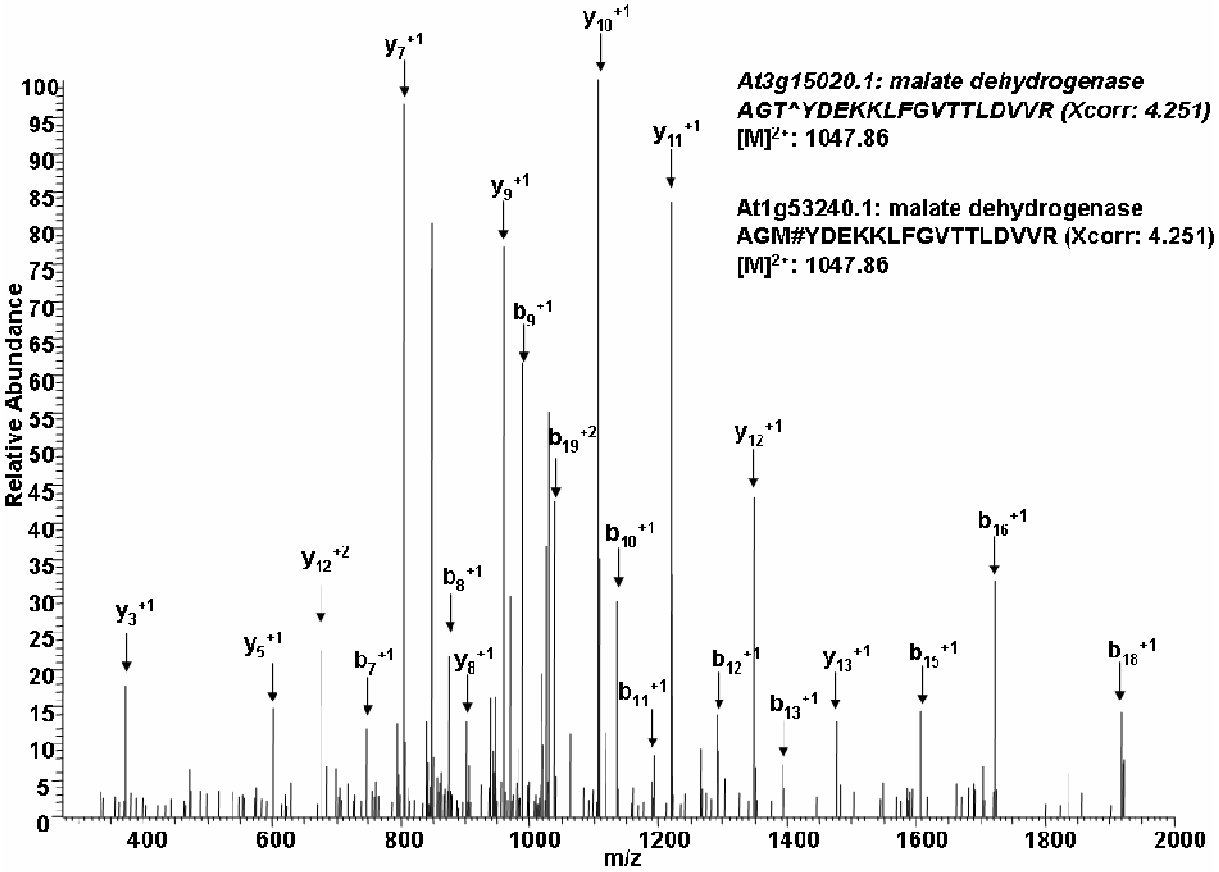
Fig.3: MS^2 spectrum identified after database search to belong to the malate dehydrogenase At1g53240.1 containing oxidized methionine M*.

Figure 1


A
HC—OH
CH3
Phos.
A
O
HC—O—P—OH
CH3
OH
CID
A
CH
CH2
+
O
HO—P—OH
OH
Phosphoric acid
A
CH2
CH2
S
CH3
Ox.
A
CH2
CH2
S=O
CH3
CID
A
CH
CH2
+
HO—S—CH3
Peptide containing
dehydrobutyric acid
Methane sulfenic acid
A: any peptide

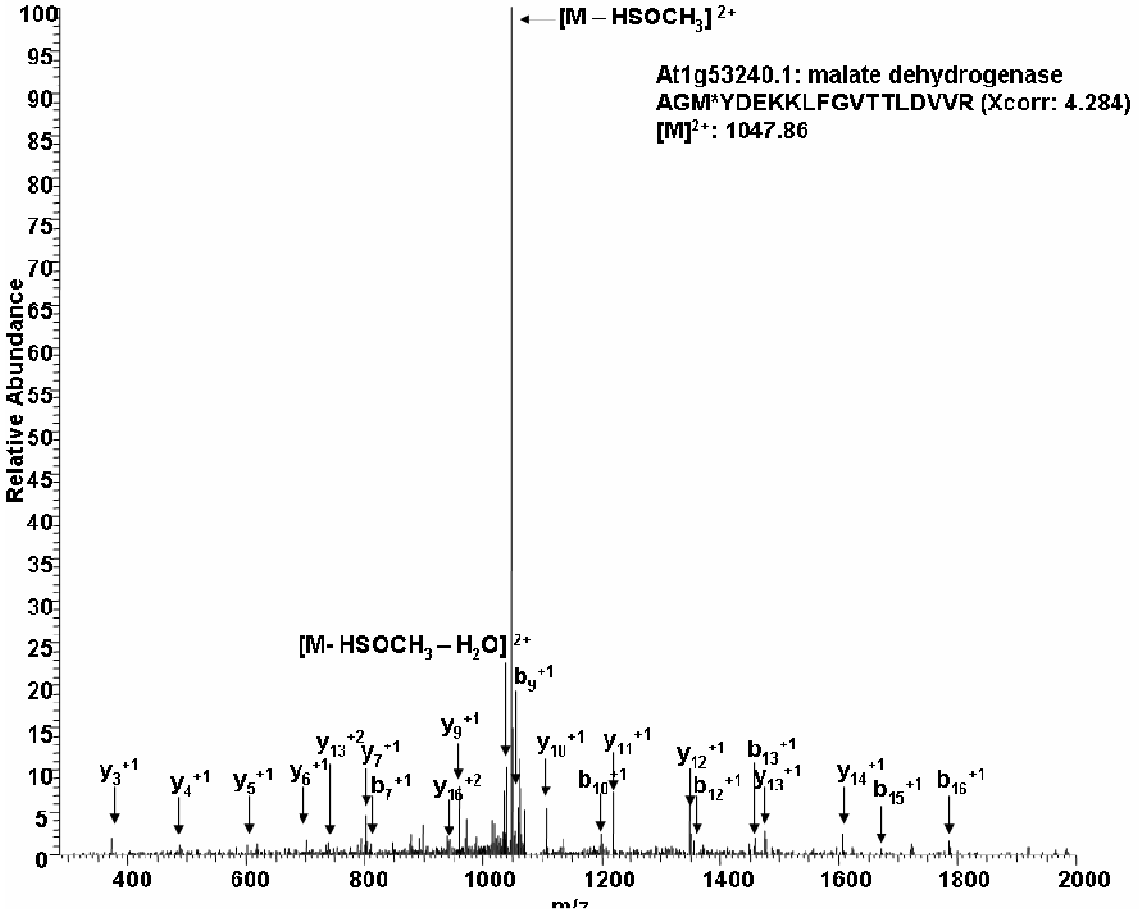
145

Figure 2


At3g15020.1: malate dehydrogenase
AGT^YDEKKLFGVTTLDVVR (Xcorr: 4.251)
[M]2+: 1047.86
At1g53240.1: malate dehydrogenase
AGM#YDEKKLFGVTTLDVVR (Xcorr: 4.251)
[M]2+: 1047.86
Relative Abundance
m/z
y7+1
y10+1
y11+1
y9+1
b9+1
b19+2
y12+1
b10+1
b16+1
y12+2
b8+1
y3+1
y5+1
b7+1
y8+1
b12+1
y13+1
b15+1
b18+1
b11+1
b13+1

150

Figure 3



155

160

165

Chapter IV: Quantification of protein phosphorylation

While the detection of phosphorylation and the identification of phosphorylation sites is highly important to get a first glance at the phosphorylation state of proteins and delivers the fundament of phosphorylation research, quantification of these phosphorylation events gives an even deeper insight into complex protein regulation. This quantification can be performed on the proteome, protein, peptide, and site-specific level. However, the quantification of protein phosphorylation imposes great challenges to the experimenter because of the low abundance of phosphoproteins and the complex fragmentation of phosphopeptides observed in tandem mass spectrometry based analyses.

Methods based on ^{32}P labelling or global anti-phospho amino acid antibodies only detect and quantify the phosphorylated version of a protein but leave the unphosphorylated forms undetected, while other methods can also determine phosphorylation stoichiometries (the percentage of phosphorylated proteins). Several approaches have been developed to cope with the inherent difficulties associated with quantification and these approaches are used to unravel complex protein regulation achieved by phosphorylation/dephosphorylation events.

IV.1. Imaging methods

One example is a general approach relying on the attachment of a fluorescent dye to phosphorylated residues followed by an intensity based measurement of the phosphoprotein/-peptide dye complex. An increase in the number of phosphorylated residues or the phosphorylation stoichiometry results in a measurable increase of the fluorescence [184]. Non-phosphorylated species of the same proteins can be observed and quantified using dyes like Sypro-ruby or silver stain. An equivalent method is the one that makes use of antibodies (see chapter III). Phosphorylation is visualised using a secondary antibody coupled to an enzyme or a fluorescent dye and increase or decrease of phosphorylation is determined by the changes in the product produced by the enzyme or by the changes in the fluorescence. Global phospho-specific antibodies as well as antibodies directed against specific phosphorylated antigens and their non-phosphorylated counterparts can be used as primary antibodies in this approach to obtain results on phosphorylation levels. However, the drawbacks of the use of these antibodies, which were discussed in the previous chapters, have to be taken into account.

Similarly, a change in phosphorylation can be monitored by comparing scintillation counts of two samples labelled with ^{32}P-orthophosphate *in vivo* or with radioactive γ-ATP *in vitro*. Nevertheless, uptake and incorporation of ^{32}P is probably never complete and endogenous as well as previously bound non-radioactive ^{33}P always presents a problem for quantification. Relative quantification of the global phosphorylation state of a protein or peptide can be obtained readily using these methods and always bearing in mind the limitations of these techniques discussed in previous chapters.

However, all of these approaches except the one relying on specific antibodies, have the major disadvantage that site-specific information is lacking. This issue becomes especially important, when one is looking at a protein or peptide, which can be phosphorylated at multiple residues and each phosphorylation event may represent another activity state. This is also true for approaches were ICP (Inductively coupled plasma) is used to determine the content of phosphorus in a protein [185, 186]. Therefore, the described methods are only suitable if one is interested in global changes of phosphorylation for example in whole proteomes or if one can rule out the possibility of multisite phosphorylation.

IV.2. Mass spectrometry based approaches

More recent methods rely on the use of mass spectrometry because of its selectivity, sensitivity and capability of direct determination of phosphorylation sites. There is a wide variety of different methods based on mass spectrometry ranging from methods depending on chemical derivatisation to the use of external and internal standards for the investigation of changes in phosphorylation some of which have been applied to plant protein phosphorylation. While many methods have been developed for the quantification of peptides and proteins in general, descriptions in the following paragraphs are restricted to the ones which were primarily designed for phosphopeptides/phosphoproteins. For a more general discussion on protein quantification the reader is referred to [187] and to the introduction of this thesis.

IV.2.1. Strategies relying on chemical derivatisation

Overall there are two different kinds of derivatisation techniques, which are commonly used for the relative quantification of phosphorylation. The first is based on beta elimination and Michael addition (see chapters II and III). In this approach two samples

(proteins or peptides) to be compared are derivatised with differently deuterated forms (i.e. d0/d5) or different isotopes (i.e. C12/C13) of the same nucleophile, mixed, proteins are digested and peptides are subjected to chromatography, which is directly coupled to a mass spectrometer. The ratio of the corresponding peptide pairs leads to a relative quantification of phosphorylation. In the second derivatisation strategy peptides derived from two different samples (control and treated) are derivatised with the undeuterated and the deuterated form of methanol thus converting carboxyl residues into their corresponding methyl esters in the process of methylesterification (see chapter II). Phosphopeptides are then usually enriched to overcome the low level of phosphorylation and peptide pairs are analysed by mass spectrometry as described above. Major drawbacks of these approaches are the side reactions observed in chemical derivatisation (see chapter II) and elution differences of the peptide pairs in chromatography depending on the type of derivatization agent. Especially deuterated forms of the peptides are known to exhibit a slight retention time shift in chromatography when compared to their non-deuterated counterparts and this sometimes impairs direct comparison of peptide pairs [188]. Also, the number of samples that can be compared is limited due to the limited availability of differentially deuterated or isotopically coded tags.

Other strategies involving stable isotope labelling by metabolic labelling or tagging (like SILAC or iTRAQ) are not primarily applied to phospho-proteins/peptides and are not further discussed (for reviews see [187, 189]).

IV.2.2. Methods based on inductively coupled plasma mass spectrometry

Inductively coupled plasma mass spectrometry (ICP-MS) is a technique that is capable of analysing atoms instead of molecules. A very powerful implementation of this technique was presented by Lehmann and coworkers who used peptide separation via HPLC directly coupled to ICP-MS to assess the global phosphorylation degree in a sample by monitoring the phosphorus and the sulphur ratio [190-192]. In this approach peptides are eluted sequentially into the ICP-MS and decomposed into their atoms in the plasma. In real time, sulphur as well as phosphorus traces are recorded and related to each other.

Verification of phosphoprotein enrichment and global phosphorylation degree of A. thaliana and C. reinhardtii

To evaluate the success of the developed phosphoprotein enrichment method depending on aluminum hydroxide and to assess the phosphorylation degree of different tissue types of *A. thaliana* the approach consisting of HPLC separation coupled to ICP-MS

designed by Lehmann and cowokers was used [190]. To this end about 40 µg of total protein from the same sources collected before and after enrichment desalted by subjecting it to SDS-PAGE for 15 minutes at 120 V. Lanes (about 1 cm x 0.3 cm) were cut out of the gels and digested with trypsin. Peptides were separated on a C18 column and analysed using ICP-MS by Dr. Ralf Krüger (DKFZ, Heidelberg).

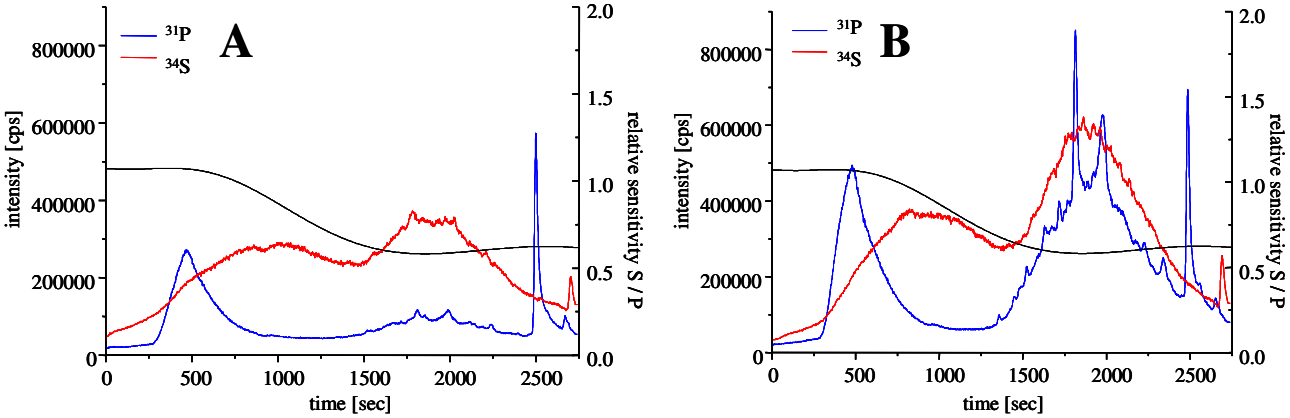


Fig. 12: ***LC-ICP measurements of digested proteins before (A) and after (B) enrichment using aluminum hydroxide. The red line indicates the sulphur (^{34}S) and the blue line the phosphorus (^{31}P) trace. The black line represents the S/P sensitivity factor, which is dependent on solvent composition.***

The ICP measurements confirmed an enrichment of phosphorylated proteins. This can be clearly seen when the ^{34}S trace, which serves as a general elution reference of peptides, is compared to the ^{31}P trace in the example (figure 12). After enrichment (figure 12 B) the ^{31}P trace and the ^{34}S trace show a considerably higher overlap than before enrichment (14 A) strongly indicating an increase of peptide bound phosphate. It was possible to calculate an enrichment factor by comparing moles of phosphorus per moles of sulphur in these proteins before and after enrichment for various samples. This factor was determined to vary between 2 - and over 20 fold enrichment leading to a final degree of phosphorylation between 40 and 100 mol phosphorus per mol protein (table V). This means that 40 to 100 % of all proteins are phosphorylated at one site after enrichment. The enrichment factor varies between tissue types and organisms, which can be explained by the following arguments:

➢ Only heat and acid stable amino acid-phosphate bonds (namely serine, threonine, and tyrosine –phosphate) can be detected using this method since proteins are heated before SDS-PAGE, proteins are fixed and destained using acidic solutions, and peptides are separated and

analysed under acidic conditions. Consequently, the phosphorylation degree is underestimated in samples containing a higher degree of heat and acid labile amino acid phosphates.

➢ Different tissue types and organisms have different overall protein patterns with varying complexity. It seems probable that in some cases a highly complex pattern interferes to a larger extent with affinity chromatography than a pattern of low complexity since the possibility that the highly complex mixture contains proteins that interfere with the desired enrichment is higher.

Table V: Phosphorylation degree before and after MOAC enrichment

	P/protein [mol/mol]	*$^{31}P/^{34}S$ [mol/mol]*	*P/S [mol/mol]*	*Enrichment factor P/S*
A. thaliana leaf	*0.075 (S)* *0.051 (FT)* *0.42 (E)*	*0.09222* *0.06307* *0.51468*	*0.003956* *0.002706* *0.022080*	*5.6*
	0.048 (S)	*0.05885*	*0.002525*	*-*
	0.061 (S)	*0.07466*	*0.003203*	*-*
A. thaliana seeds	*0.019 (S1)* *0.433 (E1)*	*0.02332* *0.53170*	*0.001000* *0.022810*	*22.8*
	0.021 (S2) *0.434 (E2)*	*0.02583* *0.53241*	*0.001108* *0.022841*	*20.7*
A. thaliana cell culture	*0.145 (S2)* *0.61 (E2)*	*0.17788* *0.74841*	*0.007612* *0.032107*	*4.2*
	0.253 (S3) *0.572 (E3)*	*0.31067* *0.701812*	*0.013328* *0.030108*	*2.3*
C. reinhardtii cell culture	*0.267 (S1)* *1.0 (E1)*	*0.32779* *1.22728*	*0.014062* *0.052650*	*3.7*
	0.25 (S2)	*0.30614*	*0.013133*	*-*
	0.225 (S3)	*0.27644*	*0.011859*	*-*

S = sample, E = eluate, FT = flow-through.

It should be noted that these calculations are based on the assumption that the compositions of the proteomes before and after enrichment are comparable with respect to the sulphur content. Since phosphostaining (see chapter III) and the use of antibodies directed against the phosphorylated version of a MAP-kinase (about 8-fold enrichment from A. thaliana cell cultures; personal communication, Dr. Jan Heise) independently confirmed successful enrichment by aluminum hydroxide this assumption seems to be justified.

Besides the confirmation of phosphoprotein enrichment these experiments for the first time delivered information for plant protein phosphorylation on a whole organism scale. Interestingly, the global phosphorylation degrees vary considerably (see figure 13). This gives insights into the importance protein phosphorylation might have throughout different developmental stages. The lowest degree was calculated for *A. thaliana* seeds and the highest for *A. thaliana* cell cultures and *C. reinhardtii*. It seems consequent that rapidly dividing tissue has a higher degree of phosphorylation. While *A. thaliana* cell cultures are immotile, *C. reinhardtii* has a flagellum and is motile. Since flagellar dependent movement of prokaryotes

as well as of eukaryotes seems to be inherently linked to protein phosphorylation [193-196] this may explain the slightly higher degree of phosphorylation in *C. reinhardtii* compared to *A. thaliana* cell cultures. This assumption is supported by the fact that seven flagellar proteins could be identified after MOAC enrichment and that a very recent study on protein phosphorylation in *C. reinhardtii* shares the impression that protein phosphorylation is quite abundant in flagella of *C. reinhardtii* [177]. However, the values only differ slightly and these differences might not be statistically valid.

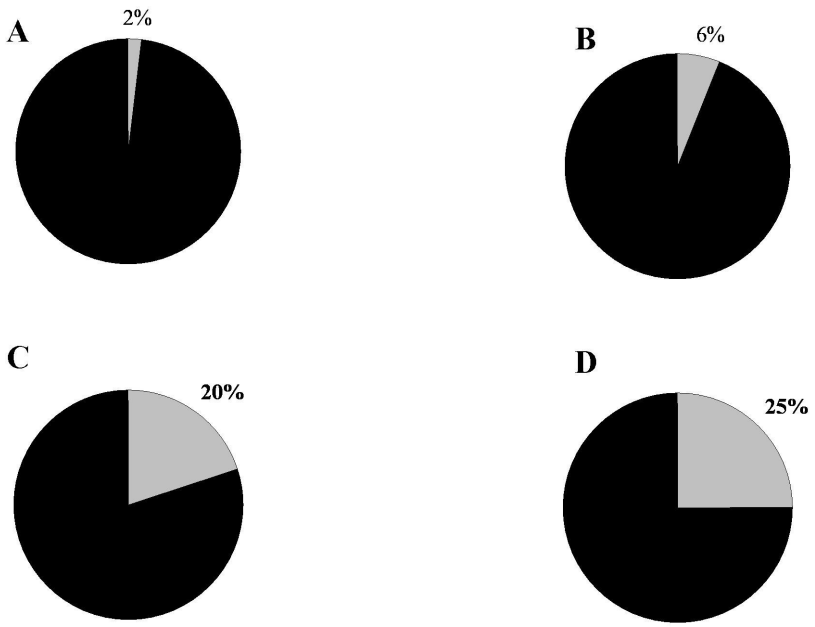


Fig. 13: ***Visualisation of the values from table V. The P/ S values were converted into percentages of phosphorylated proteins under the assumption that each phosphoprotein is phosphorylated at one site completely. A: A. thaliana seeds; B: A. thaliana leaf; C: A. thaliana cellcultures; D: C. reinhardtii.***

It is interesting to note that the cells of a usually multicellular organism (indeed, also the cell cultures of *A. thaliana* are not just single cells but conglomerations of several cells) obviously have a degree of phosphorylation comparable to the unicellular alga *C. reinhardtii*. When compared to other eukaryotic cells, which are postulated to have phosphorylation degrees of about 30 %, the values obtained for cell cultures do not deviate to a large extent. This suggests that a certain phosphorylation level is crucial throughout different kingdoms of live and that this value is indeed about 30 %.

On the other hand, the low degree for seed tissue surprises because of the necessity for nutrient storage. Apparently, the stored phosphorus is not bound to proteins to a large extent in seeds. Indeed, it is known that the major source of phosphorus in seeds is phytic acid [197], which probably explains why plants do not have to rely on phosphorus bound to proteins for seed germination.

IV.2.3. Methods involving standard peptides

The methods involving internal standards can be divided in two sections: one involving the addition of a standard, which is not derived from the sample [198] and one where peptides derived from the protein of interest are used [67, 68, 199-201]. Integrated ion traces of the phosphopeptide are then standardised using the ion traces of other, non-phosphorylated peptides. These ratios can be compared for multiple samples and thereby differences in the extent of phosphorylation can be determined. Since these methods are non-invasive in that they do not require chemical modification of the protein of interest they are especially well suited to disclose the native phosphorylation of a protein.

Phosphorylation of sucrose-phosphate synthase from A. thaliana

Despite being a well studied example of regulation by protein phosphorylation (see introduction) many questions remain open. For example, no temperature dependency of SPS phosphorylation had ever been reported even though sucrose is an important osmoprotectant and no phosphorylation site of *A. thaliana* was known. To investigate phosphorylation of sucrose-phosphate synthase from *A. thaliana* an approach was developed to relatively quantify phosphorylation changes in any serine or threonine phosphorylated protein of interest in a robust manner. A recombinant version of the enzyme was incubated *in vitro* with several protein extracts derived from *A. thaliana* adapted to different temperature conditions. First results were obtained after SDS-PAGE followed by Pro-Q phosphospecific staining. Proteins were then digested and separated by HPLC. MS^2 and MS^3 ion traces of the phosphopeptide were integrated and standardized using unmodified peptides of the same protein found in the same MS run.

The standardized ion traces were compared for the protein derived from different treatments (for which the unmodified peptides do not change but the phosphorylated peptides do) and thus differences in phosphorylation could be monitored. It was found that quantification based on MS^3 data is generally preferable over the less selective MS^2 data integration and can in principle be done automatically. However, when intensities are not sufficient for quantification on the MS^3 level, as was the case for the 32°C samples MS^2

manual integration gives a good measure for relative quantification. A raw estimation of differences in the extent of phosphorylation could be made from the signal obtained by phosphorus-specific fluorescent labelling (Pro-Q, see chapter III). The obtained values confirmed the results from the phosphostain (figure 14). However, no value could be determined for the 4°C samples due to over saturation.

Since this was an *in vitro* experiment one can only extrapolate these results to the *in vivo* situation. It should be noted that in spinach SPS the phosphorylation of the homologues serine 158 is reported to lead to a decreased activity of this enzyme. Since sucrose is regarded as an osmoprotectant, theory would suggest SPS phosphorylation at serine 159 to be decreased in the cold.

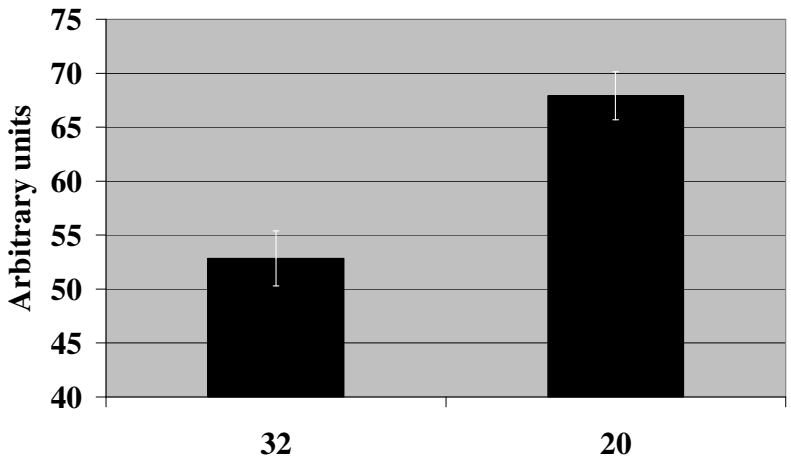


Fig. 14: Evaluation of phosphostain signal intensity: SPS incubated with extracts derived from plants treated at 32°C; 20: SPS incubated with extracts derived from plants treated at 20°C. For SPS incubated with extracts derived from plants treated at 4°C no data was obtained due to over-saturation. Y-axis: arbitrary units related to pixel intensities X-axis: temperature [°C].

As the opposite was observed three possible explanations are possible: the *in vitro* findings do not adequately reflect the *in vivo* situation, the single time point of sampling only delivered a snapshot and might not adequately reflect the overall situation, or regulation under cold stress in *A. thaliana* differs markedly from spinach. For a more detailed discussion see the relevant publication at the end of this chapter.

In addition to the published results phosphorylation stoichiometry was calculated. For the samples generated with plant extracts treated at 32 °C, which displayed the lowest phosphorylation level for all samples and thus the highest signal for the non-phosphorylated peptide, a calculation of stoichiometry was possible. The respective standard peptides (ISSVDVFENWFAQHK in its phosphorylated and unphosphorylated version) were measured and the peak values obtained for the 32 °C samples corrected as described in [201]. A phosphorylation degree of about 40% was determined in this way. In the other two samples

low signal-to-noise ratios and interfering peaks made efficient quantification of the non-phosphorylated peptide impossible. However, since the relative phosphorylation calculated for the 20°C and 4°C samples was about 1.7 fold respectively 2.6 fold of the value determined for the 32°C samples, this gives phosphorylation stoichiometries of about 68 % and 104 %. The 104 % can easily be explained by considering the standard deviations of about 20 % for every mean value. These results show that phosphorylation stoichiometries of about 100 % were achieved in these experiments with the 4°C samples suggesting once more that the observed reaction was not an unspecific artefact. Indeed, the absolute value of the peak area for the phosphopeptide derived from the 4 °C samples was the same (within error bars) as the combined peak area values for the phosphorylated plus unphosphorylated peptide of the 32 °C sample.

Appendix chapter IV

Wolschin, F., Lehmann, U., Glinski, M. & Weckwerth, W. An integrated strategy for identification and relative quantification of site-specific protein phosphorylation using liquid chromatography coupled to MS^2/MS^3. *Rapid Commun Mass Spectrom* 19, 3626-3632 (2005).

RAPID COMMUNICATIONS IN MASS SPECTROMETRY
Rapid Commun. Mass Spectrom. 2005; **19**: 3626–3632
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/rcm.2236

RCM

An integrated strategy for identification and relative quantification of site-specific protein phosphorylation using liquid chromatography coupled to MS^2/MS^3

Florian Wolschin, Ute Lehmann, Mirko Glinski and Wolfram Weckwerth*
Max Planck Institute of Molecular Plant Physiology, 14424 Potsdam, Germany

Received 12 August 2005; Revised 12 October 2005; Accepted 12 October 2005

Reversible and differential multisite protein phosphorylation is an important mechanism controlling the activity of cellular proteins. Here we describe a robust and highly selective approach for the identification and relative quantification of site-specific phosphorylation events. This integrated strategy has three major parts: visualisation of phosphorylated proteins using fluorescently stained polyacrylamide gels, determination of the phosphorylation site(s) using automatic MS^3 triggered by the loss of phosphoric acid, and relative quantification of phosphorylation by integrating MS^2- and MS^3-extracted ion traces using a fast-scanning, linear ion trap mass spectrometer. As a test case, recombinant sucrose-phosphate synthase (SPS) from *Arabidopsis thaliana* (At5g1110) was used for identification and quantification of site-specific phosphorylation. The identified phosphorylation site of the actively expressed protein coincides with the major regulatory *in vivo* phosphorylation site in spinach SPS. Site-specific differential *in vitro* phosphorylation of native protein was demonstrated after incubation of the recombinant protein with cold-adapted plant leaf extracts from *A. thaliana*, suggesting regulatory phosphorylation events of this key enzyme under stress response. Copyright © 2005 John Wiley & Sons, Ltd.

Covalent multisite protein phosphorylation has long been known to be a key post-translational regulatory modification; the ability to quantitatively assess the site-specific phosphorylation state of a protein, however, has held many challenges for investigators. Protein-specific phosphorylation can be detected by separating proteins on gels and staining with a phosphate-specific fluorescent dye.[1,2] Site-specific phosphorylation can best be addressed using selective mass spectrometric methods,[3] but neutral loss of phosphoric acid during the collision-induced dissociation (CID) fragmentation process often impairs thorough fragmentation of the peptide backbone and thus results in ambiguous database identifications.

Different approaches have been applied to circumvent this problem. A promising technique used by Syka *et al*.[4] makes use of a fragmentation approach called electron-transfer dissociation (ETD), which leaves the phosphorylated residue intact and thus allows more definitive site identification. The drawback of this approach, however, is its limited availability due to the sophisticated instrument modifications needed. An alternative is the replacement of the phosphate moiety with another, more stable group, a process that consists of two parts: removing the phosphate by beta-elimination and adding another nucleophile to the generated double bond in a Michael addition reaction.[5] However, this process sometimes leads to unwanted side reactions.[6,7] Another approach detects the characteristic neutral loss of phosphoric acid automatically and then triggers an additional MS^3 fragmentation step on the remaining precursor.[8,9] The resulting spectrum is often informative enough to determine the site of phosphorylation. As reported recently, however, to make reliable identifications using this approach, the data derived from MS^2 and MS^3 spectra should be combined.[10,11]

Even though phosphorylation site determination is an important step towards understanding site-specific protein regulation, phosphorylation *quantification* at the site of interest is the ultimate step revealing biological relevance. To obtain quantitative data on protein phosphorylation, different approaches have been described. One powerful method is based on liquid chromatography (LC) coupled to inductively coupled plasma mass spectrometry (ICP-MS).[12–14] With this approach, it is possible to determine the ratio of phosphate to other elements present in the protein of interest as well as of the absolute amount of phosphate present in a protein/peptide. Localisation of the phosphorylated residue, however, has to be performed by LC/MS in a separate experiment.

Other MS-based methods rely on beta-elimination and stable isotope labelling mentioned above,[5,15,16] or on chemical tagging of other amino acid residues.[17–19] After differential tagging, samples are mixed and the ratios of

**Correspondence to*: W. Weckwerth, Max Planck Institute of Molecular Plant Physiology, 14424 Potsdam, Germany.
E-mail: weckwerth@mpimp-golm.mpg.de
Contract/grant sponsor: The Max Planck Society.

Copyright © 2005 John Wiley & Sons, Ltd.

peptides are determined. However, with these derivatisation strategies only a limited number of samples can be compared and undesired side reactions during the procedure can lead to erroneous results (see above). Other approaches use internal standard peptides for quantitative analyses. Ruse *et al.*[20] added standard peptides to the tryptic digest of protein kinase C to obtain ratios of the phosphopeptide of interest to the synthetic peptide. A different strategy described by Tsay *et al.*[21] involves the digestion of a phosphorylated protein followed by peak integration from MS ion traces of the phosphopeptide and its non-phosphorylated cognate for relative quantification. These ratios can be compared for multiple samples. A similar but more refined approach was published recently by Steen *et al.*,[22] who used multiple peptides of the same protein rather than only the cognate non-phosphorylated peptide, thereby better describing the increase or decrease of phosphorylation.

Here, a novel integrative strategy is described for site-specific identification and quantification of protein phosphorylation using a fast-scanning linear ion trap mass spectrometer and precursor-targeted MS^2 and MS^3 capabilities. Using this strategy the *in vitro* phosphorylation site of recombinant *Arabidopsis thaliana* sucrose-phosphate synthase (SPS), a key enzyme in plant sucrose metabolism, was identified as serine 159. This site coincides with the phosphorylation site in spinach SPS known to be phosphorylated during a diurnal rhythm of the plant.[23] Incubation with cold-adapted leaf extracts led to differential site-specific phosphorylation of the recombinant native *A. thaliana* enzyme, suggesting a role of this phosphorylation site also in cold-acclimation.

EXPERIMENTAL

Chemicals

Microcystin was supplied by Cyano-Biotech (Berlin, Germany). Recombinant SPS was provided by Ute Lehmann. Sequencing-grade trypsin was from Roche (Mannheim, Germany). All other chemicals were from Sigma (München, Germany).

Growing and treatment of plants

A. thaliana plants (ecotype Columbia) were grown in a Percival growth chamber for 5 weeks at 20°C at a 12 h day/night rhythm. Light intensity was adjusted to 130 μE. After 5 weeks, plants were transferred to 4°C or to 32°C while control plants remained at 20°C. After another 72 h, 8 h into the light phase, plants were harvested directly into liquid nitrogen.

Three plants were pooled for each sample and the material was ground with liquid nitrogen.

Native protein extraction from leaves for *in vitro* phosphorylation of recombinant SPS

A. thaliana proteins were extracted from 300 mg of frozen leaf with 400 μL of incubation buffer containing 20% glycerol, 50 mM Tris pH 7.5, protease inhibitor mix from Sigma (final concentration: 100 μM bestatin, 20 μM pepstatin A, 2 mM (4-(2-aminoethyl)benzenesulfonyl fluoride, 30 μM (2*S*,3*S*)-3-(*N*-{(*S*)-1-[*N*-(4-guanidinobutyl)carbamoyl]3-methylbutyl}carbamoyl)oxirane-2-carboxylic acid, and 5 mM 1,10-phenanthroline), 20 mM leupeptin, 1 mM benzamidine, 60 mM NaF, 0.3 μM microcystin, and 5 mM dithiothreitol (DTT), for 20 min at 4°C. The material was centrifuged for 5 min at 14 000 rpm and the supernatant desalted on a Sephadex G-25 column using the buffer described above without glycerol. Protein content was determined using the Bradford test with ovalbumin as a standard.

In vitro phosphorylation assay of SPS using temperature-adapted leaf extracts

The *in vitro* assay was performed essentially as described previously.[5] Briefly, 35 μL of SPS in Tris buffer (~0.5 mg/mL) was mixed with 100 μL desalted leaf extract, 2 μL $MgCl_2$ (1 M), 2 μL $CaCl_2$ (0.1 M), and 10 μL adenosine triphosphate (ATP) (0.1 M) and incubated in a water bath for 15 min at 25°C. In the controls, ATP or SPS was replaced with extraction buffer (without glycerol).

The reaction was stopped by adding four volumes of methanol followed by methanol/chloroform precipitation according to Wessel.[24]

SDS-PAGE, phosphostaining, and Coomassie staining

Precipitated proteins were dissolved in 1× SDS sample buffer (45 mM Tris, pH 6.9, 10% Glycerin, 1% SDS, 0.01% bromophenol-blue, and 50 mM DTT) and approximately 30 μg was loaded per lane of 10% acrylamide gels. The separation was followed by a Pro-Q Diamond staining step (Invitrogen, Karlsruhe, Germany), essentially following the instructions in the manual. For increased specificity, it was necessary to change the fixation solution once after 30 min. The gels were then stored in the fixation solution overnight. Phosphostaining was performed the next day following the instructions of the manufacturer. Phosphorylated proteins were visualised using a Biorad Geldoc station (München, Germany) and an amber filter (Clare Chemicals, Dolores, USA).

The phosphostaining was followed by three washes with double-distilled H_2O. Proteins were then stained with Coomassie.

In-gel tryptic digest

Protein spots corresponding to SPS were excised, reduced with DTT, and alkylated with iodoacetamide, followed by overnight digestion with trypsin as described previously.[25] The next day, peptides were extracted from the gel in three consecutive steps using increasing percentages of acetonitrile (5, 50, and 90%, each containing 1% formic acid).

Chromatographic peptide separation and phosphorylation site identification using nano-LC coupled to a linear ion trap mass spectrometer

Peptides were loaded directly onto a ProteoSpher® Micro column (0.5 mm × 15 mm; Merck, Darmstadt, Germany) at a flow rate of 3 μL/min, and separated over 85 min through a gradient ranging from 20% to 100% buffer B (0.1% formic acid in methanol) with buffer A (0.1% formic acid and 2.5% acetonitrile in water). Separation and measurements were

Copyright © 2005 John Wiley & Sons, Ltd.

performed using an Agilent 1100 nano-LC pump (Böblingen, Germany) coupled to an LTQ ion trap (ThermoElectron, San Jose, USA) equipped with a nano-ESI source. The voltage was applied directly to the analyte solution using a T-piece. To identify tryptic peptides, phosphopeptides, and phosphorylation sites, automatic data-dependent acquisition was performed consisting of acquisition of a full MS^1 scan, a subsequent MS^2 scan, and a neutral loss MS^3 scan. The neutral loss MS^3 scan was triggered automatically when one of the five most intense MS^2 fragments corresponded to the precursor ion with a loss of 32.7, 49.0, or 98.0 Da (H_3PO_4 for the +3, +2, and +1 charged ions, respectively). Peptides were identified by searching the spectra against an *A. thaliana* database[26] containing trypsin and keratin sequences, using the Sequest algorithm (ThermoElectron, Dreieich, Germany) and filtering the results with Xcorr values of 2.0, 2.5, and 3.5 for singly, doubly, and triply charged ions, respectively. The spectra derived from phosphopeptides were inspected manually (charge state and identification of MS^2 and MS^3 spectra were checked for their concordance).

Relative quantification and normalisation of SPS phosphorylation

For normalisation of SPS phosphorylation to the total amount of SPS protein in the gel, additional tryptic peptides of the protein were identified. In the same LC/MS run precursor-targeted MS^2 scans for these internal reference peptides, as well as for the newly identified phosphopeptide ISpSVDFENWFAQHK (m/z 944.4), were recorded. Additionally, MS^3 scans were recorded for the phosphopeptide (transition m/z 944.4 → 895.3). Precursor-targeted MS^2 and MS^3 ion traces were integrated, and internal reference peptides for the normalisation procedure were identified as follows: individual integrated ion traces from each group of samples (4, 20, and 32°C) were divided by the median of all replicates. Peptides giving the most consistent results among all replicates (i.e., values after division by the median were close to one) were regarded as reliable internal reference peptides. Using sequential SPS protein dilutions and internal reference peptides it was ensured that the measured values were well within the linear range of the instrument. At least ten scans per peak were used for peak integration. The resulting peak area ratios were further normalised to total protein content in the *in vitro* phosphorylation assay. The peak area of the phosphopeptide of each sample was normalised by dividing the calculated area by the area of the internal reference peptides, and results were compared. Also, the ratios of the different reference peptides were checked for each run. The mean and standard deviation for each peptide ratio were calculated from at least four independent replicates.

RESULTS

Approach outline: detection and relative quantification of site-specific protein phosphorylation

An overview on the general strategy for the identification and relative quantification of serine/threonine phosphorylation is shown in Fig. 1. In this strategy, phosphorylation events

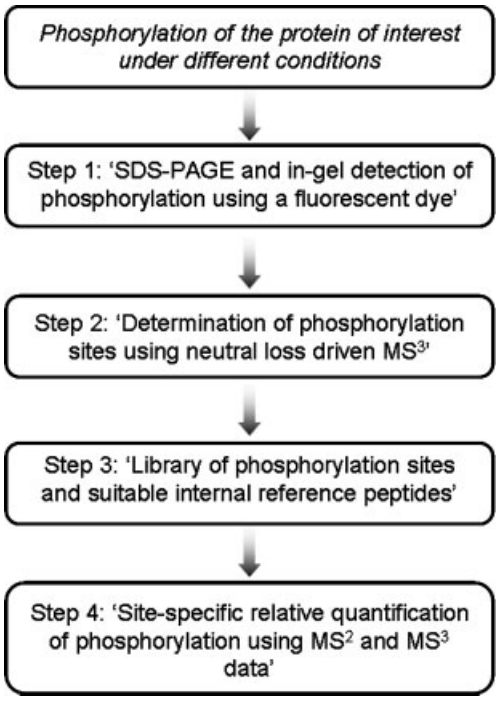


Figure 1. General overview of the approach used for detection and relative quantification of specific phosphorylation sites.

were detected using a fluorescent stain. Subsequently, phosphorylation sites were identified using neutral loss-driven MS^3. Phosphorylation levels were quantified in different samples by determining the ratio of precursor-targeted MS^2 or MS^3 ion traces of the phosphopeptide to the selected internal standard peptides. Detailed descriptions of these steps are given in the following sections.

Detection of *in vitro* phosphorylation of recombinant native SPS

Recombinant SPS was incubated *in vitro* with ATP and with protein extracts derived from plants conditioned at different temperatures (32, 20, and 4°C) for 72 h. Extracted proteins were then separated by sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS-PAGE) and stained with a phosphate-specific stain and with Coomassie (Fig. 2). The amount of SPS loaded per lane was constant for all of the samples (see Fig. 2). Different extents of phosphorylation in samples from plants conditioned at different temperatures (32, 20, and 4°C; Fig. 3(A)) were observed. The control treated without ATP does not show a signal in phosphostaining.

Identification of *in vitro* phosphorylation sites of SPS

The phosphorylated SPS protein band observed after SDS-PAGE was digested with trypsin, and the resulting peptides were subjected to a neutral loss experiment for phosphoric acid using an LTQ ion trap. In this experiment, the dominant neutral loss fragment observed in the MS^2 spectrum was automatically dissociated resulting in a more informative MS^3 spectrum of the phosphopeptide. Using this process the *in vitro* phosphorylation site of recombinant

Copyright © 2005 John Wiley & Sons, Ltd.

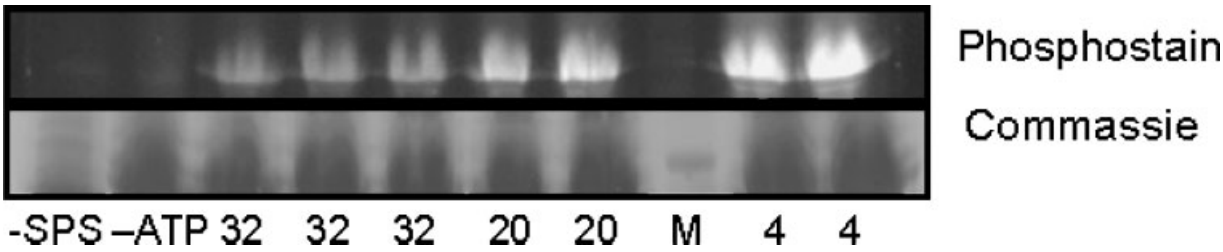


Figure 2. Visualisation of SPS phosphorylation using Pro-Q staining and imaging of total protein content by Coomassie staining. M = protein molecular weight marker; –SPS = incubation without SPS; –ATP = incubation without ATP; 32 = SPS incubated with extracts derived from plants treated at 32°C for 72 h; 20 = SPS incubated with extracts derived from plants treated at 20°C for 72 h; 4 = SPS incubated with extracts derived from plants treated at 4°C for 72 h: (A) Pro-Q staining and (B) Coomassie staining (~5 μg SPS were loaded per lane).

A. thaliana SPS was determined. Figure 3 shows MS^2 and MS^3 spectra of the tryptic SPS phosphopeptide IS-pS-VDVFENW-FAQHK. As a further control the charge states of the precursors and the detected fragment ions following the specified neutral losses were compared, to verify fragment spectra derived from neutral loss-driven MS^3.

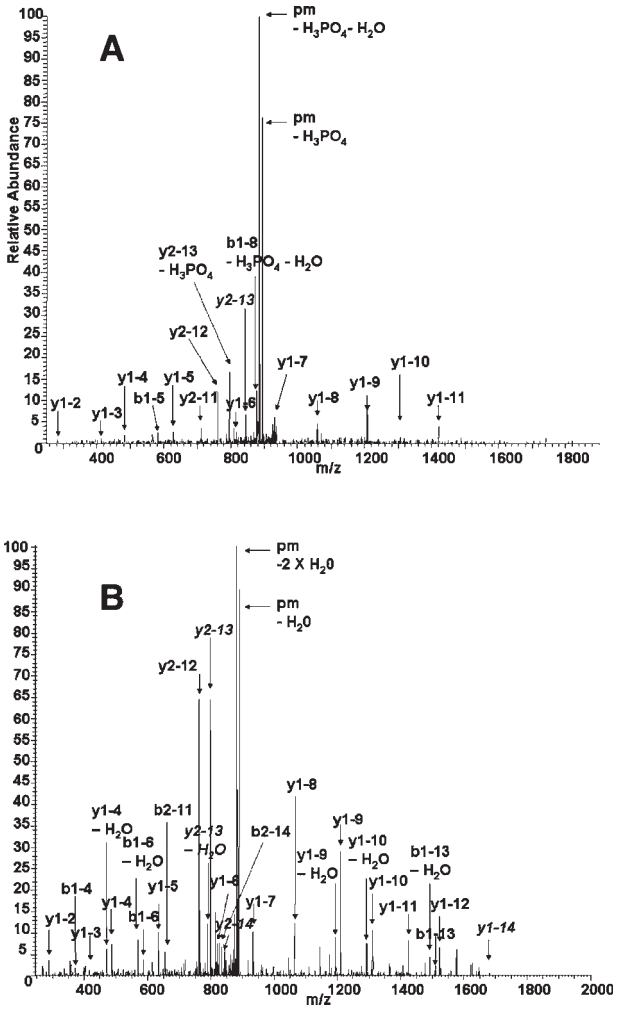


Figure 3. (A) MS^2 spectrum of the doubly charged phosphopeptide IS-pS-VDVFENWFAQHK (*m/z* 944.4) and (B) cognate neutral loss MS^3 spectrum triggered by the observation of loss of H_3PO_4 (*m/z* 944.4 → 895.3). Ion peaks annotated in italics denote phosphorylation in the MS^2 spectrum, and dehydroalanine in the MS^3 spectrum. Pm = precursor ion.

Relative quantification of site-specific SPS phosphorylation using highly selective precursor-targeted MS^2 or MS^3 scanning

Relative quantification was performed on the *in vitro* phosphorylated SPS. After the identification of the SPS phosphopeptide and other non-phosphorylated SPS peptides in a survey experiment consisting of MS, MS^2, and MS^3, we set up an experiment based solely on MS^2 and MS^3 scanning of precursor ions determined in the survey experiment. As can be seen in Fig. 4, the MS^3 extracted ion trace for the identified phosphopeptide is much more selective than MS^2, thereby increasing peak-purity and the accuracy of peak integration. However, at the limit of detection (LOD), as in the case for the 32°C samples, not enough scans per peak are sampled and peak integration is hampered.

MS^2 and MS^3 ion traces were integrated and ratios of (phosphopeptides/internal standard peptide) peak areas were calculated. Figure 5 shows the ratio of the phosphopeptide to an internal standard peptide. Normalisation to other internal peptides gave similar results throughout all LC/MS runs (data not shown, see also Experimental section for the selection of an appropriate internal standard). The non-phosphorylated peptide cognate ISSVDFENWFAQHK of the phosphorylated peptide could be observed, but was not used for normalisation purposes since reliable peak integration was impossible. Internal standard peptides showed constant peak areas for all experiments. Peptides that were found to have high deviations contained methionine and cysteine; these peptides are not suitable because of the variations in methionine oxidation and cysteine modification, and are therefore not recommended for quantification purposes.

The SPS phosphorylation site serine 159 showed varying degrees of phosphorylation depending on the temperature-adapted leaf tissue (see Fig. 5). The increase of phosphorylation correlates with cold-acclimated plants. This observation also holds true for the MS^3 trace, which could only be integrated for the 20 and 4°C samples since the number of scans/peak for the 32°C samples was not sufficient for reliable integration. Problems with low scan numbers per peak can, in principle, be overcome using a triple-quadrupole mass spectrometer with multiple reaction monitoring (MRM), as shown in a recent study investigating peptide

Copyright © 2005 John Wiley & Sons, Ltd.

library phosphorylation;[27] however, identification of the peptide is not possible using MRM. The MS^3 spectrum is a reliable source for identifying and quantifying phosphorylation sites with an optimised selectivity in terms of signal-to-noise (S/N) ratio (see Fig. 4). For low phosphorylation stoichiometries, however, the intensities of the MS^3 traces are often not strong enough, so MS^2 data are complementary. In our recent study the direct comparison of high-resolution phosphopeptide MRM transitions using a triple-quadrupole mass spectrometer showed higher sensitivity than precursor-targeted MS^2 and MS^3 scans on a linear ion trap.[27]

DISCUSSION

The approach described in this work is suitable to reliably identify and relatively quantify serine/threonine phosphorylation of a protein of interest. Several steps are used to achieve a high level of robustness and selectivity. Visualisation of phosphorylation is accomplished using a fluorescent dye which selectively stains phosphorylated proteins. At this level it is possible to obtain a preliminary impression of changes in phosphorylation. However, one can infer nothing about the site(s) of phosphorylation based on these data. Site-specific information is gained using a highly selective approach based on MS in which two different types of LC/MS^n experiments are conducted to increase selectivity and S/N ratio. In the first experiment, the peptides themselves are identified; in the second, the relative levels of phosphorylated and non-phosphorylated forms of the peptides are quantified at the MS^2 or MS^3 level based on the determined *m/z* values. Only peptides derived from the same sample as the phosphopeptides are used for normalisation, thus minimising run-to-run inconsistencies.

In this work, we took a close look at the site-specific dynamic phosphorylation pattern of sucrose-phosphate synthase (SPS) in *A. thaliana*. This enzyme catalyses a regulatory key step in sucrose metabolism in plants. In *A. thaliana*, SPS consists of a gene family with four members (At5g20280, At5g11110, At4g10120, and At1g04920). For our study we chose the isoform At5g1110 containing the putative phosphorylation site motif $RISS^{158}$ known from spinach SPS.[23] Following the proposed strategy we first identified the *in vitro* phosphorylation site of SPS as serine 159 using recombinant enzyme and an *in vitro* phosphorylation assay with plant leaf tissue.[5,27] The identified phosphorylation site in the *A. thaliana* SPS protein coincides with the conserved serine 158 from the spinach protein, indicating regulatory *in vivo* importance in *A. thaliana* as well.

Subsequently, we used the information of the exact phosphorylation site to set up a site-specific quantification procedure. A fast-scanning linear ion trap was used to perform MS^2 and MS^3 fragmentation on the phosphopeptide precursor ion. The resulting ion traces were integrated and normalised to internal standard peptides derived from the same protein digest and LC/MS run. Using this approach we investigated whether phosphorylation of serine 159 has any relation to abiotic stress responses such as cold-acclimation. SPS catalyses the formation of sucrose phosphate from the substrates UDP-glucose and fructose-6-phosphate; changes in its activity thus result in an altered ratio of monosaccharides to disaccharides. Since saccharides are known to function as osmoprotectants, changes in this ratio might be an adaptive response to cold stress. Indeed, the investigation of the amount of phosphorylation of serine 159 showed increased levels under cold-acclimation (see Fig. 5). Phosphorylation of the SPS serine 158 in spinach has been reported

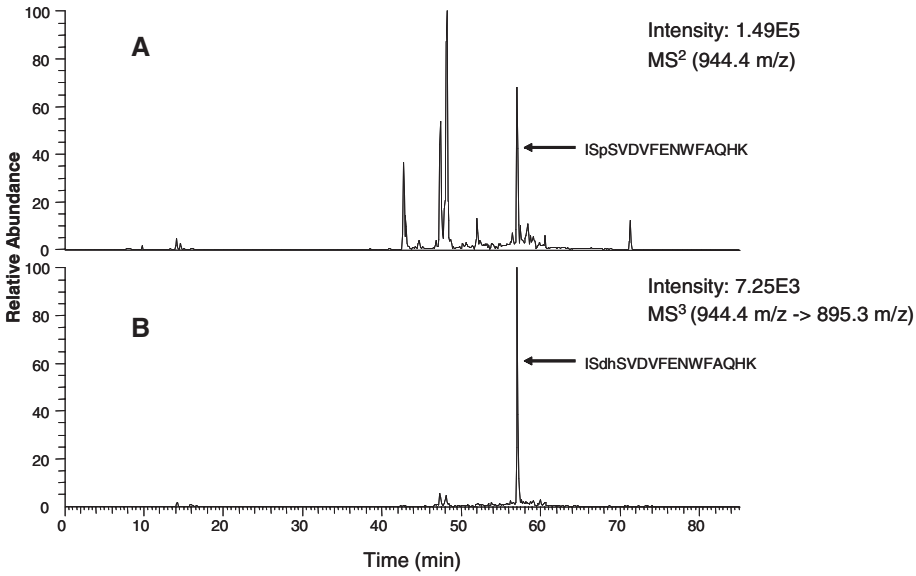


Figure 4. Chromatograms for precursor-targeted MS^2 and MS^3 scans. (A) Chromatogram acquired using MS^2 scans of the doubly charged ion of the phosphopeptide IS-pS-VDVFENWFAQHK (*m/z* 944.4). (B) Chromatogram acquired using MS^3 scans of the dehydroalanine peptide IS-dhS-VDVFENWFAQHK formed by neutral loss of H_3PO_4 in MS^2. Integration of the MS^3 ion trace leads to a better S/N ratio compared to integration of the MS^2 ion trace.

Copyright © 2005 John Wiley & Sons, Ltd.

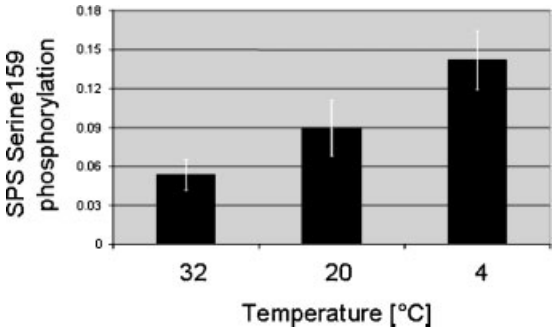


Figure 5. Relative quantification of *in vitro* phosphorylation of SPS. X-axis: temperature at which the plants were acclimated for 72 h. Y-axis: peak area ratios from the MS^2 ion traces of the phosphopeptide IS-pS-VDVFENWFAQHK (*m/z* 944.4) and of the internal reference peptide ELLWPHI-PEFVDR (*m/z* 827.2) used for normalisation. At least four independent replicates were used to determine each mean and standard deviation.

to decrease activation of SPS at night,[23,28,29] and chilling has been reported to delay SPS activity.[30] However, for interpretation of our results it is important to consider this as an *in vitro* experiment. Therefore, is it likely that an appropriate balance in SPS activity is achieved *in vivo* through both a specific kinase and a phosphatase(s). At present we are not able to discriminate the effects these counterbalancing forces have on the regulation of SPS activity in *A. thaliana*. However, site-specific relative quantification indicates that serine 159 plays a role in differential modification of SPS (At5g1110) under cold-acclimation.

Two other phosphorylation sites, which have been suggested for spinach SPS,[31] could not be confirmed for *A. thaliana* SPS in our studies. While the non-phosphorylated but not the phosphorylated homologue of the site surrounding S424 was found, the peptide containing the site homologue surrounding S229 could not be identified (neither in its phosphorylated nor in its non-phosphorylated form). The contribution of the two remaining possible sites to overall SPS phosphorylation cannot be excluded. However, since the differences in phosphorylation observed in the MS-based studies corresponded well with the observations on the gel, and since the observed site is a homologue of the major regulatory site in spinach,[29] we assume that the phosphorylation event with the main impact on SPS regulation was monitored.

While we used an *in vitro* phosphorylated recombinant protein for our analysis, the strategy discussed can also be applied to proteins purified from their *in vivo* sources as long as the phosphorylation is located on a serine or threonine and is preserved during purification, and as long as the stoichiometry is high enough to reliably detect the phosphopeptide. If the abundance of the phosphorylated protein is not high enough for detection, enrichment procedures can help to separate phosphorylated proteins or peptides from the non-phosphorylated ones and thus lead to mixtures of lower complexity. This can be achieved by using anti-phosphotyrosine/serine/threonine antibodies, immobilised metal affinity chromatography (IMAC), or the more recently introduced metal oxide affinity chromatography (MOAC).[32,33] Then, of course, one would have to add an external standard for quantification as described by Cutillas *et al*.[34] However, all other steps (detection of phosphorylation, phosphorylation site identification, and site-specific quantification by integrating MS^2 and MS^3 data) can be performed as described here. If the non-phosphosphorylated cognate can also be used for quantification purposes (i.e., the ion trace is suitable for reliable integration) it is possible to determine phosphorylation stoichiometries in analogy to the approach of Steen *et al*.[22] but using increased selectivity with MS^2 and MS^3 ion traces instead of MS ion traces (see Fig. 4(B)). A further extension is the use of MRM on a triple quadrupole.[27]

In conclusion, the method presented here enabled phosporylation site identification and relative quantification, revealing differential regulation of cold-acclimated sucrose-phosphate synthase in *A. thaliana*. Using widespread methods and instrumentation such as phosphostaining and ion trap or triple-quadrupole LC/MS, the overall process is adaptable to a wide range of similar experimental setups to monitor differences in site-specific protein phosphorylation.

Acknowledgements

We thank Megan McKenzie for revising the manuscript. We are indebted to the Max Planck Society for financial support and to Stefanie Wienkoop and Mark Stitt for valuable discussion. Furthermore, we thank Stanislav Laptev for technical assistance.

REFERENCES

1. Martin K, Steinberg TH, Goodman T, Schulenberg B, Kilgore JA, Gee KR, Beechem JM, Patton WF. *Comb. Chem. High Throughput Screen*. 2003; **6**: 331.
2. Steinberg TH, Agnew BJ, Gee KR, Leung WY, Goodman T, Schulenberg B, Hendrickson J, Beechem JM, Haugland RP, Patton WF. *Proteomics* 2003; **3**: 1128.
3. Glinski M, Weckwerth W. *Mass Spectrom. Rev*. 2005; in press.
4. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. *Proc. Natl. Acad. Sci. USA* 2004; **101**: 9528.
5. Glinski M, Romeis T, Witte CP, Wienkoop S, Weckwerth W. *Rapid Commun. Mass Spectrom*. 2003; **17**: 1579.
6. Whitaker JR, Feeney RE. *Crit. Rev. Food Sci. Nutr*. 1983; **19**: 173.
7. McLachlin DT, Chait BT. *Anal. Chem*. 2003; **75**: 6826.
8. DeGnore JP, Qin J. *J. Am. Soc. Mass Spectrom*. 1998; **9**: 1175.
9. Gruhler A, Olsen JV, Mohammed S, Mortensen P, Faergeman NJ, Mann M, Jensen ON. *Mol. Cell. Proteomics* 2005; **4**: 310.
10. Jin WH, Dai J, Zhou H, Xia QC, Zou HF, Zeng R. *Rapid Commun. Mass Spectrom*. 2004; **18**: 2169.
11. Giorgianni F, Beranova-Giorgianni S, Desiderio DM. *Proteomics* 2004; **4**: 587.
12. Wind M, Edler M, Jakubowski N, Linscheid M, Wesch H, Lehmann WD. *Anal. Chem*. 2001; **73**: 29.
13. Wind M, Feldmann I, Jakubowski N, Lehmann WD. *Electrophoresis* 2003; **24**: 1276.
14. Wind M, Gosenca D, Kubler D, Lehmann WD. *Anal. Biochem*. 2003; **317**: 26.
15. Vosseller K, Hansen KC, Chalkley RJ, Trinidad JC, Wells L, Hart GW, Burlingame AL. *Proteomics* 2005; **5**: 388.
16. Weckwerth W, Willmitzer L, Fiehn O. *Rapid Commun. Mass Spectrom*. 2000; **14**: 1677.
17. Ibarrola N, Kalume DE, Gronborg M, Iwahori A, Pandey A. *Anal. Chem*. 2003; **75**: 6043.
18. Bonenfant D, Schmelzle T, Jacinto E, Crespo JL, Mini T, Hall MN, Jenoe P. *Proc. Natl. Acad. Sci. USA* 2003; **100**: 880.

Copyright © 2005 John Wiley & Sons, Ltd.

19. Oda Y, Huang K, Cross FR, Cowburn D, Chait BT. *Proc. Natl. Acad. Sci. USA* 1999; **96**: 6591.
20. Ruse CI, Willard B, Jin JP, Haas T, Kinter M, Bond M. *Anal. Chem*. 2002; **74**: 1658.
21. Tsay YG, Wang YH, Chiu CM, Shen BJ, Lee SC. *Anal. Biochem*. 2000; **287**: 55.
22. Steen H, Jebanathirajah JA, Springer M, Kirschner MW. *Proc. Natl. Acad. Sci. USA* 2005; **102**: 3948.
23. McMichael RW Jr, Klein RR, Salvucci ME, Huber SC. *Arch. Biochem. Biophys*. 1993; **307**: 248.
24. Wessel D, Flugge UI. *Anal. Biochem*. 1984; **138**: 141.
25. Otto A, Thiede B, Muller EC, Scheler C, Wittmann-Liebold B, Jungblut P. *Electrophoresis* 1996; **17**: 1643.
26. Available: http://www.arabidopsis.org/.
27. Glinski M, Weckwerth W. *Mol. Cell. Proteomics* 2005; **4**: 1614.
28. Huber JL, Huber SC. *Biochem. J*. 1992; **283**: 877.
29. Huber SC, Huber JL. *Annu. Rev. Plant Physiol. Plant Mol. Biol*. 1996; **47**: 431.
30. Jones TL, Ort DR. *Plant Physiol*. 1997; **113**: 1167.
31. Winter H, Huber SC. *Crit. Rev. Biochem. Mol. Biol*. 2000; **35**: 253.
32. Wolschin F, Wienkoop S, Weckwerth W. *Proteomics* 2005; in press.
33. Wolschin F, Weckwerth W. *Plant Methods* 2005; in press.
34. Cutillas PR, Geering B, Waterfield MD, *Vanhaesebroeck B. Mol. Cell Proteomics* 2005.

Copyright © 2005 John Wiley & Sons, Ltd.

Concluding Remarks

Broad scale research on protein phosphorylation in plants, often called plant phosphoproteomics is still in its infancy. However, examples exist describing various aspects of the importance of protein phosphorylation as a regulatory mechanism. In nature, in addition to protein phosphorylation other protein modifications contribute to the complex signalling pathways. Thus, phosphorylation is only one, however very important, layer of posttranslational modification influencing interaction and regulation of proteins.

While it is still not as sensitive and selective as ^{32}P labelling, fluorescent dye technology has the potential to replace radioactive methods in the future. Because it is not invasive and applied after protein extraction from their *in vivo* sources it does not alter the phosphorylation state of proteins in a consequence of the experimental setup as might be the case when using ^{32}P. Mass spectrometry based methods are increasingly becoming indispensable for the study of protein phosphorylation because of their sensitivity and selectivity. The neutral loss driven MS^3 approach, which was extensively used in this thesis, proved to be a good means to obtain high quality data on protein phosphorylation, qualitatively as well as quantitatively, but only when considering the MS^2 data. My work clearly demonstrated that stringent filtering criteria are necessary for unambiguous identification of phosphorylation sites with high levels of confidence. Although, this sort of approach results in a low number of identified phosphopeptides, it delivers high quality and reliable data, which can be used as a solid basis for further research on phosphorylation in plants. This is in contrast to some published lists of phosphoproteins for which only moderate criteria were applied. In future, even higher sophisticated instruments will be available and mass spectrometry based research will probably become part of any laboratory working on protein biochemistry. Mutation/knockout analyses will continue to be important when specific functions are to be assigned to site-specific protein phosphorylations.

Further research on phosphorylation is likely to detect important regulatory mechanisms and signalling pathways in plants and by comparison to the better studied animal models it can reveal major similarities and differences between the different kingdoms of life. My own work was restricted to phosphorylation events on serine and threonine since these are regarded as the most important phosphorylation events in plants. However, it should be noted that histidine/aspartate phosphorylation is also reported for plants and is in fact an important event that triggers signal transduction in response to extracellular stimuli including cytokinin

and ethylene signalling [202, 203]. Therefore, it seems likely, that this kind of phosphorylation will be an important research area in the future.

The methods I developed and applied proved to be well suited to investigate phosphorylation on a broad scale in non-targeted approaches (general identification of novel protein phosphorylation sites) but also to conduct detailed hypothesis-driven experiments (as could be shown for PEPC and SPS). MOAC based on aluminum hydroxide is a cost effective and robust alternative for the enrichment of phosphoproteins out of complex protein mixtures. In combination with neutral loss driven approaches it can greatly facilitate the tricky analysis of *in vivo* protein phosphorylation. The quantitation method I developed serves as a robust approach to investigate phosphorylation levels of individual proteins and the ICP-MS approach applied proved to be useful for the investigation of organism or tissue wide phosphorylation levels. Thereby, all important steps of protein phosphorylation analysis were covered and important contributions to most of these steps were made: from the enrichment of phosphoproteins to the detection of protein phosphorylation then to phosphorylation site determination and finally to the quantification of protein phosphorylation.

Appendix: Experimental outline for the analysis of protein phosphorylation

The work described in this thesis resulted in a robust procedure to analyse protein phosphorylation in plants. An outline of the general strategy is given below.

Enrichment of phosphorylated proteins:

The material to be used is frozen in liquid nitrogen and ground in a pre-chilled mortar.

Up to 2 g of ground sample (leaf, whole plant tissue, or cell cultures) or the pellet from about 100 ml of *C. reinhardtii* in the exponential phase are mixed with 5 ml extraction buffer containing freshly add beta-mercaptoethanol (1-5 %) and 15 ml TE (tris edta, pH 7.5 provided by Roth, Germany) buffered phenol (material may be divided into two falcon tubes for convenience). The mixture is incubated for 30 min at 4°C and centrifuged for 8 min at 4000 rpm. The upper phase (containing the proteins) is transferred to a fresh falcon tube and precipitated over night with five volumes ice-cold acetone.

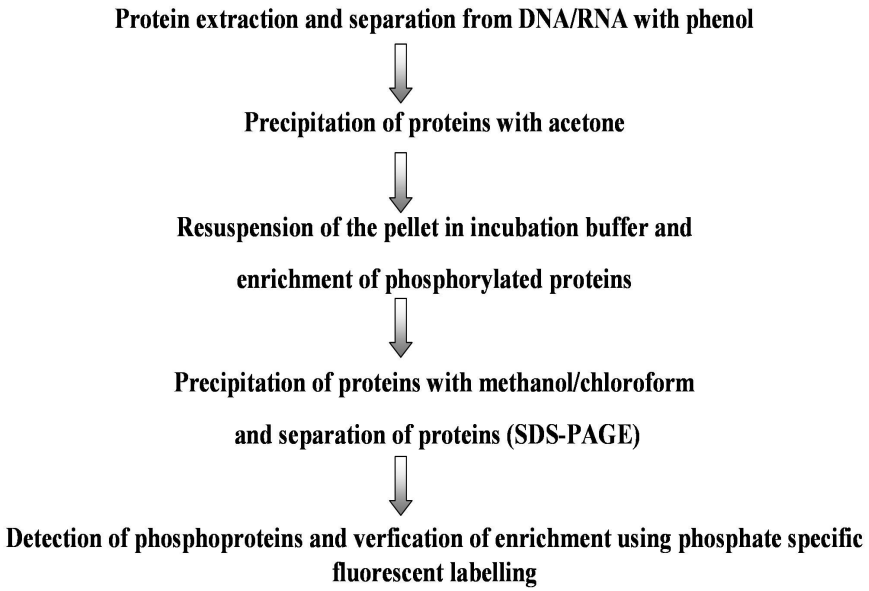


Fig. 15: ***General workflow for the analysis of protein phosphorylation part I.***

On the following day pellets are washed twice with ice-cold methanol (5 ml and 2.5 ml) and dried for 30 min at room temperature. Pellets are then dissolved in 1.8 ml MB (see below) under vigorous vortexing for 5 min. It may be useful to further soften the pellet by using a plastic pistil. Aluminum hydroxide (Sigma A-1577) is preequilibrated with MB by a short washing step (80 mg matrix are mixed with 1.8 ml MB, centrifuged at 10.000 rpm for 2

min and the supernatant is discarded). Up to 1 mg of protein in 1.8 ml of MB is loaded onto 80 mg of $Al(OH)_3$ matrix.

The mixture is rotated at 4°C for 30 min and consecutively centrifuged for 1 min at 10.000 rpm. The supernatant is discarded or saved for the analysis of unbound protein and the matrix is washed five times by adding 1.8 ml MB in each step and centrifuging at 10.000 rpm for one minute). Proteins are eluted from the matrix by adding 800 µl of elution buffer and incubating the mixture for 30 min at room temperature followed by centrifugation at 10.000 rpm for 2 min. The supernatant is precipitated by methanol/chloroform precipitation [204].

Buffer preparation for MOAC (phosphoprotein enrichment):

Extraction buffer:

50 mM Hepes

40 % sucrose

60 mM sodium fluoride

On the day of extraction dissolve buffer in water and adjust pH to 7.5 using KOH.

Dissolvation and binding buffer:

0.2 M sodium glutamate

0.2 M potassium aspartate

30 mM MES buffer

20 mM imidazole

0.25 % Chaps

8 M urea (should be omitted if not necessary for protein solubilisation)

On the day of extraction buffer is dissolved in water and adjust pH to 6.1 using HCl (be careful not to use phosphoric acid for pH adjustment since it interferes with phosphoprotein binding).

Elution buffer:

0.1 M potassium pyrophosphate

8 M urea

On the day of extraction buffer is dissolved in water and pH adjusted to 9.0 using phosphoric acid.

Protein digestion and peptide analysis

Protein digestion

Protein pellets are dissolved in 1 x SDS-PAGE sample buffer under heating and subjected to SDS-PAGE or dissolved in digestion buffer directly. Proteins are then digested using the protocols described in [205] and [33], respectively.

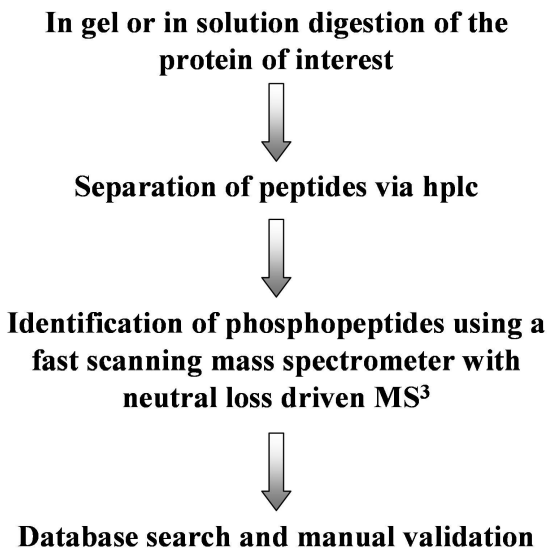


Fig. 16: General workflow for the analysis of protein phosphorylation part II.

Separation of peptides

For the separation of peptides many different methods and columns were used. However, some general recommendations are listed below.

Sample solvation: dissolvation of the peptides in 2 % TFA (trifluoro acetic acid) and 5 % acetonitrile (put 8 µl of this mixture to the peptide pellet, vortex, let the sample stay on ice for 5 min, vortex again, and centrifuge the sample for 2 min at 10000 rpm).

Columns: monolithic columns (50 – 100 µm ID; e.g. Merck; Darmstadt, Germany) were found to deliver good performance. In addition, a special column designed for the investigation of phosphopeptides (500 µm ID; Proteospher, Merck, Darmstadt, Germany) was useful.

Flow rates: 300 – 500 nl/min (monolithic columns), 3 µl/min (Proteospher).

Nano-flow tips for ESI: Coated fused silica tips. ID 30 µm for flow rates starting at 500 nl/min and ID 10 µm for lower flow rates (New objective; Woburn, USA).

Instrument type and instrument settings: several different mass spectrometers were used during the thesis (MALDI-TOF (Applied biosystems), LCQ ion trap, TSQ, LTQ ion trap (all thermo electron).The one that delivered best results and was most widely used was the LTQ linear ion trap. Instrument settings varied depending on the type of method used. Typical values for phosphopeptide analysis are shown below.

Mass window (window in which two precursors are treated as identical): 2 or 3, Normalised collision energy: 35 (measured in arbitrary units); Q value (value with which an ion oscillates in the trap): 0.25; wideband activation: off; activation time: 30 ms; Minimal MS signal: 1000; Minimal MS^n signal: 500; Repeat count (how many times a peak is recorded before the m/z value is put on the dynamic exclusion list): 2; Repeat duration (the amount of time an ion stays on a “pre-exclusion” list): 30 s; Exclusion list size (the maximum number of masses that can be on the dynamic exclusion list): 100; Exclusion duration (the amount of time a mass stays on the exclusion list): 80 s.

Scans: MS scan followed by MS^2 scans for the three most intense ions and MS^3 scans if one of the five most intense fragment ions in a MS^2 scan can be explained by a loss of 32.7, 49, or 98 Da from the precursor ion.

Useful links:

Database searching:

Sequest: commercial database searching program with many features. Together with mascot the most widely used program. Should be used in combination with DTASelect

(http://fields.scripps.edu/sequest/).

Mascot: one of the most extensively used database search programs with a limited version being publicly available (http://www.matrixscience.com/).

OMSSA: a very fast and useful open source program with limited graphical user interface

(http://pubchem.ncbi.nlm.nih.gov/omssa/).

GPM (global proteome machine): a powerful database searching program with an additional spectrum library. Nice graphical user interface, limitations in the search for modifications (two modifications can not be set as variable for the same residue in the initial search. Might be changed in future releases)(http://www.thegpm.org/).

Protein phosphorylation site databases:

PlantsP: Plant protein phosphorylation database (http://plantsp.genomics.purdue.edu/).

Protein phosphorylation site database: Information on phosphoproteins in prokaryotic organisms (http://vigen.biochem.vt.edu/xpd/xpd.htm).

Prokaryotic Protein Phosphatase Database: Information on protein phosphatases from prokaryotic organisms (http://vigen.biochem.vt.edu/p3d/p3d.htm).

PhosphoSite: Data repository containing protein phosphorylation sites of mouse and human (http://www.phosphosite.org/Login.jsp).

Phospho.ELM: contains a collection of experimentally verified Serine, Threonine and Tyrosine sites in eukaryotic proteins (http://phospho.elm.eu.org/).

Protein Kinase Resource: General data repository for protein kinases

(http://www.kinasenet.org/pkr/).

IV.3. Literature

1. B. L. Gutman and K. K. Niyogi *Chlamydomonas* and *Arabidopsis*. A dynamic duo. *Plant Physiol* 135, 607-10 (2004).

2. '*Arabidopsis* Genome Initiative' Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796-815 (2000).

3. E. H. Harris *Chlamydomonas* as a Model Organism. *Annu Rev Plant Physiol Plant Mol Biol* 52, 363-406 (2001).

4. A. E. Hall, A. Fiebig and D. Preuss Beyond the *Arabidopsis* genome: opportunities for comparative genomics. *Plant Physiol* 129, 1439-47 (2002).

5. H. B. Vickery The origin of the word protein. *Yale J Biol Med* 22, 387-93 (1950).

6. T. Schwend, I. Redwanz, T. Ruppert, A. Szenthe and M. Wink Analysis of proteins in the spent culture medium of Lupinus albus by electrospray ionisation tandem mass spectrometry. *J Chromatogr A* 1009, 105-10 (2003).

7. W. D. Lehmann Massenspektrometrie in der Biochemie, Vol., Edn. (Spektrum akademischer Verlag, Heidelberg; 1996).

8. C. Meng, M. Mann and J. Fenn Electrospray ionization of some polypeptides and small proteins. *Proceedings of the 36th American Society for Mass Spectrometry Conference on Mass Spectrometry and Allied Topics*, 711-772 (1988).

9. P. Kebarle A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *J. Mass Spectrom.* 35, 804-817 (2000).

10. J. F. Nemeth-Cawley, B. S. Tangarone and J. C. Rouse "Top Down" characterization is a complementary technique to peptide sequencing for identifying protein species in complex mixtures. *J Proteome Res* 2, 495-505 (2003).

11. V. Zabrouskov, L. Giacomelli, K. J. van Wijk and F. W. McLafferty A new approach for plant proteomics: characterization of chloroplast proteins of *Arabidopsis thaliana* by top-down mass spectrometry. *Mol Cell Proteomics* 2, 1253-60 (2003).

12. N. Kelleher, H. Lin, G. Valaskovic, D. Aaserud, E. Fridriksson and F. McLafferty Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. *J. Am. Chem. Soc.* 121, 806-812 (1999).

13. A. R. Dongre, J. K. Eng and J. R. Yates, 3rd Emerging tandem-mass-spectrometry techniques for the rapid identification of proteins. *Trends Biotechnol* 15, 418-25 (1997).

14. R. Craig, J. P. Cortens and R. C. Beavis The use of proteotypic peptide libraries for protein identification. *Rapid Commun Mass Spectrom* 19, 1844-50 (2005).

15. R. Craig, J. P. Cortens and R. C. Beavis Open source system for analyzing, validating, and storing protein identification data. *J Proteome Res* 3, 1234-42 (2004).

16. L. Y. Geer, S. P. Markey, J. A. Kowalak, L. Wagner, M. Xu, D. M. Maynard, X. Yang, W. Shi and S. H. Bryant Open mass spectrometry search algorithm. *J Proteome Res* 3, 958-64 (2004).

17. J. R. Yates, 3rd, J. K. Eng, A. L. McCormack and D. Schieltz Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal Chem* 67, 1426-36 (1995).

18. B. Cox, T. Kislinger and A. Emili Integrating gene and protein expression data: pattern analysis and profile mining. *Methods* 35, 303-14 (2005).

19. H. Liu, R. G. Sadygov and J. R. Yates, 3rd A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* 76, 4193-201 (2004).

20. K. B. Rechinger, H. Siegumfeldt, I. Svendsen and M. Jakobsen "Early" protein synthesis of Lactobacillus delbrueckii ssp. bulgaricus in milk revealed by [35S] methionine labeling and two-dimensional gel electrophoresis. *Electrophoresis* 21, 2660-9 (2000).

21. Y. Oda, K. Huang, F. R. Cross, D. Cowburn and B. T. Chait Accurate quantitation of protein expression and site-specific phosphorylation. *Proc Natl Acad Sci U S A* 96, 6591-6 (1999).

22. S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey and M. Mann Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1, 376-86 (2002).

23. S. E. Ong, I. Kratchmarova and M. Mann Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J Proteome Res* 2, 173-81 (2003).

24. S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb and R. Aebersold Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 17, 994-9 (1999).

25. A. Schmidt, J. Kellermann and F. Lottspeich A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* 5, 4-15 (2005).

26. K. S. Lilley and P. Dupree Methods of quantitative proteomics and their application to plant organelle characterization. *J Exp Bot* 57, 1493-9 (2006).

27. J. Klose Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* 26, 231-43 (1975).

28. Stewart, II, T. Thomson and D. Figeys 18O labeling: a tool for proteomics. *Rapid Commun Mass Spectrom* 15, 2456-65 (2001).

29. P. L. Ross, Y. N. Huang, J. N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson and D. J. Pappin Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 3, 1154-69 (2004).

30. D. M. Desiderio and M. Kai Preparation of stable isotope-incorporated peptide internal standards for field desorption mass spectrometry quantification of peptides in biologic tissue. *Biomed Mass Spectrom* 10, 471-9 (1983).

31. J. C. Silva, R. Denny, C. A. Dorschel, M. Gorenstein, I. J. Kass, G. Z. Li, T. McKenna, M. J. Nold, K. Richardson, P. Young and S. Geromanos Quantitative proteomic analysis by accurate mass retention time pairs. *Anal Chem* 77, 2187-200 (2005).

32. G. Wang, W. W. Wu, W. Zeng, C. L. Chou and R. F. Shen Label-free protein quantification using LC-coupled ion trap or FT mass spectrometry: Reproducibility, linearity, and application with complex proteomes. *J Proteome Res* 5, 1214-23 (2006).

33. K. Morgenthal, S. Wienkoop, F. Wolschin and W. Weckwerth in Metabolomics: Methods and Protocols, Vol., Edn. (ed.^(eds. 57-75 (Humana Press, Totowa, NJ; 2006).

34. W. Wang, H. Zhou, H. Lin, S. Roy, T. A. Shaler, L. R. Hill, S. Norton, P. Kumar, M. Anderle and C. H. Becker Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal Chem* 75, 4818-26 (2003).

35. D. Radulovic, S. Jelveh, S. Ryu, T. G. Hamilton, E. Foss, Y. Mao and A. Emili Informatics platform for global proteomic profiling and biomarker discovery using liquid chromatography-tandem mass spectrometry. *Mol Cell Proteomics* 3, 984-97 (2004).

36. M. Glinski and W. Weckwerth The role of mass spectrometry in plant systems biology. *Mass Spectrom Rev* 25, 173-214 (2006).

37. O. N. Jensen Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr Opin Chem Biol* 8, 33-41 (2004).

38. W. D. Lehmann in Handbook of Toxicogenomics, Vol., Edn. 1. (ed.^(eds. J. Borlack) 115-136 (Wiley-VCH, 2005).

39. S. J. Kwon, E. Y. Choi, Y. J. Choi, J. H. Ahn and O. K. Park Proteomics studies of post-translational modifications in plants. *J Exp Bot* 57, 1547-51 (2006).

40. L. R. Zeng, M. E. Vega-Sanchez, T. Zhu and G. L. Wang Ubiquitination-mediated protein degradation and modification: an emerging theme in plant-microbe interactions. *Cell Res* 16, 413-26 (2006).

41. S. Kepinski and O. Leyser Ubiquitination and auxin signaling: a degrading story. *Plant Cell* 14 Suppl, S81-95 (2002).

42. B. Eisenhaber, M. Wildpaner, C. J. Schultz, G. H. Borner, P. Dupree and F. Eisenhaber Glycosylphosphatidylinositol lipid anchoring of plant proteins. Sensitive prediction from sequence- and genome-wide studies for Arabidopsis and rice. *Plant Physiol* 133, 1691-701 (2003).

43. M. J. Omaetxebarria, P. Hagglund, F. Elortza, N. M. Hooper, J. M. Arizmendi and O. N. Jensen Isolation and Characterization of Glycosylphosphatidylinositol-Anchored Peptides by Hydrophilic Interaction Chromatography and MALDI Tandem Mass Spectrometry. *Anal Chem* 78, 3335-3341 (2006).

44. C. Lindermayr, G. Saalbach and J. Durner Proteomic identification of S-nitrosylated proteins in Arabidopsis. *Plant Physiol* 137, 921-30 (2005).

45. D. P. Dixon, M. Skipsey, N. M. Grundy and R. Edwards Stress-induced protein S-glutathionylation in Arabidopsis. *Plant Physiol* 138, 2233-44 (2005).

46. L. Johnson, S. Mollah, B. A. Garcia, T. L. Muratore, J. Shabanowitz, D. F. Hunt and S. E. Jacobsen Mass spectrometry analysis of Arabidopsis histone H3 reveals distinct combinations of post-translational modifications. *Nucleic Acids Res* 32, 6511-8 (2004).

47. R. Grimm, M. Grimm, C. Eckerskorn, K. Pohlmeyer, T. Rohl and J. Soll Postimport methylation of the small subunit of ribulose-1,5-bisphosphate carboxylase in chloroplasts. *FEBS Lett* 408, 350-4 (1997).

48. http://nobelprize.org/medicine/laureates/1992/press.html.

49. P. Cohen The regulation of protein function by multisite phosphorylation--a 25 year update. *Trends Biochem Sci* 25, 596-601 (2000).

50. P. J. Roach Multisite and hierarchal protein phosphorylation. *J Biol Chem* 266, 14139-42 (1991).

51. H. Saito Histidine phosphorylation and two-component signaling in eukaryotic cells. *Chem Rev* 101, 2497-509 (2001).

52. T. Hunter and J. A. Cooper Protein-tyrosine kinases. *Annu Rev Biochem* 54, 897-930 (1985).

53. A. Sickmann, M. Mreyen and H. E. Meyer Identification of modified proteins by mass spectrometry. *IUBMB Life* 54, 51-7 (2002).

54. H. R. Matthews Protein kinases and phosphatases that act on histidine, lysine, or arginine residues in eukaryotic proteins: a possible regulator of the mitogen-activated protein kinase cascade. *Pharmacol Ther* 67, 323-50 (1995).

55. T. Hunter The Croonian Lecture 1997. The phosphorylation of proteins on tyrosine: its role in cell growth and disease. *Philos Trans R Soc Lond B Biol Sci* 353, 583-605 (1998).

56. S. Klumpp and J. Krieglstein Phosphorylation and dephosphorylation of histidine residues in proteins. *Eur J Biochem* 269, 1067-71 (2002).

57. R. W. McMichael, Jr. and J. C. Lagarias Phosphopeptide mapping of Avena phytochrome phosphorylated by protein kinases in vitro. *Biochemistry-Us* 29, 3872-8 (1990).

58. V. N. Lapko, X. Y. Jiang, D. L. Smith and P. S. Song Mass spectrometric characterization of oat phytochrome A: isoforms and posttranslational modifications. *Protein Sci* 8, 1032-44 (1999).

59. V. N. Lapko, X. Y. Jiang, D. L. Smith and P. S. Song Posttranslational modification of oat phytochrome A: phosphorylation of a specific serine in a multiple serine cluster. *Biochemistry-Us* 36, 10595-9 (1997).

60. V. N. Lapko, T. A. Wells and P. S. Song Protein kinase A-catalyzed phosphorylation and its effect on conformation in phytochrome A. *Biochemistry-Us* 35, 6585-94 (1996).

61. V. Rubio and X. W. Deng Phy tunes: phosphorylation status and phytochrome-mediated signaling. *Cell* 120, 290-2 (2005).

62. J. I. Kim, Y. Shen, Y. J. Han, J. E. Park, D. Kirchenbauer, M. S. Soh, F. Nagy, E. Schafer and P. S. Song Phytochrome phosphorylation modulates light signaling by influencing the protein-protein interaction. *Plant Cell* 16, 2629-40 (2004).

63. X. Wang, M. B. Goshe, E. J. Soderblom, B. S. Phinney, J. A. Kuchar, J. Li, T. Asami, S. Yoshida, S. C. Huber and S. D. Clouse Identification and functional analysis of in vivo phosphorylation sites of the Arabidopsis BRASSINOSTEROID-INSENSITIVE1 receptor kinase. *Plant Cell* 17, 1685-703 (2005).

64. S. Yoshida and M. Parniske Regulation of plant symbiosis receptor kinase through serine and threonine phosphorylation. *J Biol Chem* 280, 9203-9 (2005).

65. S. Stracke, C. Kistner, S. Yoshida, L. Mulder, S. Sato, T. Kaneko, S. Tabata, N. Sandal, J. Stougaard, K. Szczyglowski and M. Parniske A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* 417, 959-62 (2002).

66. I. Carlberg, M. Hansson, T. Kieselbach, W. P. Schroder, B. Andersson and A. V. Vener A novel plant protein undergoing light-induced phosphorylation and release from the photosynthetic thylakoid membranes. *Proc Natl Acad Sci U S A* 100, 757-62 (2003).

67. A. V. Vener, A. Harms, M. R. Sussman and R. D. Vierstra Mass spectrometric resolution of reversible protein phosphorylation in photosynthetic membranes of *Arabidopsis thaliana*. *J Biol Chem* 276, 6959-66 (2001).

68. E. M. Aro, A. Rokka and A. V. Vener Determination of phosphoproteins in higher plant thylakoids. *Methods Mol Biol* 274, 271-85 (2004).

69. A. V. Vener, I. Ohad and B. Andersson Protein phosphorylation and redox sensing in chloroplast thylakoids. *Curr Opin Plant Biol* 1, 217-23 (1998).

70. V. Citovsky, B. G. McLean, J. R. Zupan and P. Zambryski Phosphorylation of tobacco mosaic virus cell-to-cell movement protein by a developmentally regulated plant cell wall-associated protein kinase. *Genes Dev* 7, 904-10 (1993).

71. E. Kolobova, A. Tuganova, I. Boulatnikov and K. M. Popov Regulation of pyruvate dehydrogenase activity through phosphorylation at multiple sites. *Biochem J* 358, 69-77 (2001).

72. J. Avruch, Y. Lin, X. Long, S. Murthy and S. Ortiz-Vega Recent advances in the regulation of the TOR pathway by insulin and nutrients. *Curr Opin Clin Nutr Metab Care* 8, 67-72 (2005).

73. A. J. Williams, J. Werner-Fraczek, I. F. Chang and J. Bailey-Serres Regulated phosphorylation of 40S ribosomal protein S6 in root tips of maize. *Plant Physiol* 132, 2086-97 (2003).

74. I. F. Chang, K. Szick-Miranda, S. Pan and J. Bailey-Serres Proteomic characterization of evolutionarily conserved and variable proteins of Arabidopsis cytosolic ribosomes. *Plant Physiol* 137, 848-62 (2005).

75. H. Reyes de la Cruz, R. Aguilar and E. Sanchez de Jimenez Functional characterization of a maize ribosomal S6 protein kinase (ZmS6K), a plant ortholog of metazoan p70(S6K). *Biochemistry-Us* 43, 533-9 (2004).

76. C. Garcia Flores, R. Aguilar, H. Reyes de la Cruz, M. Albores and E. Sanchez de Jimenez A maize insulin-like growth factor signals to a transduction pathway that regulates protein synthesis in maize. *Biochem J* 358, 95-100 (2001).

77. E. Beltran-Pena, R. Aguilar, A. Ortiz-Lopez, T. D. Dinkova and E. S. De Jimenez Auxin stimulates S6 ribosomal protein phosphorylation in maize thereby affecting protein synthesis regulation. *Physiol Plant* 115, 291-297 (2002).

78. E. Sanchez-de-Jimenez, R. Aguilar and T. Dinkova S6 ribosomal protein phosphorylation and translation of stored mRNA in maize. *Biochimie* 79, 187-94 (1997).

79. M. M. Mahfouz, S. Kim, A. J. Delauney and D. P. Verma Arabidopsis TARGET OF RAPAMYCIN Interacts with RAPTOR, Which Regulates the Activity of S6 Kinase in Response to Osmotic Stress Signals. *Plant Cell*, (2005).

80. A. Tovar-Mendez, J. A. Miernyk and D. D. Randall Regulation of pyruvate dehydrogenase complex activity in plant cells. *Eur J Biochem* 270, 1043-9 (2003).

81. K. P. Rao and D. D. Randall Plant pyruvate dehydrogenase complex: inactivation and reactivation by phosphorylation and dephosphorylation. *Arch Biochem Biophys* 200, 461-6 (1980).

82. N. V. Bykova, A. Stensballe, H. Egsgaard, O. N. Jensen and I. M. Moller Phosphorylation of formate dehydrogenase in potato tuber mitochondria. *J Biol Chem* 278, 26021-30 (2003).

83. K. Izui, H. Matsumura, T. Furumoto and Y. Kai Phosphoenolpyruvate carboxylase: a new era of structural biology. *Annu Rev Plant Biol* 55, 69-84 (2004).

84. B. Li, X. Zhang and R. Chollet Phosphoenolpyruvate carboxylase kinase in tobacco leaves is activated by light in a similar but not identical way as in maize. *PLANT PHYSIOLOGY* 111, 497-505 (1996).

85. S. Duff and R. Chollet In Vivo Regulation of Wheat-Leaf Phosphoenolpyruvate Carboxylase by Reversible Phosphorylation. *Plant Physiol* 107, 775-782 (1995).

86. X. Q. Zhang, B. Li and R. Chollet In Vivo Regulatory Phosphorylation of Soybean Nodule Phosphoenolpyruvate Carboxylase. *Plant Physiol* 108, 1561-1568 (1995).

87. A. Gousset-Dupont, B. Lebouteiller, J. Monreal, C. Echevarria, J. Pierre, M. Hodges and J. Vidal Metabolite and post-translational control of phosphoenolpyruvate carboxylase from leaves and mesophyll cell protoplasts of *Arabidopsis thaliana*. *Plant Sci.* 169, 1096-1101 (2005).

88. H. Winter and S. C. Huber Regulation of sucrose metabolism in higher plants: localization and regulation of activity of key enzymes. *Crit Rev Biochem Mol Biol* 35, 253-89 (2000).

89. S. C. Huber and J. L. Huber Role and Regulation of Sucrose-Phosphate Synthase in Higher Plants. *Annu Rev Plant Physiol Plant Mol Biol* 47, 431-444 (1996).

90. R. W. McMichael, Jr., R. R. Klein, M. E. Salvucci and S. C. Huber Identification of the major regulatory phosphorylation site in sucrose-phosphate synthase. *Arch Biochem Biophys* 307, 248-52 (1993).

91. D. Toroser, G. S. Athwal and S. C. Huber Site-specific regulatory interaction between spinach leaf sucrose-phosphate synthase and 14-3-3 proteins. *FEBS Lett* 435, 110-4 (1998).

92. B. Alberts, D. Bray, K. Lewis, M. Raff and e. al. Molecular Biology of the Cell, 3rd ed. *Garland Publishing, New York*, 195-222 (1994).

93. M. Mann, S. E. Ong, M. Gronborg, H. Steen, O. N. Jensen and A. Pandey Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. *Trends Biotechnol* 20, 261-8 (2002).

94. H. Steen, J. A. Jebanathirajah, J. Rush, N. Morrice and M. W. Kirschner Phosphorylation Analysis by Mass Spectrometry: Myths, Facts, and the Consequences for Qualitative and Quantitative Measurements. *Mol Cell Proteomics* 5, 172-181 (2006).

95. G. K. Taylor and D. R. Goodlett Rules governing protein identification by mass spectrometry. *Rapid Commun Mass Spectrom* 19, 3420 (2005).

96. E. Salih Phosphoproteomics by mass spectrometry and classical protein chemistry approaches. *Mass Spectrom Rev* 24, 828-46 (2005).

97. S. Laugesen, A. Bergoin and M. Rossignol Deciphering the plant phosphoproteome: tools and strategies for a challenging task. *Plant Physiol Biochem* 42, 929-36 (2004).

98. G. V. Novikova, I. E. Moshkov, A. R. Smith and M. A. Hall The effect of ethylene on MAPKinase-like activity in *Arabidopsis thaliana*. *FEBS Lett* 474, 29-32 (2000).

99. A. C. Roque and C. R. Lowe Lessons from nature: On the molecular recognition elements of the phosphoprotein binding-domains. *Biotechnol Bioeng* 91, 546-55 (2005).

100. F. C. Milne, G. Moorhead, M. Pozuelo Rubio, B. Wong, A. Kulma, J. E. Harthill, D. Villadsen, V. Cotelle and C. MacKintosh Affinity purification of diverse plant and human 14-3-3-binding partners. *Biochem Soc Trans* 30, 379-81 (2002).

101. A. Kulma, D. Villadsen, D. G. Campbell, S. E. Meek, J. E. Harthill, T. H. Nielsen and C. MacKintosh Phosphorylation and 14-3-3 binding of Arabidopsis 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. *Plant J* 37, 654-67 (2004).

102. F. Thaler, B. Valsasina, R. Baldi, J. Xie, A. Stewart, A. Isacchi, H. M. Kalisz and L. Rusconi A new approach to phosphoserine and phosphothreonine analysis in peptides and proteins: chemical modification, enrichment via solid-phase reversible binding, and analysis by mass spectrometry. *Anal Bioanal Chem* 376, 366-73 (2003).

103. W. J. Qian, M. B. Goshe, D. G. Camp, 2nd, L. R. Yu, K. Tang and R. D. Smith Phosphoprotein isotope-coded solid-phase tag approach for enrichment and quantitative analysis of phosphopeptides from complex mixtures. *Anal Chem* 75, 5441-50 (2003).

104. K. Vosseller, K. C. Hansen, R. J. Chalkley, J. C. Trinidad, L. Wells, G. W. Hart and A. L. Burlingame Quantitative analysis of both protein expression and serine / threonine post-translational modifications through stable isotope labeling with dithiothreitol. *Proteomics* 5, 388-98 (2005).

105. P. van der Veken, E. H. Dirksen, E. Ruijter, R. C. Elgersma, A. J. Heck, D. T. Rijkers, M. Slijper and R. M. Liskamp Development of a Novel Chemical Probe for the Selective Enrichment of Phosphorylated Serine- and Threonine-Containing Peptides. *Chembiochem* 6, 2271-2280 (2005).

106. M. B. Goshe, T. P. Conrads, E. A. Panisko, N. H. Angell, T. D. Veenstra and R. D. Smith Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses. *Anal Chem* 73, 2578-86 (2001).

107. M. B. Goshe, T. D. Veenstra, E. A. Panisko, T. P. Conrads, N. H. Angell and R. D. Smith Phosphoprotein isotope-coded affinity tags: application to the enrichment and identification of low-abundance phosphoproteins. *Anal Chem* 74, 607-16 (2002).

108. Y. Oda, T. Nagasu and B. T. Chait Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat Biotechnol* 19, 379-82 (2001).

109. H. Zhou, J. D. Watts and R. Aebersold A systematic approach to the analysis of protein phosphorylation. *Nat Biotechnol* 19, 375-8 (2001).

110. W. A. Tao, B. Wollscheid, R. O'Brien, J. K. Eng, X. J. Li, B. Bodenmiller, J. D. Watts, L. Hood and R. Aebersold Quantitative phosphoproteome analysis using a dendrimer conjugation chemistry and tandem mass spectrometry. *Nat Methods* 2, 591-8 (2005).

111. A. Amoresano, G. Marino, C. Cirulli and E. Quemeneur Mapping phosphorylation sites: a new strategy based on the use of isotopically-labelled dithiothreitol and mass spectrometry. *Eur J Mass Spectrom* 10, 401-412 (2004).

112. L. Wells, K. Vosseller, R. N. Cole, J. M. Cronshaw, M. J. Matunis and G. W. Hart Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. *Mol. Cell. Proteomics* 1, 791-804 (2002).

113. K. F. Medzihradszky, Z. Darula, E. Perlson, M. Fainzilber, R. J. Chalkley, H. Ball, D. Greenbaum, M. Bogyo, D. R. Tyson, R. A. Bradshaw and A. L. Burlingame O-sulfonation of serine and threonine: mass spectrometric detection and characterization of a new posttranslational modification in diverse proteins throughout the eukaryotes. *Mol Cell Proteomics* 3, 429-40 (2004).

114. D. T. McLachlin and B. T. Chait Improved beta-elimination-based affinity purification strategy for enrichment of phosphopeptides. *Anal Chem* 75, 6826-36 (2003).

115. J. R. Whitaker and R. E. Feeney Chemical and physical modification of proteins by the hydroxide ion. *Crit Rev Food Sci Nutr* 19, 173-212 (1983).

116. J. Porath, J. Carlsson, I. Olsson and G. Belfrage Metal Chelate Affinity Chromatography, a New Approach to Protein Fractionation. *Nature* 258, 598-599 (1975).

117. R. Anguenot, S. Yelle and B. Nguyen-Quoc Purification of tomato sucrose synthase phosphorylated isoforms by Fe(III)-immobilized metal affinity chromatography. *Arch Biochem Biophys* 365, 163-9 (1999).

118. D. Heintz, V. Wurtz, A. A. High, A. Van Dorsselaer, R. Reski and E. Sarnighausen An efficient protocol for the identification of protein phosphorylation in a seedless plant, sensitive enough to detect members of signalling cascades. *Electrophoresis* 25, 1149-1159 (2004).

119. T. S. Nuhse, A. Stensballe, O. N. Jensen and S. C. Peck Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry. *Mol. Cell. Proteomics* 2, 1234-1243 (2003).

120. M. Zachariou and M. T. W. Hearn Protein Selectivity in Immobilized Metal Affinity-Chromatography Based on the Surface Accessibility of Aspartic and Glutamic-Acid Residues. *J. Protein Chem.* 14, 419-430 (1995).

121. E. K. Ueda, P. W. Gout and L. Morganti Current and prospective applications of metal ion-protein binding. *J Chromatogr A* 988, 1-23 (2003).

122. T. He, K. Alving, B. Feild, J. Norton, E. G. Joseloff, S. D. Patterson and B. Domon Quantitation of phosphopeptides using affinity chromatography and stable isotope labeling. *J Am Soc Mass Spectr* 15, 363-373 (2004).

123. C. E. Haydon, P. A. Eyers, L. D. Aveline-Wolf, K. A. Resing, J. L. Maller and N. G. Ahn Identification of Novel Phosphorylation Sites on *Xenopus laevis* Aurora A and Analysis of Phosphopeptide Enrichment by Immobilized Metal-affinity Chromatography. *Mol Cell Proteomics* 2, 1055-1067 (2003).

124. L. Andersson and J. Porath Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. *Anal Biochem* 154, 250-4 (1986).

125. M. O. Collins, L. Yu, H. Husi, W. P. Blackstock, J. S. Choudhary and S. G. Grant Robust enrichment of phosphorylated species in complex mixtures by sequential protein and peptide metal-affinity chromatography and analysis by tandem mass spectrometry. *Sci STKE* 2005, pl6 (2005).

126. M. O. Collins, L. Yu, M. P. Coba, H. Husi, I. Campuzano, W. P. Blackstock, J. S. Choudhary and S. G. Grant Proteomic analysis of in vivo phosphorylated synaptic proteins. *J Biol Chem* 280, 5972-82 (2005).

127. S. Y. Huang, J. L. Hsu and S. H. Chen Two-step immobilized metal affinity chromatography (IMAC) for phosphoproteomics using mass spectrometry. *J Chin Chem Soc-Taip* 52, 765-772 (2005).

128. E. Kinoshita, A. Yamada, H. Takeda, E. Kinoshita-Kikuta and T. Koike Novel immobilized zinc(II) affinity chromatography for phosphopeptides and phosphorylated proteins. *J Sep Sci* 28, 155-62 (2005).

129. H. Strunz and E. Nickel Strunz Mineralogical Tables.
Chemical-Structural Mineral Classification System., Vol., Edn. 9th. (E. Schweizerbart`sche Verlagsbuchhandlung, Stuttgart; 2001).

130. C. T. Chen and Y. C. Chen Fe3O4/TiO2 core/shell nanoparticles as affinity probes for the analysis of phosphopeptides using TiO2 surface-assisted laser desorption/ionization mass spectrometry. *Anal Chem* 77, 5912-9 (2005).

131. I. Kuroda, Y. Shintani, M. Motokawa, S. Abe and M. Furuno Phosphopeptide-selective column-switching RP-HPLC with a titania precolumn. *Anal Sci* 20, 1313-9 (2004).

132. M. R. Larsen, T. E. Thingholm, O. N. Jensen, P. Roepstorff and T. J. Jorgensen Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics* 4, 873-86 (2005).

133. A. Sano and H. Nakamura Titania as a chemo-affinity support for the column-switching HPLC analysis of phosphopeptides: Application to the characterization of phosphorylation sites in proteins by combination with protease digestion and electrospray ionization mass Spectrometry. *Anal Sci* 20, 861-864 (2004).

134. A. Sano and H. Nakamura Chemo-affinity of titania for the column-switching HPLC analysis of phosphopeptides. *Anal Sci* 20, 565-566 (2004).

135. M. W. H. Pinkse, P. M. Uitto, M. J. Hilhorst, B. Ooms and A. J. R. Heck Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-nanoLC-ESI-MS/MS and titanium oxide precolumns. *Anal. Chem.* 76, 3935-3943 (2004).

136. S. Miyazaki, K. Morisato, N. Ishizuka, H. Minakuchi, Y. Shintani, M. Furuno and K. Nakanishi Development of a monolithic silica extraction tip for the analysis of proteins. *J Chromatogr A* 1043, 19-25 (2004).

137. Y. Ikeguchi and H. Nakamura Determination of organic phosphates by column-switching high performance anion-exchange chromatography using on-line preconcentration on titania. *Anal Sci* 13, 479-483 (1997).

138. K. Murayama, H. Nakamura, T. Nakajima, K. Takahashi and A. Yoshida Preparation and evaluation of octadecyl titania as column-packing material for high-performance liquid-chromatography. *Microchem J* 49, 362-367 (1994).

139. H. Matsuda, H. Nakamura and T. Nakajima New ceramic-titania - selective adsorbent for organic-phosphates. *Anal Sci* 6, 911-912 (1990).

140. H. K. Kweon and K. Hakansson Selective zirconium dioxide-based enrichment of phosphorylated peptides for mass spectrometric analysis. *Anal Chem* 78, 1743-9 (2006).

141. S. Rinalducci, M. R. Larsen, S. Mohammed and L. Zolla Novel protein phosphorylation site identification in spinach stroma membranes by titanium dioxide microcolumns and tandem mass spectrometry. *J Proteome Res* 5, 973-82 (2006).

142. N. W. Baylor, W. Egan and P. Richman Aluminum salts in vaccines--US perspective. *Vaccine* 20 Suppl 3, S18-23 (2002).

143. J. N. Francis and S. R. Durham Adjuvants for allergen immunotherapy: experimental results and clinical perspectives. *Curr Opin Allergy Clin Immunol* 4, 543-8 (2004).

144. S. Iyer, H. HogenEsch and S. L. Hem Effect of the degree of phosphate substitution in aluminum hydroxide adjuvant on the adsorption of phosphorylated proteins. *Pharm Dev Technol* 8, 81-6 (2003).

145. M. F. Chang, J. L. White, S. L. Nail and S. L. Hem Role of the electrostatic attractive force in the adsorption of proteins by aluminum hydroxide adjuvant. *PDA J Pharm Sci Technol* 51, 25-9 (1997).

146. Y. Shi, H. HogenEsch and S. L. Hem Change in the degree of adsorption of proteins by aluminum-containing adjuvants following exposure to interstitial fluid: freshly prepared and aged model vaccines. *Vaccine* 20, 80-5 (2001).

147. H. Kato, M. Shibano, T. Saito, J. Yamaguchi, S. Yoshihara and N. Goto Relationship between hemolytic activity and adsorption capacity of aluminum hydroxide and calcium phosphate as immunological adjuvants for biologicals. *Microbiol Immunol* 38, 543-8 (1994).

148. M. A. Colettipreviero and A. Previero Alumina Phosphate Complexes for Immobilization of Biomolecules. *Anal. Biochem.* 180, 1-10 (1989).

149. A. Tullberg, G. Hakansson and H. L. Race A protein tyrosine kinase of chloroplast thylakoid membranes phosphorylates light harvesting complex II proteins. *Biochem Biophys Res Commun* 250, 617-22 (1998).

150. A. K. Azad, Y. Sawa, T. Ishikawa and H. Shibata Phosphorylation of plasma membrane aquaporin regulates temperature-dependent opening of tulip petals. *Plant Cell Physiol* 45, 608-17 (2004).

151. S. Pursiheimo, E. Rintamaki, E. Baena-Gonzalez and E. M. Aro Thylakoid protein phosphorylation in evolutionally divergent species with oxygenic photosynthesis. *FEBS Lett* 423, 178-82 (1998).

152. E. Barizza, F. Lo Schiavo, M. Terzi and F. Filippini Evidence suggesting protein tyrosine phosphorylation in plants depends on the developmental conditions. *FEBS Lett* 447, 191-4 (1999).

153. M. Glinski, T. Romeis, C. P. Witte, S. Wienkoop and W. Weckwerth Stable isotope labeling of phosphopeptides for multiparallel kinase target analysis and identification of phosphorylation sites. *Rapid Commun Mass Spectrom* 17, 1579-84 (2003).

154. J. A. Bond, K. Webley, F. S. Wyllie, C. J. Jones, A. Craig, T. Hupp and D. Wynford-Thomas p53-Dependent growth arrest and altered p53-immunoreactivity following metabolic labelling with 32P ortho-phosphate in human fibroblasts. *Oncogene* 18, 3788-92 (1999).

155. A. L. Craig, J. P. Blaydes, L. R. Burch, A. M. Thompson and T. R. Hupp Dephosphorylation of p53 at Ser20 after cellular exposure to low levels of non-ionizing radiation. *Oncogene* 18, 6305-12 (1999).

156. D. Heintz, V. Wurtz, A. A. High, A. Van Dorsselaer, R. Reski and E. Sarnighausen An efficient protocol for the identification of protein phosphorylation in a seedless plant, sensitive enough to detect members of signalling cascades. *Electrophoresis* 25, 1149-59 (2004).

157. T. Furumoto, M. Teramoto, N. Inada, M. Ito, I. Nishida and A. Watanabe Phosphorylation of a bifunctional enzyme, 6-phosphofructo-2-kinase/fructose-2,6-bisphosphate 2-phosphatase, is regulated physiologically and developmentally in rosette leaves of *Arabidopsis thaliana*. *Plant Cell Physiol* 42, 1044-8 (2001).

158. M. W. Pierce, J. L. Palmer, H. T. Keutmann, T. A. Hall and J. Avruch The insulin-directed phosphorylation site on ATP-citrate lyase is identical with the site phosphorylated by the cAMP-dependent protein kinase in vitro. *J Biol Chem* 257, 10681-6 (1982).

159. N. Blom, T. Sicheritz-Ponten, R. Gupta, S. Gammeltoft and S. Brunak Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4, 1633-49 (2004).

160. D. Toroser, R. McMichael, Jr., K. P. Krause, J. Kurreck, U. Sonnewald, M. Stitt and S. C. Huber Site-directed mutagenesis of serine 158 demonstrates its role in spinach leaf sucrose-phosphate synthase modulation. *Plant J* 17, 407-13 (1999).

161. E. Kinoshita, E. Kinoshita-Kikuta, K. Takiyama and T. Koike Phosphate-binding tag: A new tool to visualize phosphorylated proteins. *Mol Cell Proteomics*, (2005).

162. K. Martin, T. H. Steinberg, T. Goodman, B. Schulenberg, J. A. Kilgore, K. R. Gee, J. M. Beechem and W. F. Patton Strategies and solid-phase formats for the analysis of protein and peptide phosphorylation employing a novel fluorescent phosphorylation sensor dye. *Comb Chem High T Scr* 6, 331-339 (2003).

163. K. Martin, T. H. Steinberg, L. A. Cooley, K. R. Gee, J. M. Beechem and W. F. Patton Quantitative analysis of protein phosphorylation status and protein kinase activity on microarrays using a novel fluorescent phosphorylation sensor dye. *Proteomics* 3, 1244-55 (2003).

164. S. L. Tobey, B. D. Jones and E. V. Anslyn C3v symmetric receptors show high selectivity and high affinity for phosphate. *J Am Chem Soc* 125, 4026-7 (2003).

165. M. S. Han and D. H. Kim Visual detection of AMP and real-time monitoring of cyclic nucleotide phosphodiesterase (PDE) activity in neutral aqueous solution. Chemosensor-coupled assay of PDE and PDE inhibitors. *Bioorg Med Chem Lett* 13, 1079-82 (2003).

166. M. S. Han and D. H. Kim Naked-eye detection of phosphate ions in water at physiological pH: a remarkably selective and easy-to-assemble colorimetric phosphate-sensing probe. *Angew Chem Int Ed Engl* 41, 3809-11 (2002).

167. J. P. DeGnore and J. Qin Fragmentation of phosphopeptides in an ion trap mass spectrometer. *J Am Soc Mass Spectrom* 9, 1175-88 (1998).

168. M. Edelson-Averbukh, R. Pipkorn and W. D. Lehmann Phosphate group-driven fragmentation of multiply charged phosphopeptide anions. Improved recognition of peptides phosphorylated at serine, threonine, or tyrosine by negative ion electrospray tandem mass spectrometry. *Anal Chem* 78, 1249-56 (2006).

169. A. Tholey, J. Reed and W. D. Lehmann Electrospray tandem mass spectrometric studies of phosphopeptides and phosphopeptide analogues. *J Mass Spectrom* 34, 117-23 (1999).

170. M. Salek, S. Costagliola and W. D. Lehmann Protein tyrosine-O-sulfation analysis by exhaustive product ion scanning with minimum collision offset in a NanoESI Q-TOF tandem mass spectrometer. *Anal Chem* 76, 5136-42 (2004).

171. H. Steen, B. Kuster, M. Fernandez, A. Pandey and M. Mann Detection of tyrosine phosphorylated peptides by precursor ion scanning quadrupole TOF mass spectrometry in positive ion mode. *Anal Chem* 73, 1440-8 (2001).

172. J. J. Coon, B. Ueberheide, J. E. Syka, D. D. Dryhurst, J. Ausio, J. Shabanowitz and D. F. Hunt Protein identification using sequential ion/ion reactions and tandem mass spectrometry. *Proc Natl Acad Sci U S A* 102, 9463-8 (2005).

173. J. E. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz and D. F. Hunt Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A* 101, 9528-33 (2004).

174. S. D. Shi, M. E. Hemling, S. A. Carr, D. M. Horn, I. Lindh and F. W. McLafferty Phosphopeptide/phosphoprotein mapping by electron capture dissociation mass spectrometry. *Anal Chem* 73, 19-22 (2001).

175. A. Stensballe, O. N. Jensen, J. V. Olsen, K. F. Haselmann and R. A. Zubarev Electron capture dissociation of singly and multiply phosphorylated peptides. *Rapid Commun Mass Spectrom* 14, 1793-800 (2000).

176. R. A. Zubarev, N. L. Kelleher and F. W. McLafferty Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.* 120, 3265-3266 (1998).

177. V. Wagner, G. Gessner, I. Heiland, M. Kaminski, S. Hawat, K. Scheffler and M. Mittag Analysis of the phosphoproteome of *Chlamydomonas reinhardtii* provides new insights into various cellular pathways. *Eukaryot Cell* 5, 457-68 (2006).

178. R. A. Bradshaw Revised draft guidelines for proteomic data publication. *Mol Cell Proteomics* 4, 1223-5 (2005).

179. S. Carr, R. Aebersold, M. Baldwin, A. Burlingame, K. Clauser and A. Nesvizhskii The need for guidelines in publication of peptide and protein identification data: Working Group on Publication Guidelines for Peptide and Protein Identification Data. *Mol Cell Proteomics* 3, 531-3 (2004).

180. M. Taniguchi, Y. Taniguchi, M. Kawasaki, S. Takeda, T. Kato, S. Sato, S. Tabata, H. Miyake and T. Sugiyama Identifying and characterizing plastidic 2-oxoglutarate/malate and dicarboxylate transporters in *Arabidopsis thaliana*. *Plant Cell Physiol* 43, 706-17 (2002).

181. I. I. Hanning, K. Baumgarten, K. Schott and H. W. Heldt Oxaloacetate transport into plant mitochondria. *Plant Physiol* 119, 1025-32 (1999).

182. C. Zoglowek, S. Kromer and H. W. Heldt Oxaloacetate and Malate Transport by Plant Mitochondria. *Plant Physiol* 87, 109-115 (1988).

183. G. E. Reid, K. D. Roberts, E. A. Kapp and R. I. Simpson Statistical and mechanistic approaches to understanding the gas-phase fragmentation behavior of methionine sulfoxide containing peptides. *J Proteome Res* 3, 751-9 (2004).

184. T. H. Steinberg, B. J. Agnew, K. R. Gee, W. Y. Leung, T. Goodman, B. Schulenberg, J. Hendrickson, J. M. Beechem, R. P. Haugland and W. F. Patton Global quantitative phosphoprotein analysis using Multiplexed Proteomics technology. *Proteomics* 3, 1128-44 (2003).

185. M. Wind, D. Gosenca, D. Kubler and W. D. Lehmann Stable isotope phospho-profiling of fibrinogen and fetuin subunits by element mass spectrometry coupled to capillary liquid chromatography. *Anal. Biochem.* 317, 26-33 (2003).

186. M. Wind, I. Feldmann, N. Jakubowski and W. D. Lehmann Spotting and quantification of phosphoproteins purified by gel electrophoresis and laser ablation-element mass spectrometry with phosphorus-31 detection. *Electrophoresis* 24, 1276-1280 (2003).

187. S. E. Ong and M. Mann Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* 1, 252-62 (2005).

188. R. Zhang and F. E. Regnier Minimizing resolution of isotopically coded peptides in comparative proteomics. *J Proteome Res* 1, 139-47 (2002).

189. I. C. Guerrera and O. Kleiner Application of mass spectrometry in proteomics. *Biosci Rep* 25, 71-93 (2005).

190. R. Kruger, D. Kubler, R. Pallisse, A. Burkovski and W. D. Lehmann Protein and proteome phosphorylation stoichiometry analysis by element mass spectrometry. *Anal Chem* 78, 1987-94 (2006).

191. M. Wind, M. Edler, N. Jakubowski, M. Linscheid, H. Wesch and W. D. Lehmann Analysis of protein phosphorylation by capillary liquid chromatography coupled to element mass spectrometry with 31P detection and to electrospray mass spectrometry. *Anal Chem* 73, 29-35 (2001).

192. M. Wind, H. Wesch and W. D. Lehmann Protein phosphorylation degree: determination by capillary liquid chromatography and inductively coupled plasma mass spectrometry. *Anal Chem* 73, 3006-10 (2001).

193. L. M. DiBella and S. M. King Dynein motors of the Chlamydomonas flagellum. *Int Rev Cytol* 210, 227-68 (2001).

194. J. Pan and W. J. Snell Signal transduction during fertilization in the unicellular green alga, Chlamydomonas. *Curr Opin Microbiol* 3, 596-602 (2000).

195. M. Eisenbach Control of bacterial chemotaxis. *Mol Microbiol* 20, 903-10 (1996).

196. J. S. Tash Protein phosphorylation: the second messenger signal transducer of flagellar motility. *Cell Motil Cytoskeleton* 14, 332-9 (1989).

197. R. Pilu, D. Panzeri, G. Gavazzi, S. K. Rasmussen, G. Consonni and E. Nielsen Phenotypic, genetic and molecular characterization of a maize low phytic acid mutant (lpa241). *Theor Appl Genet* 107, 980-7 (2003).

198. C. I. Ruse, B. Willard, J. P. Jin, T. Haas, M. Kinter and M. Bond Quantitative dynamics of site-specific protein phosphorylation determined using liquid chromatography electrospray ionization mass spectrometry. *Anal Chem* 74, 1658-64 (2002).

199. Y. G. Tsay, Y. H. Wang, C. M. Chiu, B. J. Shen and S. C. Lee A strategy for identification and quantitation of phosphopeptides by liquid chromatography/tandem mass spectrometry. *Anal Biochem* 287, 55-64 (2000).

200. F. Wolschin, U. Lehmann, M. Glinski and W. Weckwerth An integrated strategy for identification and relative quantification of site-specific protein phosphorylation using liquid chromatography coupled to MS(2)/MS(3). *Rapid Commun Mass Spectrom* 19, 3626-32 (2005).

201. H. Steen, J. A. Jebanathirajah, M. Springer and M. W. Kirschner Stable isotope-free relative and absolute quantitation of protein phosphorylation stoichiometry by MS. *Proc Natl Acad Sci U S A* 102, 3948-53 (2005).

202. H. Sakakibara, M. Taniguchi and T. Sugiyama His-Asp phosphorelay signaling: a communication avenue between plants and their environment. *Plant Mol Biol* 42, 273-8 (2000).

203. H. Sakakibara Nitrate-specific and cytokinin-mediated nitrogen signaling pathways in plants. *J Plant Res* 116, 253-7 (2003).

204. D. Wessel and U. I. Flugge A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* 138, 141-3 (1984).

205. A. Otto, B. Thiede, E. C. Muller, C. Scheler, B. WittmannLiebold and P. Jungblut Identification of human myocardial proteins separated by two-dimensional electrophoresis using an effective sample preparation for mass spectrometry. *Electrophoresis* 17, 1643-1650 (1996).