Dimitri Heintz[1]
Virginie Wurtz[2]
Anthony A. High[2]
Alain Van Dorsselaer[2]
Ralf Reski[1]
Eric Sarnighausen[1]

[1]University of Freiburg, Plant Biotechnology, Freiburg, Germany
[2]Laboratoire de Spectrométrie de Masse Bio-Organique, Strasbourg, France

An efficient protocol for the identification of protein phosphorylation in a seedless plant, sensitive enough to detect members of signalling cascades

We describe a reproducible protocol to explore for the first time the phosphoproteome of a seedless plant, the moss *Physcomitrella patens*. Following tryptic digestion of a total protein extract, phosphorylated peptides were isolated using the combination of C18 reverse-phase chromatography (RP-C18), immobilized Fe^{3+} metal affinity chromatography (IMAC), capillary zone electrophoresis (CZE), liquid chromatography-tandem mass spectrometry (LC-MS/MS) and matrix assisted laser desorption/ionization-time of flight-mass spectrometry (MALDI-TOF-MS) analysis. The total protein extracts were first prepared as usually made for plant two-dimensional gel electrophoresis, the tryptic digest was desalted and concentrated by reverse phase chromatography, and from this mixture the phosphorylated peptides were captured by IMAC. Subsequently, the complex phosphopeptide mixture was separated into ten fractions by RP-C18-HPLC and each analyzed by CZE. This permitted the detection of 253 distinct phosphopeptides. These were identified by nano-LC-MS/MS and MALDI-TOF-MS analysis in conjunction with alkaline phosphatase treatment to remove covalently bound phosphate to specifically identify the phosphopeptides. Among others, several kinases and a transcription factor were identified. This protocol will be taken as a basis to unravel early events in plant signal transduction known to occur *via* rapid phosphorylation/dephosphorylation of proteins.

Keywords: Immobilized metal affinity chromatography / Mass spectrometry / Moss / Phosphorylation / *Physcomitrella patens* / Proteomics DOI 10.1002/elps.200305795

1 Introduction

Phosphorylation represents an important post-translational modification of proteins; in eukaryotes ~ 30% of cellular proteins contain covalently bound phosphate [1]. Reversible phosphorylation is a key mode of signal transduction, a central mechanism in the modulation of protein function, able to regulate almost all aspects of cell life in both prokaryotes and eukaryotes. Phosphorylation of a protein can alter its behavior in almost every conceivable way: modulation of its intrinsic biological activity, half-life, subcellular location, and docking with other proteins [2]. The extent of its importance is illustrated by the hundreds of conventional protein kinases and phosphatases detected in various eukaryotic genomes [3, 4]. Phosphorylation of proteins occurs also in prokaryotes and is mediated by kinases substantially different from eukaryotic serine/threonine kinases [5].

The seedless plant *Physcomitrella patens* offers several advantages as model organism for signal transduction studies involving phosphorylation and dephosphorylation events in plants [6]. Transformation studies in this moss have revealed a totally unique feature for plants, *i.e.*, that foreign DNA sequences integrate into the nuclear genome preferentially at targeted locations by homologous recombination [7, 8], enabling for the first time in plants the application of powerful molecular genetic approaches used routinely in bacteria, yeast and, since 1989, mouse embryonic stem cells [9, 10]. Recently, two different expressed sequence tags databases have been built, rendering proteomics studies feasible in *Physcomitrella patens*, one with about 50 000 EST sequences being in the public domain [11], another database with ~ 110 000 sequences established from our group in collaboration with an industrial partner [12]. The effects of the different phytohormones on cell differentiation in *Physcomitrella patens* are well investigated [13], and numerous mutants

Correspondence: Eric Sarnighausen, University of Freiburg, Plant Biotechnology, Schaenzlestr. 1, D-79104 Freiburg, Germany
E-mail: Eric.Sarnighausen@biologie.uni-freiburg.de
Fax: +49-761-203-6945

Abbreviations: DHB, 2,5-dihydroxybenzoic acid; **IDA**, imidodiacetic acid; **IMAC**, immobilized metal affinity chromatography; **4HCCA**, α-cyano-4-hydroxycinnamic acid; **NTA**, nitrilotriacetic acid; **TBP**, tributylphosphine

are available for different stages of development making it an ideal system to study phosphorylation/dephosphorylation events.

Methods for analysis of the "phosphorylome", *i.e*., the subset of proteins in the proteome that become modified *in vivo* by phosphorylation, involve the use of radioactive labelling either *in vivo* or *in vitro* [14]. However, uneven uptake of the label in complex multicellular organisms and the large pools of endogenous free phosphate often limit conclusions [14, 15]. Phosphoamino acid antibodies have been exploited recently but their use is restricted to tailor-made immunological applications and because these antibodies cannot detect the nonphosphorylated form they are unable to determine stoichiometry [16, 17].

Currently, two-dimensional gel electrophoresis (2-DE) remains the most common methodology for assessing wide-scale changes in phosphorylation [18]. However, this methodology is limited by a number of well-described problems as well as the technique being relatively slow. Limitations include its inability to directly detect medium and low-abundance proteins from a whole cell lysate [19]. This is relevant in the case of phosphorylation studies as kinases and phosphatases often exist at very low copy numbers per cell. Also, in most cases the use of 2-DE to study phosphorylation results in the use of further techniques to characterize the proteins isolated from the gels. These techniques include but are not limited to Edman degradation [20], 2-D tryptic phosphopeptide mapping [21], and mass spectrometry (MS).

Several recent publications have described MS techniques to study changes in phosphorylation patterns in combination with immobilized metal affinity chromatography (IMAC) as an alternative to 2-DE [22–24]. The recent introduction of derivatization methods applied on phosphopeptides by fine chemistry before MS analysis eliminate limits of the IMAC method: nonspecific adsorption of nonphosphorylated peptides, low efficiency of adsorption of phosphopeptides in many cases, and suppressive effect of nonphosphorylated peptides [25]. Recently, Ficarro *et al*. [22] attained a much higher specificity using methyl esterification of acidic residues before IMAC enrichment. In this way, they were able to identify 383 phosphorylation sites from over 1000 phosphopeptides that were detected from *Saccharomyces cerevisiae* cell lysates. Two other chemical derivatization methods to enrich phosphorylated species from complex mixtures have been reported. The method of Oda *et al.* [23] designed a strategy in which the phosphate group on serine and threonine was replaced by β-elimination and further tagged with biotin in order to be selectively fished out by using immobilized streptavidin. Nevertheless, if phosphorylated serine residues undergo this reaction quite easily it is not reliable for threonine residues. This method does not distinguish between *O*-glycosylated and phosphorylated serine/threonine residues requiring additional experiments to confirm phosphorylation. Zhou *et al*. [24] made use of a more complex reaction scheme to capture peptides containing phosphorylated serine, threonine, and tyrosine residues on a solid support containing immobilized iodoacetyl groups. The major disadvantage of these two last methods is that a large amount of sample is required, precluding analysis of low-abundance proteins [25, 26].

We have used *Physcomitrella patens* as a model organism to expand and improve current analytical methods that use MS for the detection and identification of phosphoproteins/phosphopeptides. Here, we describe a multidimensional approach to detect phosphorylation sites without inducing any artificial modifications in an attempt to improve total protein extraction and gain access to low-abundance proteins. Emphasis is placed on sample pretreatment and IMAC protocols to reduce the amount of nonphosphopeptide binding often seen in other published papers. We have developed a sensitive protocol suited for exploring for the first time the phosphoproteome of a seedless plant that is compatible with subsequent MS analysis. We have applied this protocol to the analysis of a tryptic-digested cell extract. We report the detection of 253 distinct phosphopeptides by immobilized Fe^{3+} metal affinity chromatography and 2-D LC combining a gradient elution on HPLC-C18 reverse-phase chromatography with capillary zone electrophoresis (CZE). The phosphorylation states of the phosphopeptides were confirmed by MALDI-TOF-MS in positive mode (following the loss of 80 Da (owing to HPO_3 loss) or neutral loss of 98 Da (owing to H_3PO_4 loss) or multiples of 80 or 98 Da mass shifts from the phosphopeptides after alkaline phosphatase treatment), or by LC-MS/MS. The phosphopeptides were sequenced using MS/MS spectra. The MS/MS spectra were subjected to database searching and *de novo* sequencing tools for the identification of all proteins. Our results clearly demonstrate the power of this technique to study phosphorylation events that occur in *Physcomitrella patens* as well as provide a new analytical methodology that should be applicable to the study of phosphorylation/dephosphorylation in any biological system.

2 Materials and methods

2.1 Materials

Trizol was from Gibco-BRL (Carlsbad, CA, USA). Plant protease inhibitor, urea (electrophoresis reagent), 3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate

(CHAPS), and the model protein β-casein were purchased from Sigma-Aldrich (St. Louis, MO, USA). Trifluoroacetic acid, tributylphosphine (TBP), and iron(III) chloride hexahydrate were from Fluka (Buchs, Switzerland). Enzymes: sequencing-grade modified trypsin was from Promega (Madison, WI, USA), alkaline phosphatase (1×10^3 U/mL) was from Roche Applied Science (Mannheim, Germany). Reversed-phase packing POROS® 50 R2 was from PerSeptive Biosystems (Framingham, MA, USA) and oligo R3 reversed-phase packing from Applied Biosystems (Foster City, CA, USA). IMAC matrices: Ni-NTA Superflow was from Qiagen (Hilden, Germany), Fractogel® EMD chelate (S) was from Merck (Darmstadt, Germany) and SelfPack POROS® 20 MC was from Applied Biosystems. Capillary electrophoresis buffers and reagents were purchased from Bio-Rad (Hercules, CA, USA). MALDI matrices α-cyano-4-hydroxycinnamic acid (4HCCA) and 2,5-dihydroxybenzoic acid (DHB) were from Sigma-Aldrich. MS calibration standards angiotensin (1046.54 Da), substance P (1347.74 Da), bombesin (1620.81 Da), ATCH (2465.20 Da) and poly-DL-alanine (M_r 1000–5000; for MS/MS calibration) were from Sigma-Adrich. All other chemicals were of reagent grade. All solutions were prepared with HPLC-grade (bidistilled) water.

2.2 Growth of plant material

Physcomitrella patens protonema was grown in semicontinuous 5 L bioreactors in modified Knop medium as described by Hohe *et al.* [27]. Plant material was harvested from 1 L of culture medium by paper filtration in a buchner funnel (diameter, 80 mm) with suction. Afterwards, the material was immediately frozen in liquid nitrogen and stored at −80°C. On average, 0.9 g of protonema (fresh weight) could be obtained from 1 L of bioreactor culture.

2.3 Extraction of total proteins

Frozen material was ground to a fine powder in a ball mill (Mikro-Dismembrator S, B. Braun Biotech International, Melsungen, Germany) at 2600 rpm for 1 min. Plant material, PTFE shaking flasks, and grinding balls were precooled in liquid nitrogen in order to avoid thawing of the material during the process. 60 mg of ground moss was extracted with 1.9 mL of acetone, 13 mM DTT at room temperature and kept at −20°C for 2 h. After centrifugation at $19\,000 \times g$ at −10°C for 15 min, the pellet was washed with ice-cold acetone, 13 mM DTT twice, and subsequently dried under vacuum for 15 min. Total proteins were extracted from the dry material with 100 μL 9 M urea, 2% w/v CHAPS, 5 mM TBP with shaking at room temperature for 1 h. Insoluble material was removed by centrifugation at $19\,000 \times g$ at room temperature for 15 min. The average protein content was 10 mg/mL, determined according to a modified Bradford protocol [28]. Alternatively, the ground material was either extracted directly in Laemmli sample buffer [29] with heating to 95°C for 5 min or protonema was subjected to Trizol extraction as described by Ficcaro *et al.* [22]. In the latter case, 100 mg of frozen material were ground in the ball mill in the presence of 1.5 mL Trizol and precipitation as well as washing of the proteins was performed strictly following the manufacturer's protocol. The proteins were subsequently solubilized with 100 μL of 1% SDS and dialyzed against 1% SDS using a molecular weight cutoff of 10 000.

2.4 Tryptic digestion

Tryptic digestion and subsequent procedures towards the identification of phosphopeptides were performed in the same manner for *Physcomitrella* proteins and the model protein β-casein. One milligram of protein in 100 μL of 9 M urea, 2% CHAPS, 5 mM TBP was diluted with 1.5 mL of 50 mM NH_4HCO_3, pH 7.8. Plant protease inhibitor cocktail was added at this point at a concentration of 1% v/v only when total proteins were to be analyzed by CE. After the addition of 20 μg trypsin (≥ 100 U), digestion was performed at 37°C for 16 h. The reaction was stopped by the addition of formic acid to a final concentration of 5%. Only if proteins had been extracted with Trizol was the digestion protocol modified according to Ficcaro *et al.* [22]. In this case, 50 μL of the total protein in 1% SDS (corresponding to 500 μg) were diluted with 500 μL of 100 mM ammonium acetate, pH 7.8, and digestion was performed after the addition of 20 μg trypsin at 37°C for 16 h. Peptides were dried by lyophilization, resolubilized in 400 μL of 2 M methanolic hydrochloric acid and allowed to stand at room temperature for 2 h. The resulting peptide methyl esters were again dried under vacuum and subsequently dissolved in 120 μL of a solution containing equal parts of methanol/water/acetonitrile. Only in this case IMAC was performed directly, omitting the otherwise essential desalting step.

2.5 Desalting and concentration of total peptides

Desalting and the concentration of the peptide mixture was accomplished using a combination of two different reverse-phase chromatography systems according to Neubauer and Mann [30]. Self-made chromatography columns 0.5 cm in diameter were packed with 60 μL of POROS® 50 R2 reversed-phase packing or 100 μL of oligo R3 reversed-phase packing, respectively, and equi-

librated with 0.5% formic acid. The peptide solution (1.6 mL) acidified with a final concentration of 5% formic acid was applied to the oligo R2 column, which was then washed once with 500 μL of 0.5% formic acid. This flow-through was applied directly onto the oligo R3 column. The oligo R2 column was subsequently eluted with 100 μL of 20% acetonitrile in 0.5% formic acid, 100 μL of 30% acetonitrile in 0.5% formic acid and 500 μL of 80% acetonitrile in 0.5% formic acid. Elution of peptides from the oligoR3 column was accomplished by application of 250 μL 50% methanol followed by 250 μL of 50% methanol in 5% ammonia. The eluted peptides were dried under vacuum, redissolved in 50 μL of 0.1 M acetic acid, and applied to the IMAC column.

2.6 Isolation of phosphopeptides

The phosphopeptides were enriched from the peptide mixtures by IMAC. Self-made chromatography columns (0.5 cm in diameter) were filled with 100 μL of Fractogel® EMD chelate (S), SelfPack POROS® 20 MC or Ni-NTA Superflow, respectively. In the case of the latter, nickel was displaced with 10 volumes of 100 mM EDTA, 50 mM Tris-HCl, pH 8.0, buffer according to Hart *et al.* [31] and the column was subsequently rinsed with 10 volumes of bidistilled water. The matrices were washed with two bed volumes of 0.1 M acetic acid prior to the application of the same volume of 100 mM iron(III) chloride in 0.1 M acetic acid. Excess iron was removed with two bed volumes of 0.1 M acetic acid prior to the application of the peptides. In order to remove unspecific binding of the applied peptides, the column was washed with 600 μL of 100 mM NaCl in water:acetonitrile:acetic acid (75/25/1 v/v/v). Following a further wash step with 200 μL of 0.1 M acetic acid, phosphopeptides were eluted with 500 μL of 50 mM Na_2HPO_4, pH 9.4, according to Ficarro *et al.* [22].

2.7 Phosphopeptide separation

The subsequent fractionation of phosphopeptides was performed by reversed-phase HPLC. A precolumn (ChromCart, Nucleosil, 8/3, 300-5C18; Macherey-Nagel) was connected to the analytical HPLC column (Nucleosil, 300–5 μm, 125 × 2 mm ID; Macherey-Nagel). The sample was injected *via* a sample loop of 500 μL (Waters 2790 sample loop 500 μL). The chromatography solvents were 0.1% TFA in H_2O (solvent I), and 0.1% TFA, 99.9% acetonitrile (solvent II). In order to remove Na_2HPO_4, the column was washed with 2% solvent II for 10 min and subsequently with 4% solvent II for 6 min. The HPLC column was then developed with a linear gradient from 4 to 80% solvent II at a flow rate of 0.25 mL/min over 50 min. The dead volume from the top of the HPLC column to the fraction collector corresponds to a time delay of 2 min and 30 s. Fractions were collected every 3 min starting with the wash step to the end of the gradient. The collected fractions were dried under vacuum prior to CZE and MS analysis. Elution of phosphopeptides occurred within a total time of 30 min starting from minute 16 to minute 46 covering ten distinct fractions (A to J in Fig. 5).

2.8 Capillary electrophoresis

2.8.1 Capillary zone electrophoresis

Phosphopeptide profiles were analyzed by CZE (Bio-Focus 2000 Capillary Electrophoresis System; Bio-Rad, Hercules, CA, USA). Total peptide pellets were dissolved in 10 μL of 0.5% formic acid and electrophoretic separation was performed in 0.1 M phosphate buffer, pH 2.5, using a coated capillary with an inner diameter of 50 μm and an effective length of 50 cm (BioCAP™ LPA coated capillary; Bio-Rad). Samples were introduced into the capillary by electrophoretic injection at 10 kV for 5 s. The polarity was set from positive to negative and the capillary temperature was kept at 20°C. Peptides were detected by monitoring the absorbance at 200 nm. At a constant voltage of 13 kV a complete peptide profile could be obtained within 60 min.

2.8.2 SDS-capillary electrophoresis

Total protein patterns were analyzed either directly after extraction in 9 M urea, 2% CHAPS, 5 mM TBP or following the subsequent 15-fold dilution with 50 mM NH_4HCO_3 pH 7.8. Total protein extracts in Laemmli sample buffer [29] served as a positive control. A negative control was applied by injecting buffer without protein for baseline detection. Protein separation was performed employing a CE-SDS Protein-Kit (Bio-Rad) as described by the manufacturer using a noncoated capillary, 50 μm ID × 375 μm OD, with 29.5 cm total length and 24 cm effective length. An aliquot of each fraction from the different conditions mentioned above and corresponding to 5 μg total protein to be analyzed was diluted with 30 μL of CE-SDS protein sample buffer and introduced into the capillary by electrophoretic injection at 10 kV for 5 s. The polarity was set from negative to positive at an applied voltage of 15 kV, the capillary temperature was set to 20°C and the absorbance was monitored at 220 nm. The running time to stop was 15 min.

2.9 MALDI-TOF-MS and MALDI-TOF-MS/MS

MALDI-TOF-MS and MALDI-TOF-MS/MS were performed on an Ultraflex™ TOF-TOF mass spectrometer (Bruker-Daltonik, Bremen, Germany). The instrument was used at a maximum accelerating potential of 20 kV and was operated either in reflector positive or reflector negative mode. The matrix used in positive mode was a freshly saturated solution of recrystallized 4HCCA in 50% acetonitrile. Alternatively, DHB in water was used in negative operation mode and when phosphopeptides were to be treated with alkaline phosphatase. MS calibration was performed in external mode with four peptides, angiotensin (1046.54 Da), substance P (1347.74 Da), bombesin (1620.81 Da), and ATCH (2465.20 Da). For MS/MS (positive mode only) a 0.5 μL aliquot of sample was spotted on the target, followed by 0.5 μL of saturated matrix solution in acetronitrile.

2.10 Nano-LC-MS/MS

The sample (6.4 μL) was injected into a CapLC system (Waters, Milford, MA, USA) equipped with an autosampler, gradient and auxiliary pump *via* the "microliter pick-up" mode. The samples were desalted on-line through a 300 μm × 5 mm C_{18} trapping cartridge (LC Packings, San Fransisco, CA, USA) at high flow (30 μL/min) for 3 min. Peptides were separated on a 75 μm × 15 cm 3 μm C_{18} 100 Å PepMap™ column (LC Packings) prior to introduction into the mass spectrometer. A typical reversed-phase gradient from 0% to 95% acetonitrile in 0.1% formic acid in 35 min was used. The flow rate was 5 μL/min. The system utilized a split flow resulting in a column flow rate of ~ 400–500 nL/min. Mass calibration was performed using poly-DL-alanine.

2.11 Alkaline phosphatase treatment

For MALDI-TOF-MS, alkaline phosphatase treatment was performed according to Larsen *et al.* [32]. The previously analysed sample in DHB was redissolved on the target with 1.5 μL 50 mM NH_4HCO_3, pH 7.8, containing 0.05 U/μL alkaline phosphatase. The target was incubated at high humidity at 37°C for 30 min. Subsequently, the sample was acidified with 0.5 μL of 5% TFA, and the matrix was allowed to recrystallize. For LC-MS/MS, the sample was solubilized directly with 13 μL of 50 mM NH_4HCO_3, pH 7.8. 6.4 μL was taken, acidified with 0.5 μL of 5% TFA, and injected without alkaline phosphatase, and of the remaining sample 6.4 μL was taken, 0.05 U/μL alkaline phosphatase was added, and the sample was incubated at 37°C for 30 min. Subsequently, the sample was acidified with 0.5 μL of 5% TFA. The identification of phosphopeptides was possible by observing a loss of 80 Da (owing to HPO_3 loss) or neutral loss of 98 Da (owing to H_3PO_4 loss) or multiples of 80 or 98 Da mass shifts after alkaline phosphatase treatment.

2.12 Protein identification and sequence processing

MS/MS data was obtained using a Q-Tof 2 (Micromass, Manchester, UK) fitted with a Z-spray nanoflow electrospray ion source. The mass spectrometer was operated in positive ion mode with a potential of 3500 V applied to the nanoflow probe body. The collision energy was determined on the fly based on the mass and charge state of the peptide. Charge state recognition was used to switch only doubly and triply charged ions into MS/MS mode. Several trypsin autolysis ions were excluded. Mass data acquisitions were piloted by MassLynx software (Micromass) using automatic switching between MS and MS/MS modes. Mass data collected during LC-MS/MS analysis were processed and converted into a .PKL file to be submitted to the search software MASCOT (Matrix Science, London, UK). Searches were done with a tolerance on mass measurement of 0.25 Da in MS mode and 0.5 Da in MS/MS mode. If no significant hits were obtained, MS/MS ion spectra were interpreted manually to obtain amino acid tags. The amino acid tags were subsequently used to interrogate the *Physcomitrella patens* genome public EST contig database (http://moss.nibb.ac.jp/blast/blast.htmL), protein identification was accomplished based on the annotated public EST database according to Nishiyama *et al.* [11], using the NCBI (National Center for Biotechnology) nonredundant protein database (http://www.ncbi.nlm.nih.gov/blast/).

3 Results and discussion

We describe a protocol that combines protein extraction under highly denaturating conditions, protein tryptic digestion, isolation of phosphopeptides by IMAC, 2-D fractionation and mapping of phosphopeptides by HPLC/CZE, and identification by MS. To our knowledge, this is the first report addressing the application of the IMAC technique to plant tissues.

3.1 Protein extraction

The protein extraction procedure has to meet three major criteria. It has to permit the solubilization of all proteins (*e.g.*, all phosphoproteins), it must be compatible with the subsequent purification and detection procedures, and the proteins themselves have to remain in a state

that allows their detection and identification (avoidance of involuntary modifications). Effective protein denaturation will inactivate endogenous protease activities, which is a central requirement if the method is to provide reliable results when applied to total extracts rather than model proteins. We found that endogenous proteases were indeed inactivated irreversibly at 95°C in the presence of SDS. The treatment did, however, turn out to be incompatible with further purification steps due to the co-extraction of interfering compounds from the plant material (results not shown).

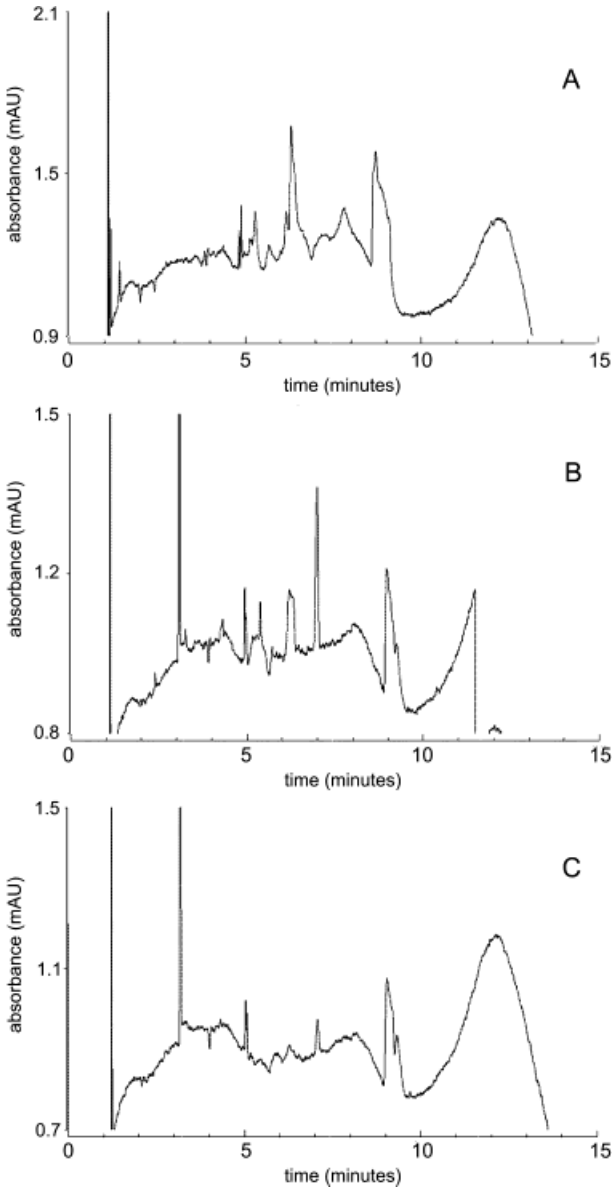

Figure 1. SDS-CE separation of total *Physcomitrella* protonema proteins. After extraction with 9 M urea, 2% CHAPS, and 5 mM TBP, protein samples were diluted 15-fold with 50 mM NH_4HCO_3 pH 7.8 (A) and incubated at 37°C for 16 h (B). Comparison with the corresponding profile of the blank control lacking protein from the start (C) indicates that most proteins are not stable during the incubation period.

Total protein solubilization involved in 2-D electrophoresis includes more gentle denaturation procedures that aim at reducing the amount of nonproteinous (interfering) compounds [33]. Therefore, we decided to extract total proteins using a modified urea denaturation protocol. IMAC purification relies on the fragmentation of proteins into peptides, usually performed by tryptic digestion. As trypsin activity is blocked by elevated concentrations of urea, digestion is only feasible after a dilution of the sample (information provided by the supplier). Here, the minimum dilution necessary to allow efficient tryptic digestion was determined as 15-fold by using bovine β-casein as a model protein (see below). However, while the total proteins extracted from moss protonema remained stable in the presence of 9 M urea for a prolonged period of time, the conditions allowing trypsin activity also restored the activity of endogenous moss proteases even after urea denaturation (Fig. 1). Addition of protease inhibitors, which does indeed provide a certain degree of protection (results not shown), is counterproductive in this case as trypsin itself belongs to the large group of serine proteases [34].

Only when proteins within the ground plant material were acetone-denaturated prior to extraction with 9 M urea, the inactivation of endogenous proteases was complete and irreversible (Fig. 2) and specific digestion could be accomplished. At the same time, the plant material was hereby depleted of acetone-soluble compounds increasing the purity of the sample facilitating subsequent isolation and identification steps.

3.2 Protein tryptic digestion

β-Casein was used as standard protein to determine the dilution steps necessary to optimize trypsin activity for complete protein digestion. Several sample dilution steps and enzyme concentrations were tested. Complete tryptic digestion in the urea buffer in the presence of CHAPS was only feasible when the concentration of the detergent was below 0.2%. Under this condition, the two expected β-casein tryptic-digested phosphopeptides with theoretical molecular masses of 3122.26 (for the multiphosphorylated peptide) and 2061.83 (for the monophosphorylated peptide), respectively, were exclusively detected after IMAC (Fig. 3). No additional phosphopeptide masses were observed, indicating that no partial digestion of β-casein leading to different phosphopeptide masses was encountered under this condition. Following tryptic digestion, the peptides were further purified and concen-

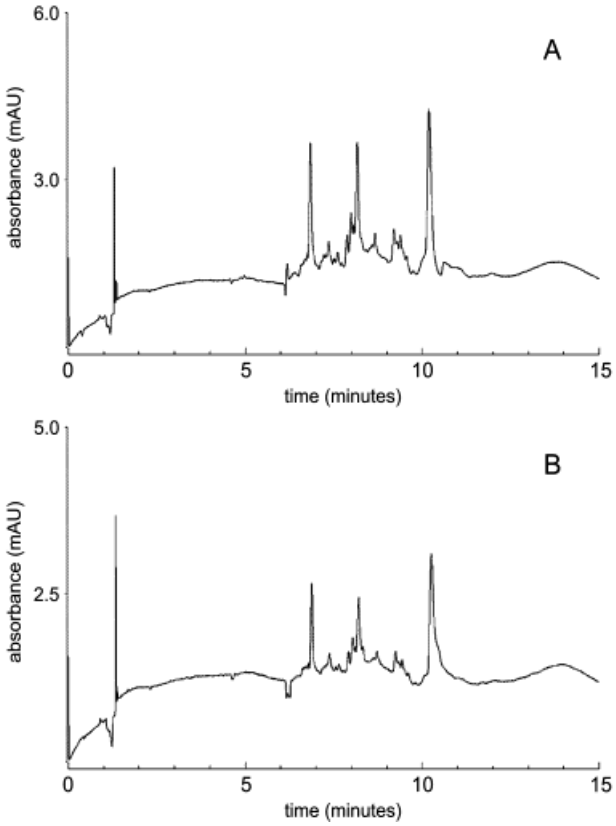

Figure 2. SDS-CE separation of total *Physcomitrella* protonema proteins. Following acetone precipitation/washing of the ground plant material, total proteins were extracted with 9 M urea, 2% CHAPS, and 5 mM TBP, the samples were diluted 15-fold with 50 mM NH_4HCO_3 pH 7.8 (A) and incubated at 37°C for 16 h (B). A comparison of the profiles confirms the stability of the proteins during the incubation period.

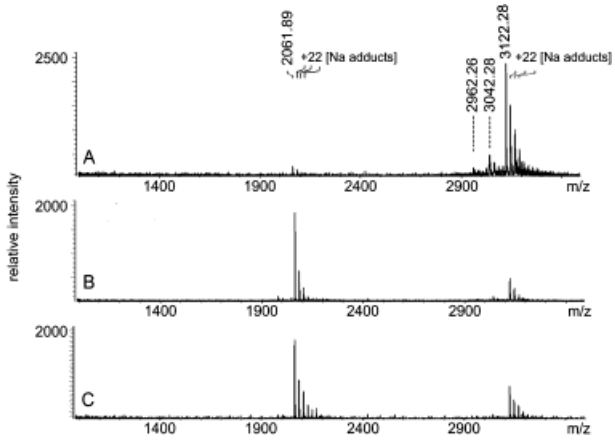

Figure 3. MALDI-TOF mass spectra of phosphopeptides from β-casein tryptic digest eluted from different IMAC matrices. In each case, 5 pmol of tryptic digest of β-casein was applied to the IMAC column. (A) Fractogel® EMD chelate (S) matrix; (B) SelfPack POROS® 20 MC matrix; (C) Fe-NTA Superflow matrix. Only in the case of the Fe-NTA Superflow matrix were both the monophosphorylated (2061.89 Da) and the polyphosphorylated (3122.26 Da) peptide detectable in corresponding amounts and this even though the quantification by means of MALDI mesurement is known to be problematic. All the other detected peaks are corresponding to Na adducts characterized by multiples of a 22 Da shift.

trated using the double column technique as described by Neubauer and Mann [30], thereby avoiding the loss of small and hydrophilic peptides during this indispensable step.

3.3 Immobilized metal affinity chromatography

IMAC, though commonly used for phosphopeptide isolation [35, 36], has frequently been reported to suffer from serious inadequacies rendering its use rather problematic. Among these were its avidity for negatively charged and hydrophobic amino acids, which gave rise to a low recovery of phosphopeptides due to a high background of nonspecific bindings [36, 37]. To avoid these problems from the start, preliminary experiments using a tryptic digest of bovine β-casein were conducted with the purpose of optimizing the IMAC performance. The well-characterized 24 kDa protein contains five phosphorylation sites, four of which are found in the tryptic peptide with a theoretical M_r of 3122.26 and the remaining one in the peptide with a theoretical M_r of 2061.83.

It has been shown by a number of research groups that nitrilotriacetic acid (NTA) is a more effective chelating agent for use in IMAC than the more widely used iminodiacetic (IDA) acid matrix [31]. Neville *et al.* [38] reported that iron(III)-NTA-type chelates had greater specificity for phosphorylated peptides than iron(III)-IDA-type chelates. In addition, iron(III)-NTA-type chelates are less prone to leaching and have greater stability than gallium(III)-NTA-type chelates. Three different IMAC matrices were chosen for comparison: two IDA matrices, the widely used IDA-type POROS® 20 MC resin, the IDA quadridendate Fractogel® EMD (S) tentacle chelate resin and an NTA quadridendate metal binding tentacle chelate resin matrix. The specificity of the matrices in their ability of retention and elution of mono- and multiphosphorylated peptides from a tryptic β-casein digest was analyzed.

The peptides were applied to the matrices at acidic pH and unspecific binding was further counteracted by including additional wash steps prior to peptide elution. The columns were rinsed with 100 mM sodium chloride in order to reduce ionic interactions between peptides and free carboxy groups of the matrix [39]. Under these con-

ditions, the three matrices tested provided clearly diverging results (Fig. 3). While elution of nonphosphorylated tryptic β-casein peptides was never observed, the NTA Superflow matrix gave the best results with respect to the isolation of mono- as well as polyphosphorylated peptides and was therefore chosen for the isolation of phosphopeptides from the total pool of tryptic peptides.

Recently, Ficcaro *et al.* [22] attained a high specificity of phosphopeptide recovery on IMAC using esterification of acidic amino acid residues. In addition, before IMAC enrichment, these authors used a strong acidic treatment to reduce nonspecific binding of the peptides to carboxylate groups. Their protocol made sequencing of hundreds of phosphopeptides from a total phenolic yeast protein extract feasible.

However, when applied to *Physcomitrella* protonema, this protocol was not applicable for two major reasons: (i) The chemical modification of the peptides, which is achieved by a strong acidic treatment with 2 M methanolic hydrochloric acid for 2 h, resulted in the loss of large parts of the material due to immediate and irreversible precipitation. (ii) The second major drawback became obvious upon sequencing of the phosphopeptides, as the MALDI-MS/MS analysis of the phosphorylated peptides revealed that some of them were obviously modified (Fig. 4), ion fragmentation with mass losses of 104, 142, or 120, respectively, were only detected when using the phenol-based method as described by Ficcaro *et al.* [22], regardless of whether peptides were converted to their corresponding methyl esters or not. Above all, the presence of these modifications rendered sequencing of the peptides impossible. For these reasons, this method was considered inapplicable at least to the analysis of protein phosphorylation in *Physcomitrella.*

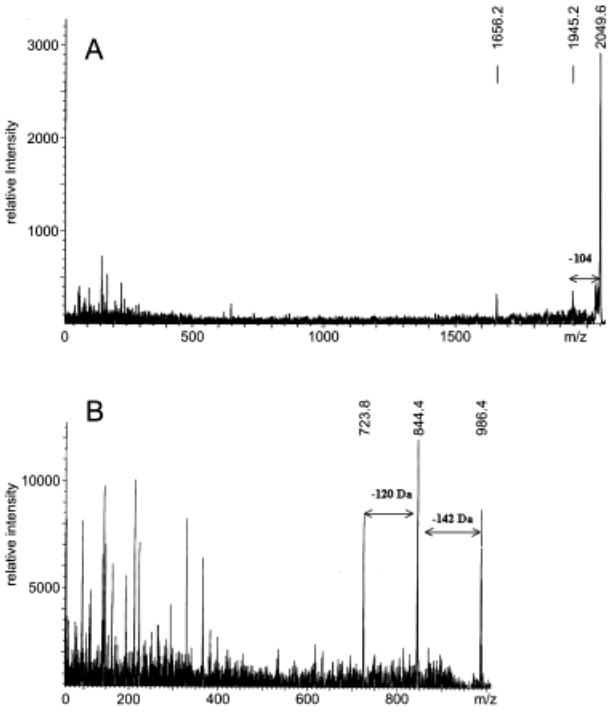

Figure 4. Detection of amino acid modification following protein extraction with Trizol. (A) MS/MS spectra of parent ion *m/z* 2049.6 obtained by MALDI-TOF-MS/MS. The loss of 104 Da could be attributed to benzoyl modification. (B) MS/MS spectra of parent ion *m/z* 986.4 obtained by MALDI-TOF-MS/MS. The loss of 142 Da could be attributed to either *N*-methyl glutamyl or *N*-methyl lysyl. The MS/MS spectra of parent ion *m/z* 844.4 was also obtained by MALDI-TOF-MS/MS. The loss of 120 Da could be attributed to benzyloxymethyl modification.

3.4 Peptide profiling

The complex mixture of peptides eluted from the IMAC column after application of tryptic peptides from total extracts of *Physcomitrella* protonema necessitated the combination of two separation techniques in order to render resolution feasible. Prefractionation of the sample into ten fractions by reversed-phase HPLC prior to CZE allowed the detection of 253 distinct peptides (Fig. 5). CZE is the method of choice for high-throughput phosphopeptide profiling due to its high sensitivity, its speed, and the requirement for only minimal amounts of sample [40, 41]. For each CZE fraction the complete analysis including the run time and automatic peak integration was achieved in less than 60 min. Comparison of CZE profiles between parallel preparations demonstrates the high degree of reproducibility achieved following the protocol described (Fig. 6). Therefore, the method fulfils the basic requirement for differential profiling and is applicable to the recording of alterations in protein phosphorylation patterns.

3.5 Confirmation of peptide phosphorylation

According to the literature [36, 37], a drawback of the IMAC method when applied to the isolation of phosphopeptides is its susceptibility to unspecific binding of negatively charged or hydrophobic amino acids. The resulting high background makes the detection of only low-abundant phosphopeptides difficult if not impossible. The efficiency of the novel protocol described here to discriminate against nonphosphorylated peptides was demonstrated by MALDI-TOF-MS implicating enzymatic dephosphorylation. Alkaline phosphatase treatment could be applied successfully to demonstrate peptide phosphorylation (Fig. 7). Out of 192 peptides analyzed by LC-MS/MS, and MALDI-TOF-MS (Heintz *et al.*, manuscript in preparation), 184 responded to the alkaline phosphatase

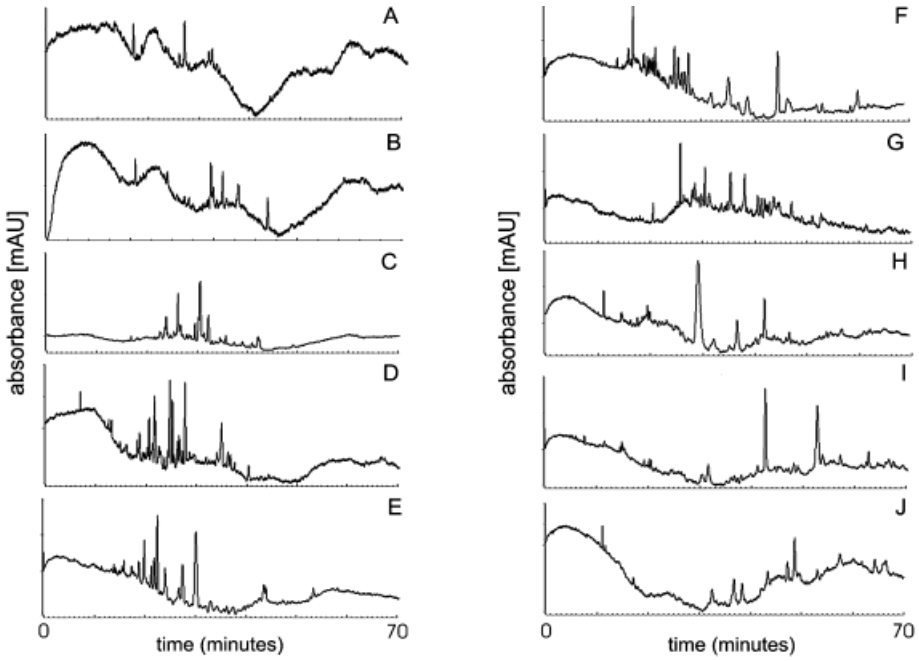

Figure 5. CZE analysis of phosphopeptides in ten different fractions (A–J) collected after linear gradient elution from HPLC-C18 reverse-phase chromatography column. The fractions were dried, and reconstituted with 10 μL of 0.5% formic acid for CZE analysis. The IMAC phosphopeptide containing fraction was applied to a HPLC-C18 reverse-phase chromatography column. Following a 16 min wash step to remove Na_2HPO_4, a linear gradient from 4 to 80% of (0.1% TFA, 99.9% acetonitrile) was formed over 50 min. The fractions from minute 16 (A) to minute 46 (J) contained phosphopeptides and were collected at 3 min intervals. Altogether 253 distinct peptides could be detected.

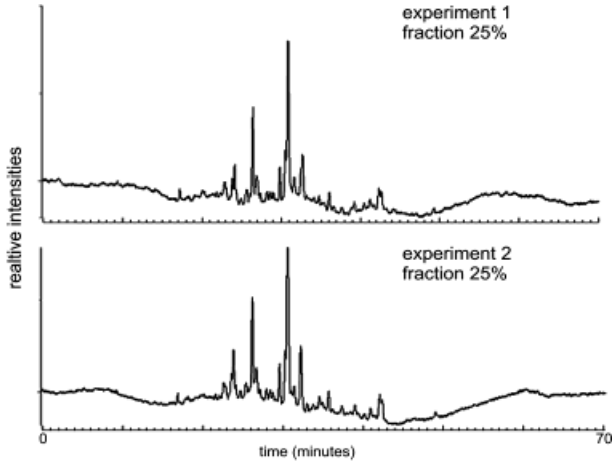

Figure 6. Electropherograms demonstrating the reproducible and highly efficient separation of *P. patens* phosphopeptides using CZE. Two aliquots of the same total peptide preparation were submitted sequentially to IMAC as described in Section 2.6. The CZE analysis of the phosphopeptide fraction eluted with 7.5% acetonitrile and 0.5% formic acid are shown.

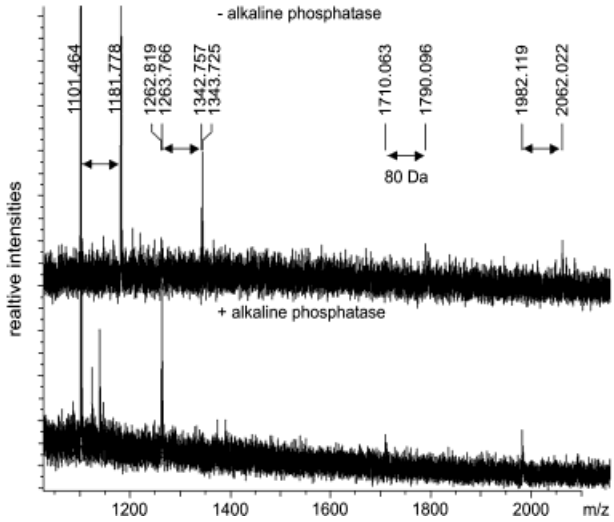

Figure 7. MALDI-TOF-MS, positive ion mass spectra of *P. patens* phosphopeptides, before and after (+A.P/-A.P) alkaline phosphatase treatment to confirm phosphorylation by a mass shift of 80 Da owing to the loss of HPO_3^-.

treatment with a mass shift corresponding to one or more phosphate groups, confirming that at least 95.8% of the peptides isolated by IMAC are phosphorylated. At this point, we are not able to say whether the remaining 4.2% of the peptides, bound unspecifically to the IMAC matrix, lost their phosphate groups prior to alkaline phosphatase treatment or display resistance towards enzymatic dephosphorylation, as described recently [42].

3.6 Identification of phosphopeptides

The sequencing of *Physcomitrella patens* phosphopeptides was accomplished by LC-MS/MS and MALDI-TOF-MS; the resulting amino acid sequences were used as queries with the tblastn program for searching ESTs, according to Jongeneel [43]. The search was restricted to the recent public *Physcomitrella patens* EST contig database, which provides the advantage that the EST

Table 1. Phosphopeptides identified in a tryptic digest of a *Physcomitrella patens* total protein extract

Protein name	Phosphopeptide partial sequence	NCBI GI No.	Contig No.	Species
Calcium/calmodulin-dependent protein kinase CaMK2	HGSLDASQTGPK[a)]	16904224	2739	*Nicotiana tabacum*
MAP kinase kinase (MAPK2)	HRIALTSMGAPE[b)]	7341300	6044	*Oryza sativa*
Protein kinase-like protein	GIEAGGESTF[a)]	13877617	2350	*Arabidopsis thaliana*
Transcription factor-like protein	LPTGFPLYGVE[a)]	15239113	5580	*Arabidopsis thaliana*
Probable protein kinase	GDWSGVTCLPK[a)]	235069	pph1o15	*Arabidopsis thaliana*
Phosphoribulokinase	KPDFDAFIDPQK[b)]	125578	2872	*Mesembryantheum crystallimum*

The listed proteins were identified by nano-LC-MS/MS, in HPLC phosphopeptide fractions obtained after IMAC of a tryptic-digested *Physcomitrella patens* total protein extract. The tryptic digested phosphopeptides were partial sequences and the positions which carry phosphorylation were not determined. (GI) Gene info Identifier number of the protein from the species identified in the nonredundant database of NCBI after a BLAST search with the corresponding EST contig from the public *Physcomitrella* EST database.
a) Monophosphorylated peptide, b) multiphosphorylated peptide

contig sequences have already been searched against a nonredundant protein database on NCBI. Among the identified and characterized proteins (Table 1) several kinases could be identified. One protein displays similarity to a calcium/calmodulin-dependent protein kinase from *Nicotiana tabacum*. This enzyme plays a role in the spatial and temporal regulation of intracellular Ca^{2+}, a key event in many signalling pathways [44]. We also identified a protein similar to a *Mesembryanthenum crystallium* (ice plant) phosphoribulokinase. This calvin cycle enzyme catalyses the reversible reaction of ATP/ADP exchange in a light-dependent manner on D-ribulose-5-phosphate/D-ribulose-1,5-bisphosphate [45]. Additionally, we also identified two kinase-like proteins and a transcription factor-like protein. These three proteins are similar to *Arabidopsis thaliana* proteins and their functions are yet unknown. Moreover, a *Physcomitrella* homologue to a MAP kinase kinase from *Oryza sativa* was identified, a protein involved in eukaryotic signal transduction [46]. The signal suppression of phosphate-containing molecules in the commonly used positive detection mode is an argument in favor of the dephosphorylation of phosphopeptides prior to sequencing. The use of alkaline phosphatase allows the analysis of peptides not amenable to positive ion MS in their phosphorylated form. The phosphorylated form of the peptides may not ionize well in the positive ion mode, preventing the identification of this peptide during analysis performed without the use of alkaline phosphatase [47]. These results suggest that dephosphorylation of phosphopeptides after IMAC enrichment prior to MS analysis is a useful tool if sequences of multiply phosphorylated peptides are to be obtained. Alkaline phosphatase treatment has been shown to enhance detection of multiply phosphorylated peptides [48].

4 Concluding remarks

We describe a reproducible protocol for identification of phosphoproteins from protonema tissue of the moss *Physcomitrella patens*. This is also the first time that such a study of protein phosphorylation is applied to a whole-cell lysate different from that of *Saccharomyces cerevisiae* [22]. We have developed a method using a 2-D buffer which should allow the extraction of all proteins present in a tissue without inducing any artificial modifications. To our knowledge, CE-SDS experiments performed to screen the stability of a total protein extract during sample preparation in a 2-D buffer have not previously been reported. One challenge in proteomics is reproducibility. The simplicity and the rapidity of the described CE-SDS method is helpful for the control of protein stability during sample preparation in proteomic applications when protease activity is suspected. In this study, we also report the utilization of CZE for the profiling of a high number of phosphopeptides. In the ten HPLC phosphopeptide fractions, more than 250 peptide peaks were automatically detected and unambiguously interpretable by the CZE. Although it has already been known for a long time that NTA-based IMAC matrices are superior to IDA for phosphopeptide enrichment, we have now shown that NTA is able to bind mono- as well as multiphosphorylated peptides, whereas the widely used IDA POROS matrix displays a higher avidity for multiphosphorylated peptides. MALDI-TOF-MS/MS and LC-MS/MS were employed for phosphopeptide sequencing. We have shown that a phenol-based method which was successful in case of the sequencing of hundreds of phosphopeptides in the yeast *Saccharomyces cerevisiae* [22], was not applicable to *Physcomitrella patens*. We cannot yet explain exactly the chemical

reactions that occur leading to the different modifications, due to which, however, the sequencing was not feasible.

Although the *de novo* sequencing of MS/MS spectra is time consuming, it is up to now the only reliable method applicable if only short enzymatic peptides are available as a source for protein identification. Finally, from the 253 sequences, a large part have been clearly identified and we report here, among these sequences, the identification of several kinases, and a transcription factor-like protein as well as one MAP kinase kinase. As the abundance of phosphorylated signalling molecules in cells is known to be low, the improvements in the resolution and sensitivity attained provides the critical information for embarking on investigation of phosphorylation in signalling pathways. Moreover, this technique goes beyond the field of a seedless plant and could potentially be used to map and compare phosphorylation in a variety of biological systems.

Financial support by the Deutsche Forschungsgemeinschaft (DFG Re 387/7 and DFG Sa 1014/1) is gratefully acknowledged.

Received October 7, 2003

5 References

[1] Zolnierowicz, S., Bollen, M., *EMBO J.* 2000, *19*, 483–488.

[2] Cohen, P., *Trends Biochem. Sci.* 2000, *25*, 596–601.

[3] Hunter, T., Plowman, G. D., *Trends Biochem. Sci.* 1997, *22*, 18–22.

[4] Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., *et al.*, *Science* 2000, *287*, 2185–2195.

[5] Cozzone, A. J., *Cell. Biochem.* 1993, *51*, 7–13.

[6] Holtorf, H., Guitton, M.-C., Reski, R., *Naturwissenschaften* 2002, *89*, 235–249.

[7] Strepp, R., Scholz, S., Kruse, S., Speth, V., Reski, R., *Proc. Natl. Acad. Sci. USA* 1998, *95*, 4368–4373.

[8] Reski, R., *Trends Plant Sci.* 1998, *3*, 209–210.

[9] Girke, T., Schmidt, H., Zähringer, U., Reski, R., Heinz, E., *Plant J.* 1998, *15*, 39–48.

[10] Thorneycroft, D., Sherson, S. M., Smith, S. M., *J. Exp. Bot.* 2001, *52*, 361, 1593–601.

[11] Nishiyama, T., Fujita, T., Shin-I, T., Seki, M. *et al.*, *Proc. Natl. Acad. Sci. USA* 2003, *100*, 8007–8012.

[12] Rensing, S. A., Rombauts, S., Van de Peer, Y., Reski, R., *Trends Plant Sci.* 2002, *7*, 535–538.

[13] Reski, R., *Bot. Acta* 1998, *111*, 1–15.

[14] Mason, G. G., Murray, R. Z., Pappin, D., Rivett, A. J., *FEBS Lett.* 1998, *430*, 269–274.

[15] Raska, C. S., Parker, C. E., Dominski, Z., Marzluff, W. F., *et al.*, *Anal. Chem.* 2002, *74*, 3429–3433.

[16] Steen, H., Kuster, B., Fernandez, M., Pandey, A., Mann, M., *Anal. Chem.* 2001, *73*, 1440–1448.

[17] Vener, A. V., Harms, A., Sussman, M. R., Viestra, R. D., *J. Biol. Chem.* 2001, *276*, 6959–6966.

[18] Sickmann, A., Meyer, H. E., *Proteomics* 2001, *1*, 200–206.

[19] Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., Aebersold, R., *Proc. Natl. Acad. Sci. USA* 2000, *97*, 9390–9395.

[20] Boyle, W. J., van der Geer, P., Hunter, T., *Methods Enzymol.* 1991, *201*, 110–149.

[21] Aebersold, R. H., Leavitt, J., Saavedra, R. A., Hood, L. E., Kent, S. B., *Proc. Natl. Acad. Sci. USA* 1987, *84*, 6970–6974.

[22] Ficarro, S. B., McCleland, M. L., Stukenberg, P. T.,Burke, D. J., *et al.*, *Nat. Biotechnol.* 2002, *19*, 301–305.

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

[24] Zhou, H., Watts, J., Aebersold, R., *Nat. Biotechnol.* 2001, *19*, 375–378.

[25] Kovarona, H., Hajduch, M., Livingstone, M., Dzubak, P., Lefkovits, I., *J. Chromatogr. B* 2003, *787*, 53–61.

[26] Kalume, D. E., Molina, H., Pandey, A., *Curr. Opin. Chem. Biol.* 2003, *7*, 64–69.

[27] Hohe, A., Reski, R., *Plant Sci.* 2002, *163*, 69–74.

[28] Ramagli, L. S., Rodriguez, L. V., *Electrophoresis* 1985, *6*, 559–563.

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

[30] Neubauer, G., Mann, M., *Anal. Chem.* 1999, *71*, 235–242.

[31] Hart, S. R., Waterfield, M. D., Burlingame, A. L., Cramer, R., *J. Am. Soc. Mass Spectrom.* 2002, *13*, 1042–1051.

[32] Larsen, M. R., Sørensen, G. L., Fey, S. J., Larsen, P. M., Roepstorff, P., *Proteomics* 2001, *1*, 223–238.

[33] Herbert, B. R., Sanchez, J.-C., Bini, L., in: Wilkins, M. R. (Ed.), *Proteome Research: New Frontiers in Functional Genomics*, Springer Verlag, Berlin 1997, pp. 13–33.

[34] Barrett, A. J. (Ed.), *Methods in Enzymology 244 – Proteolytic Enzymes: Serine and Cystein Peptidases*, Academic Press, San Diego, CA 1994, pp. 19–457.

[35] Zhou, W., Merrick, B. A., Khaledi, M. G., Tomer, K. B., *J. Am. Soc. Mass. Spectrom.* 2000, *11*, 273–282.

[36] Posewitz, M. C., Tempst, P., *Anal. Chem.* 1999. *71*, 2883–2892.

[37] Ahn, G., Resing, K. A., *Nat. Biotechnol.* 2001, *19*, 317–318.

[38] Neville, D. C. A., Rozanas, C. R., Price, E. M., Gruis, D. B., *et al.*, *Prot. Sci.* 1997, *6*, 2436–2445.

[39] http://www.chromatography.co.uk/Techniqs/Bio/bio_006.htm

[40] Grosmann, P. D., Colburn, J. C. (Eds.), *Capillary Electrophoresis: Theory and Practice*, Academic Press, San Diego, CA 1992.

[41] Kuhn, R., Hoffstetter-Kuhn, S., *Capillary Electrophoresis: Principles and Practice*, Springer-Verlag, Berlin/Heidelberg 1993.

[42] Loughrey Chen, S., Huddleston, M. J., Shou, W., Deshaies, R. J., *et al.*, *Mol. Cell. Proteomics* 2002, *1*, 186–196.

[43] Jongeneel, C. V., *Briefing in Bioinformatics* 2000, *1*, 76–92.

[44] Harmon, A. C., *Gravit Space Biol Bull.* 2003, *16*, 2, 83–90.

[45] Kobayashi, D., Tamoi, M., Iwaki, T., Shigeoka, S., Wadano A.,*Plant Cell Physiol.* 2003, *44*, 3, 269–276.

[46] Davis, R. J. in: Gutkind, J. S. (Ed.), *Signalling Networks and Cell Cycle Control: The Molecular Basis of Cancer and Other Diseases*, Humana Press, Totowa, NJ 2000, pp. 153–164.

[47] Ficarro, S., Chertinihin, O., Westbrook, A., White, F., *et al.*, *J. Biol. Chem.* 2003, *278*, 11579–11589.

[48] Mann, M., Ong, S.-E., Grønborg, M., Steen, H., *et al.*, *Trends Biotechnol.* 2002, *20*, 261–268.