Published in final edited form as:

Methods Enzymol. 2015 ; 551: 405–431. doi:10.1016/bs.mie.2014.10.022.

Sample Preparation for Phosphoproteomic Analysis of Circadian Time Series in *Arabidopsis thaliana*

Johanna Krahmer, Matthew M. Hindle, Sarah F. Martin, Thierry Le Bihan, and Andrew J. Millar¹

SynthSys and School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom

Abstract

Systems biological approaches to study the *Arabidopsis thaliana* circadian clock have mainly focused on transcriptomics while little is known about the proteome, and even less about posttranslational modifications. Evidence has emerged that posttranslational protein modifications, in particular phosphorylation, play an important role for the clock and its output.

Phosphoproteomics is the method of choice for a large-scale approach to gain more knowledge about rhythmic protein phosphorylation.

Recent plant phosphoproteomics publications have identified several thousand phosphopeptides. However, the methods used in these studies are very labor-intensive and therefore not suitable to apply to a well-replicated circadian time series. To address this issue, we present and compare different strategies for sample preparation for phosphoproteomics that are compatible with large numbers of samples. Methods are compared regarding number of identifications, variability of quantitation, and functional categorization. We focus on the type of detergent used for protein extraction as well as methods for its removal. We also test a simple two-fraction separation of the protein extract.

1. INTRODUCTION

Plant circadian transcript abundance has been extensively studied (Covington & Harmer, 2007; Edwards et al., 2006). By contrast, few studies have investigated rhythmicity of protein abundance with a systems biology approach (Robles & Mann, 2013).

Proteomic studies under diel or circadian conditions have been published for animal tissues: mouse liver (Mauvoisin et al., 2014; Reddy et al., 2006), mouse retina (Tsuji et al., 2007), and mouse SCN (Deery et al., 2009). With the exception of Mauvoisin et al. (2014), laborious 2D gels were used by these studies, which is becoming obsolete as a technique preceding mass spectrometry (Robles & Mann, 2013). Mauvoisin et al. (2014) used a SILAC approach and detected over 5000 proteins, about 200 of which were rhythmic in diurnal conditions. Baker, Kettenbach, Loros, Gerber, and Dunlap (2009) carried out circadian interaction proteomics and phosphoproteomics specifically for the clock protein FREQUENCY in *Neurospora crassa*. A recent circadian proteomics study in *Synechococcus*

elongatus (Guerreiro et al., 2014) achieved 82% coverage of the proteome (1537 proteins quantified) and detected 77 rhythmic proteins, using six-plex isobaric labeling and strong cation exchange (SCX) fractionation.

Regarding eukaryotic algae, Wagner, Fiedler, Markert, Hippler, and Mittag (2004) published a circadian proteomic time series in *Chlamydomonas reinhardtii*, and Akimoto, Wu, Kinumi, and Ohmiya (2004) studied diurnal proteomic changes in *Lingulodinium polyedrum*. Both studies used 2D-gels. In higher plants, Hwang et al. (2011) used 2D gels for a diurnal and circadian proteomic study in rice, finding approximately 3000 protein spots, 354 of which were rhythmic in diurnal conditions and 53 of these were rhythmic in constant darkness. Reiland et al. (2009) compared the chloroplast phosphoproteome at the end of day and end of night time points in a diurnal cycle. Their protocol involved SCX fractionation and two partially complementary phosphopeptide enrichment steps, yielding 16 fractions per sample.

A common finding of circadian proteomic studies is that the fraction of rhythmic proteins is lower than the fraction of rhythmic transcripts; many rhythmic transcripts do not have rhythmic proteins and vice versa (Hwang et al., 2011; Mauvoisin et al., 2014; Reiland et al., 2009; Wagner & Mittag, 2009). When RNA and protein both cycle, peak phases are often several hours apart (Mauvoisin et al., 2014). Where these have been measured, arrhythmicity of protein abundance is most often due to long protein half-lives (Mauvoisin et al., 2014).

One hypothesis created from this discrepancy between transcript and protein rhythmicity is that the total amount of a protein is not as physiologically relevant as the abundance of its posttranslational modification state at a given time of day. Phosphorylation plays a role in the plant circadian clock as well as its output (for a recent review, see Kusakina & Dodd, 2012). Phosphorylation of plant circadian clock proteins has been demonstrated by Western blotting (e.g., Farré & Kay, 2007; Sugano, Andronis, Ong, Green, & Tobin, 1999), but the pervasiveness of rhythmic phosphorylation at the *Arabidopsis* proteome scale is unknown. A circadian phosphoproteomic time series would address this question but this requires a streamlined experimental protocol.

We present an optimized sample preparation for phosphoproteomics in *Arabidopsis* that can be applied to a well-replicated time series of about 30 samples. Plant studies that yielded a large number of phosphopeptides employed techniques that are too time-consuming to be applied to 30 samples (Reiland et al. (2009): over 3000 phosphopeptides; Wang et al. (2013): over 5000; Facette, Shen, Björnsdóttir, Briggs, and Smith (2013): over 11,000).

Quantitation is obviously desirable to detect rhythmic regulation and can be facilitated by isotopic labeling. Metabolic labeling is difficult to achieve in plants due to inefficient isotope incorporation and its high cost for large numbers of samples. The latter is also the case for chemical labeling (Slade, Werth, Chao, & Hicks, 2014). Therefore, we limited our quantitation strategy to label-free quantitation.

Since only a small fraction of peptides is phosphorylated, it is difficult to identify phosphopeptides in total protein samples due to ionic suppression (Mann et al., 2002). It is

therefore necessary to enrich phosphopeptides. Popular strategies are metal oxide affinity chromatography (MOAC) and immobilized metal affinity chromatography. TiO₂ beads are commonly used in a MOAC approach. Negatively charged phosphopeptides bind to positive charges on the surface of the TiO₂ particles (Dunn, Reid, & Bruening, 2010). We used commercial TiO₂ spin columns for all our analyses.

The scope of this chapter is to find a strategy for phosphoproteomic sample preparation that balances the number of phosphopeptide identifications, variability, and workload. We test different extraction buffer compositions and detergent removal methods, including an acid-labile detergent and separation of the protein extract into two fractions. All our workflows allow us to take digest aliquots before phosphoenrichment. We therefore present the analysis of global peptide digests along with the phosphoproteomic results.

2. MATERIALS AND METHODS

2.1. Plant material

Seedlings of *Arabidopsis thaliana*, ecotype Columbia-0 were grown on plates (2.15 g/l Murashige & Skoog medium Basal Salt Mixture (Duchefa Biochemie), pH 5.8 (adjusted with KOH) and 12 g/l agar (Sigma)) for 10 days followed by 3 weeks on soil. Growth conditions were 12 h light and 12 h dark at 22 °C and a fluence rate of 120 µmol/s/m² (cool-white fluorescent tubes). Rosettes without roots and shoots were snap-frozen in liquid nitrogen. For sample OG trichloroacetate (TCA), rosettes were grown in the same way but harvested 1 week earlier.

2.2. Protein extraction for buffer optimization and the RapiGest™ SF experiment

Frozen plant material was ground using a dry ice cooled mortar and pestle. About 300 mg of plant material was used per replicate. 500 µl of extraction buffer of the following composition was added: for the detergent-free extraction, the extraction buffer contained 8 M urea, 100 mM Tris-HCl pH 7.5, 75 mM NaCl, 10 mM EDTA, 0.5% (v/v) polyvinylpyrrolidone, cross-linked (Sigma; CAS 25249-54-1), 1× PhosStop (Roche), and 1× Complete Protease inhibitor cocktail EDTA free (Roche).

For *n*-octyl-β-D-glucoside (OG) containing samples (refer to OG ethyl acetate, OG TCA MetOH, and OG TCA), 1% (w/v) OG was added. For sodium dodecyl sulfate (SDS) containing samples (SDS TCA, SDS ethyl acetate), 1% SDS (v/v) was added; for sample IGEPAL TCA, 1% (v/v) IGEPAL CA-630 (Sigma) was added. For the RapiGest™ SF method and its control at pH 8.5, the pH of Tris-HCl was adjusted to 8.5. 0.1% (v/v) RapiGest™ SF (Waters Corp.) was added for the RapiGest™ SF extracted samples.

All samples were homogenized on a Qiagen Tissue Lyzer with a metal ball of 3 mm diameter for 1 min at 30 Hz. After incubation on ice for 15 min, extracts were cleared at 20,000×g and 4 °C for 10 min twice. Protein was quantified using the Quick Start™ Bradford Protein Assay (Bio-Rad) or the SDS compatible DC™ protein assay (Bio-Rad) for SDS-containing samples.

2.3. Fractionation with polyethylene glycol

A streamlined version of a protocol published by Aryal, Krochko, and Ross (2012) was used to deplete a protein extract fraction of RuBisCO by centrifugation. In contrast to the original protocol, only one centrifugation with polyethylene glycol (PEG), at 15%, was carried out. 10 ml of Mg/Tx-100 buffer (0.5 M Tris-HCl pH 8.3, 0.5% polyvinylpolypyrrolidone, 2% (v/v) Triton X-100, 20 mM MgCl₂, 2% (v/v) β-mercaptoethanol, 1× Complete Protease Inhibitor Cocktail EDTA Free (Roche), 1× PhosStop (Roche)) were added to 2 g of crushed plant material. Samples were incubated on ice for 1 h and centrifuged at 3220×g for 20 min at 4 °C to remove cell debris. The supernatants were transferred to a fresh tube. 1.5 g PEG powder (average MW 3500, Sigma) was added to the tube. An unfractionated control sample was carried out in parallel without PEG and no second centrifugation. The samples for fractionation were incubated on ice for 30 min and centrifuged at 3220×g for 20 min. The supernatant was transferred to a fresh tube. The pellet was resuspended in 2.5 ml Mg/Tx-100 buffer.

2.4. Protein precipitation by TCA/acetone

Protein of all extracts except for the detergent-free protocols, the RapiGest™ SF samples plus their controls and the samples for ethyl acetate-based detergent removal were precipitated by a TCA and acetone-based method.

H₂O was added to the extract containing 2 mg of protein to a volume of 1250 µl. 310 µl of 100% TCA solution, 1248 µl of methanol, and 624 µl of chloroform were added. Samples were mixed and precipitated at –20 °C overnight. Centrifugation at 3220×g for 5 min at 4 °C resulted in an upper clear phase and a lower green phase. The upper phase was discarded; the green phase and interface were kept. 1 ml methanol was added; samples were mixed and centrifuged for 1 min at 3220×g at 4 °C. The supernatant was discarded and the resulting protein pellet was reconstituted in 1 ml acetone, centrifuged 1 min at 3220×g and 4 °C, repeated three times. The resulting protein pellets were finally dissolved in 8 M urea with the exception of peptides for method OG TCA MetOH, which were resuspended in methanol. Protein concentration was measured with a Bradford assay, except for the OG TCA MetOH.

2.5. Tryptic digest

500 µg of protein was subjected to protease digestion with trypsin as follows. For OG TCA MetOH, the volume required for 500 µg was estimated based on our experience that on average 60% of protein is lost in the TCA/acetone precipitation of OG extracted tissue. Volumes were adjusted to 250 µl with 8 M urea, or methanol in the case of OG TCA MetOH. 25 µl of 1 M ammonium bicarbonate and 25 µl of 200 mM DTT were added and samples were incubated to reduce disulfide bonds for 30 min at room temperature. 25 µl of 500 mM iodoacetamide were added to carbamidomethylate thiol groups followed by incubation at room temperature and in the dark for 1 h. 665 µl distilled water were added and 10 µl of 1 mg/ml trypsin (Worthington). Samples were digested at room temperature overnight.

2.6. Detergent removal by ethyl acetate for sample OG ethyl acetate and SDS ethyl acetate

The digest volume was reduced to 100 µl in a Speed-vac. A method described by Yeung and Stanley (2010) was followed which uses ethyl acetate to remove detergents.

2.7. Removal of RapiGest™ SF by acidification

For hydrolysis of RapiGest™ SF, 50 µl of 20% formic acid was added to digested samples “RapiGest” and “No RapiGest pH 8.5”, which lowered the pH below 2. Samples were incubated at 37 °C for 45 min and centrifuged at 13,000×g at room temperature for 10 min. The supernatant was transferred to a fresh tube.

2.8. Cleanup of digests

After digest or detergent removal by ethyl acetate, peptides were desalted on a reverse phase resin using BondElute columns (25MG columns, Agilent): 1 ml methanol and two times 1 ml HPLC grade water (Thermo-Fisher) were sequentially passed through the column before loading the sample. After binding of peptides, columns were washed using 1 ml of HPLC grade water and peptides were eluted in 1 ml acetonitrile. Sample eluates were split into 490 µg (for phosphopeptide enrichment) and 10 µg (for global peptide analysis). Both aliquots were vacuum dried in a Speed-vac (RC1010, Thermo).

2.9. Phosphopeptide enrichment

Preliminary tests showed that the Titanshpere™ spin tip kit (GL Sciences Inc.) worked best in our hands (data not shown), so we consistently used this method for our experiments. Therefore, all steps from the phosphopeptide enrichment step were identical for all methods.

The following solutions were prepared: sln0 (2.5% acetonitrile, 0.5% trifluoroacetic acid (TFA)), sln1 (80% acetonitrile, 0.5% TFA), sln2 (9×sln1+1× Solution B of the enrichment kit which is 100% lactic acid), sln3 (5% (v/v) ammonium hydroxide solution, prepared just before use), sln4 (5% (v/v) pyrrolidine solution). Dried peptides were sequentially dissolved and sonicated in 50 µl of sln0 and 100 µl of sln2 and centrifuged at 20,000×g for 5 min at 22 °C. The supernatant was recovered for enrichment. Spin adapters were inserted into a waste tube (2 ml Eppendorf tube without lid) and spin tips were inserted. Centrifugation of spin columns was always done at 200×g at 22 °C. Spin tips were washed twice with 20 µl of sln1 for 2 min and waste tubes were exchanged. The sample was loaded and spin tips were centrifuged for 5 min. After incubation for 15 min, spin tips were spun again for 20 min to pass the entire sample through. The flow through was reloaded and the two spins with 15 min incubation were repeated. The reloading and centrifugation steps with incubation were repeated once more. Spin tips were washed with 20 µl sln1, twice with 20 µl sln2, and again twice with 20 µl sln1. For elution, the waste tube was exchanged for a clean 1.5-ml microfuge tube. Peptides were first eluted with 50 µl of sln3 for 5 min and then with 50 µl of sln4 for 8 min. Immediately, 20 µl 20% formic acid were added to lower the pH to about 3. Phosphopeptides were desalted on reverse phase resin using Bond Elut Omix tips (Agilent). The Omix tip was washed with 130 µl (all volumes were 130 µl except for elution) Buffer B (80% acetonitrile, 0.1% TFA) twice and then with buffer A (0.1% TFA in distilled water) twice. Peptides were retained on the tip by pipetting up and down 10 times, followed by two

washes with buffer A and elution with 50 µl of buffer B. Samples were dried on a Speed-vac.

2.10. Mass spectrometry

Dried peptides were dissolved in 12 µl 0.05% TFA for phosphopeptide analysis, and 20 µl for global peptide analysis. Peptides were passed through Millex-LH 0.45 µm (Millipore) filters. 8 µl was run on an on-line capillary-HPLC-MSMS system consisting of a micropump (1200 binary HPLC system, Agilent, UK) coupled to a hybrid LTQ-Orbitrap XL instrument (Thermo-Fisher, UK). Reverse phase buffer used for LC-MS separation was buffer A (2.5% acetonitrile, 0.1% FA in H₂O) and buffer B (10% H₂O, 90% acetonitrile, 0.1% formic acid, 0.025% TFA). LC peptide separation was carried out on an initial 80 min long linear gradient from 0% to 35% buffer B, then a steeper gradient up to 98% buffer B over a period of 20 min then remaining constant at 98% buffer B (for 15 min) until a quick drop to 0% buffer B before the end of the run at 120 min. The tair Arabidopsis_1rep database was used for data-dependent detection.

2.11. Data analysis

For determination of the number of (phospho)peptides detected in each replicate, each raw file was converted to an mgf file and peptide searches were carried out with the following parameters: database Arabidopsis_1rep, trypsin as enzyme, allowing up to two missed cleavages, carbamidomethyl (C) as a fixed modification, Oxidation (M), Phospho (ST) and Phospho (Y), Acetyl(Protein N-term) as variable modifications, a peptide tolerance of 7 ppm, and MS/MS tolerance of 0.4 Da, peptide charges 2+, 3+, and 4+, on an ESI-trap instrument, and with decoy search. For export, an ion-cutoff of 20 was chosen. To determine the number of phosphopeptides for Figs. 1C, 2D, and 3C, the number of peptide entries containing phosphorylated residues was counted as well as the number of total peptides.

The Progenesis LC MS software was used for label-free quantitation. Raw files were imported into a label-free analysis experiment. Chromatograms were aligned by adding up to four manual vectors where peaks were misaligned, before starting automatic alignment and peak picking. Only charges 2+, 3+, and 4+ and data from 6 to 100 min of the runs were chosen for analysis. The exported file of MS/MS spectra was uploaded on the Mascot website and a search was carried out as described above for the single run searches. Protein and peptide measurements were exported and statistically analyzed. For phosphopeptide-enriched samples, peptide measurements were exported as .csv file and then saved as an .xlsx file for peptide merging (Hindle et al., in review).

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (Vizcaíno et al., 2014) via the PRIDE partner repository with the dataset identifier PXD001134 and DOI 10.6019/PXD001134.

For a measure of variability of each method, the coefficient of variation (CV) was calculated from the arcsinh transformed data of all replicates for each protein from the Progenesis output file or the merged phosphopeptide file for phosphopeptides. For visualization, boxplots of these CV values are presented (Figs. 1D, 2E, and 3D–E).

2.12. Gene ontology enrichment analysis

Gene ontology (GO) term enrichment analysis was applied to both pairwise comparisons (“No Detergent” vs. “IGEPAL TCA” and “RapiGest” vs. “No RapiGest pH 8.5”) in order to compare overrepresentation of functions and processes. Significance difference of the means in each pairwise comparison was determined by homoscedastic *t*-tests on an arcsinh transform of the normalized abundance as quantified in Progenesis. For each pairwise comparison, we tested enrichment of GO terms within protein entries based on the *p*-values for significance differences in the mean abundances. For the phosphoproteomics data, the phosphopeptide quantification with the lowest *p*-value was used if more than one phosphopeptide of the same protein were found.

The annotation of GO terms to *Arabidopsis* proteins was obtained from <http://www.geneontology.org/GO.downloads.annotations.shtml> (submitted: 10 February 2014). We used RStudio and the topGO library of the bioconductor package (<http://bioconductor.org/biocLite.R>, version: 2.16.0; Alexa & Rahnenführer, 2010). The Kolmogorov–Smirnov (KS) test was used to evaluate the significance of enrichment for each GO term. The KS statistic was applied with the weight01 algorithm to account for the topology of the GO tree and is a combination of a weighted and elim methods (Alexa, Rahnenführer, & Lengauer, 2006). All enrichment *p*-values of <0.05 were taken as significantly enriched terms.

3. RESULTS

3.1. Choice of extraction buffer and detergent removal method affects number of detected proteins and phosphopeptides

Protein extracts of mature *Arabidopsis* rosettes were prepared using four extraction buffers and two detergent removal strategies as summarized in Fig. 1A. A urea- and Tris-HCl-based buffer (see Section 2) without detergent provided a reference that avoided any detergent removal (“No Detergent”; Fig. 1A). Proteins were digested with trypsin following a standard in-solution protocol and phosphopeptides were enriched using a titanium dioxide-based kit. The phosphoenrichment and downstream procedures were the same for all methods.

Since detergents increase the amount of extracted protein, the detergents IGEPAL, octyl glucoside (OG), or SDS were added to this buffer. The amount of extracted protein increased more than twofold for the same amount of starting material compared to “No Detergent” (data not shown); however, the detergent needs to be removed before mass spectrometry. We took two different approaches. First, we precipitated and washed proteins using a TCA/acetone-based method and resuspended in urea before trypsination, resulting in the samples “IGEPAL TCA,” “OG TCA,” and “SDS TCA” (Fig. 1A). In the case of OG, we also tested resuspending the protein pellet in methanol instead of urea since this may increase digestion efficiency (Russell, Park, & Russell, 2001), resulting in sample “OG TCA MeOH.” The loss of protein in this step was approximately 50%. In the second approach, we digested in the presence of the detergent (OG or SDS) and removed it thereafter using an

ethyl acetate-based partitioning, which produced samples “OG ethyl acetate” and “SDS ethyl acetate” (Yeung & Stanley, 2010).

In addition to the phosphoenriched peptides, an aliquot of each non-enriched protein digest was also analyzed on the mass spectrometer. Non-enriched samples from extractions using SDS were not tested to avoid problems due to remaining detergent traces.

Figure 1B shows the average number of proteins detected in the global peptide analysis. Addition of detergent and precipitation by TCA/acetone yielded significantly more protein identifications compared to “No Detergent.” OG removal by ethyl acetate did not perform as well as the other methods. There was no significant difference in protein numbers between the three TCA/acetone extraction-based methods.

The situation was similar for the number of phosphopeptides (Fig. 1C): addition of detergent increased the number of identified phosphopeptides significantly if the detergent was removed by TCA/acetone precipitation. The ethyl acetate strategy did not result in as many phosphopeptides identified as the TCA/acetone-based methods. In both SDS extraction methods, the number of peptides was significantly lower than in “No Detergent.” In two of the three “SDS TCA” replicates, no phosphopeptides were identified. The percentage of phosphopeptides in the total number of peptides identified was not significantly different between OG TCA MetOH, OG TCA, and IGEPAL TCA.

Since SDS TCA, SDS ethyl acetate, and OG ethyl acetate did not perform as well as the other methods, we focused on the others in further analyses.

The median CV of protein abundance in the global protein analysis was between about 0.02 and 0.05, with IGEPAL TCA being the lowest (Fig. 1D). For phosphopeptide analysis, the range of the median values was larger (between 0.02 and 0.07). OG TCA MetOH had the largest median and OG TCA the smallest.

To indicate how these methods overlap regarding the identity of detected phosphopeptides, we measured the number of phosphopeptides that were identified in all replicates of one method and in none of the other methods (Fig. 1E). For each detergent extraction, the number of specific phosphopeptides is higher when comparing to the “No Detergent” samples than when comparing to other detergent-based methods, suggesting that detergents extracted a characteristic set of additional proteins. Method “OG TCA” has more specific phosphopeptides than the other methods.

The IGEPAL TCA method and the no detergent control were compared in a GO analysis (Table 1). Fewer terms were enriched in the phosphopeptide analysis than in the global protein analysis and the IGEPAL TCA method generally enriched more terms than the no detergent control. Although terms enriched by either method are from a broad spectrum, some words consistently appeared in only one of the methods: the majority of IGEPAL TCA enriched terms for cellular component were related to photosynthesis or plastids or membranes, while No Detergent was enriched for many ribosome-related terms in the global protein analysis. Membrane-related terms were also enriched by IGEPAL TCA in the phosphopeptide analysis.

3.2. Fractionation with PEG does not increase numbers of identified peptides

Fractionation can significantly improve the number of identified peptides by reducing the complexity of the sample, thus allowing the MS instrument to identify a higher proportion of the peptides. However, it also multiplies the number of samples to be processed. This is problematic for circadian time series, which already have a large number of samples. We explored a method that produces only two fractions per sample by a centrifugation step with PEG. This pellets denser proteins, in particular the RuBisCO subunits, and thereby is expected to reduce the range of protein abundance in the supernatant fraction (Aryal et al., 2012).

As shown in Fig. 2A, we extracted with a magnesium and Triton X-100 containing buffer called “Mg/Tx-100 buffer” (Aryal et al., 2012). After incubating the protein extract with PEG, centrifugation resulted in a pellet and a supernatant. The pellet was resuspended in Mg/Tx-100 buffer, and both were TCA precipitated, digested, and phosphoenriched.

As a control, we conducted one extraction in the same way but without the addition of PEG, as an unfractionated control. To control for the fact that the extraction buffer could make a difference compared to the other methods we have used, we also extracted using the method “IGEPAL TCA” (see Fig. 1A) in parallel. All extractions used tissue from the same pool of ground tissue. The number of replicates was three for “IGEPAL TCA”, five for “Unfractionated,” and “Pellet,” four for “Supernatant.”

The fractionation reduced the amount of RBCL in the supernatant 15-fold (Fig. 2B bottom). The amount of the clock output protein CCR1 is significantly higher in the supernatant fraction than the pellet (Fig. 2B middle). The only clock protein of which a phosphopeptide was detected in the phosphoenriched sample was TIME FOR COFFEE (TIC, Fig. 2B top). We did not find a significant difference in its abundance between the two fractions.

There was no statistically significant difference in the average number of quantified proteins between any of the four methods (Fig. 2C).

There is no statistically significant difference in the average number of identified phosphopeptides between the different workflows and fractions, which shows that the Mg/Tx-100 buffer is generally suitable for protein extraction but the PEG treatment did not increase the number of identified phosphopeptides in the supernatant sample even though some abundant proteins were present at lower levels.

Therefore, it is not beneficial to use only the supernatant sample instead of an unfractionated sample. However, we wanted to find out if the number of phosphopeptides increased when both fractions are analyzed. Due to the quantification procedure in Progenesis, increasing the number of samples even if they are replicates from the same extraction method will increase the number of detected peptides. Therefore, we determined the number of phosphopeptides found in combinations of two runs of the same treatment or of one pellet and one supernatant analysis. Figure 2F shows the average number of phosphopeptides for all possible pairs. The highest average was found in pairs of just the pellet fraction. The average number for a supernatant pair was significantly lower than for a pellet pair

($p=0.012$, t -test). The average number for a supernatant and pellet sample pair was not significantly different from the pellet pair or the unfractionated pair, but a pellet and supernatant pair had significantly more phosphopeptides than two supernatant samples ($p=0.020$, t -test).

With respect to variability of the phosphopeptide analysis, the box plots in Fig. 2E (bottom) show that the supernatant sample has a higher median CV than the other methods. This is not the case for the median CV of protein abundance calculated from the global peptide analysis, where the CV of the IGEPAL TCA control has the highest median.

3.3. The acid-labile detergent RapiGest™ does not increase the number of detected phosphopeptides

As detergents extracted additional proteins and phosphopeptides but required time-consuming desalting steps (Fig. 1), we tested whether detergent removal could be avoided using an acid-labile detergent. We added RapiGest™ SF surfactant to the same buffer as used in the “No Detergent” sample in Fig. 1A, except the pH was adjusted to 8.5 to make sure RapiGest™ is not hydrolyzed prematurely.

We conducted a control without RapiGest™ but still at pH 8.5. The control was also subjected to acidification after digest (workflow see Fig. 3A). To rule out the possibility that a higher pH and acidification after digest affects performance, we also carried out an extraction in the same way as “No Detergent” in Fig. 1A in parallel and from the same tissue powder as the two other methods.

Figure 3B and C shows that neither the number of proteins in the global peptide analysis nor the number of phosphopeptides identified in the phosphoenriched samples is significantly different between any of the three treatments. The median CV is slightly higher for the RapiGest™ samples in the phosphoenriched samples (Fig. 3E) but not in the protein quantification from the global peptide analysis (Fig. 3D).

In a GO analysis of the global protein data, photosynthesis and membrane-related terms were enriched in the RapiGest™ SF data compared to the control at pH 8.5. In the latter, ribosome- and translation-related terms were enriched.

4. DISCUSSION

When developing or choosing a sample preparation method for enriching a specific class of peptides (in this case phosphopeptides), several factors have to be considered: (1) the number of peptides identified (phosphorylated peptides in this case) (2) if quantitation is needed then the variability between replicates has to be low, and (3) the capability to enrich for interesting peptides. One major challenge in proteomics is its bias toward highly abundant species often masking more interesting lower abundant proteins. A common approach to tackle this issue is to fractionate a sample into several subfractions. Although this approach is appropriate for few samples to analyze, it can be quite a tedious one for more complex experimental designs, such as different drug treatments or genotypes, or circadian time series. Extensive fractionation is not a viable option in the case of circadian

phosphoproteomic time series. Methods employed by published studies that achieved high phosphoproteome coverage, such as Reiland et al. (2009) and Wang et al. (2013), are too laborious for a well-replicated circadian time series.

For a circadian (phospho)proteomics experiment, a high number of identifications as well as reliable quantification are required across several time points.

Sample preparation for quantitative phosphoproteomic experiments generally requires long protocols to be carried out accurate and carefully. In addition, there are plant-specific challenges: the presence of cell walls makes cells more resilient to protein extraction; protein content is typically lower than in animal tissues so more material is required; plant cells contain interfering compounds such as polyphenols and a range of secondary metabolites; the enzyme RuBisCO amounts to 30–60% of cellular protein, making less abundant peptides more difficult to be detected (Koroleva & Bindschedler, 2011; Slade et al., 2014).

Our aim was to compare the performance of several sample preparation methods for proteomics and phosphoproteomics, which are compatible with a well-replicated circadian time series experimental design.

4.1. Extraction with a nonionic detergent and precipitation with TCA/acetone outperforms other strategies

We demonstrate a comparison of seven different extraction methods. A detergent-free buffer is appealing since time-consuming detergent removal introduces additional steps, which is likely to increase the technical variation. In this work, we show that reliable proteomics and phosphoproteomics can be done without detergent. However, regarding the number of proteins and phosphopeptides identified, the detergent-free method yields significantly fewer phosphopeptides than methods using a nonionic detergent (IGEPAL or OG) in the extraction buffer and precipitation of protein with a TCA/acetone-based protocol.

We chose to test the strong ionic detergent SDS and the nonionic detergents IGEPAL and OG. IGEPAL (also known as NP-40) is widely used for protein extraction. OG was found to be superior to other nonionic detergents by Arachea et al. (2012). In our hands, buffers containing IGEPAL, OG and SDS extracted more than twice as much protein as the detergent-free buffer from a given amount of tissue (data not shown). However, it should be taken into account that about 50% of protein material is lost during the TCA/acetone precipitation, so roughly equal amounts of starting material are required.

In contrast to our expectations we did not observe higher variability with the additional step for detergent removal by TCA/acetone. The exception is method “OG TCA MetOH,” which resulted in phosphopeptide quantification with significantly more variability than in the other methods (Fig. 1D). We had expected this since resuspension with methanol is not compatible with the standard quantitation methods, so the appropriate amount of protein for digestion was estimated according to typical losses experienced in the OG TCA method. According to Chen, Cociorva, Norris, and Yates (2008), trypsinization efficiency is higher in mixed organic-aqueous systems compared to aqueous buffers. Our data do not support this

since we did not find a significant difference in phosphopeptide or protein numbers between OG TCA and OG TCA MeOH. However, it is possible that factors other than digestion are limiting the number of identifications in our case, making increased digestion efficiency unnoticeable.

An additional benefit of the precipitation step is removal of nonprotein contaminants from either the plant cell extract or buffer components, which allows better control of the sample composition and compatibility with the electrospray ionization. We have experienced contamination with polymers in detergent-free samples but not in TCA/acetone-precipitated samples.

Regarding the choice of detergent, the differences in global protein and phosphopeptide numbers identified in OG TCA and IGEPAL TCA are not significant. IGEPAL is cheaper and may therefore be the preferred choice. The ionic detergent SDS did not prove useful in our hands since the number of phosphopeptides identified was significantly lower than the detergent-free procedure (Fig. 1C). Possible reasons include incomplete removal and therefore potential incompatibility with the phosphoenrichment strategy or the mass spectrometric analysis.

Presence of detergents can also increase the efficiency of tryptic digestion (Chen et al., 2008). Therefore we tested a strategy of extracting with detergent and removing it after digest using ethyl acetate (Yeung & Stanley, 2010). The particularity of this method is that the detergent is removed at the peptide level, not the protein level. When using OG as a detergent, the method resulted in approximately as many global protein and phosphopeptide identifications as the detergent-free method but significantly fewer than the OG extracted and TCA/acetone-precipitated samples. Therefore, extraction with OG and its removal by ethyl acetate does not have any advantages over the simpler detergent-free extraction.

The use of acid-labile detergents such as RapiGestTM SF is a very attractive way to avoid the time-consuming detergent removal process, while benefitting from the ability of the detergent to solubilize membrane proteins and increase digestion efficiency (Chen et al., 2008). After digest, the detergent is hydrolyzed by acidification, resulting in a soluble and an insoluble product. The latter is removed by centrifugation.

We expected an increase in the number of identifications compared to a control that was subjected to the same procedure but without the addition of RapiGestTM SF. However, we found no statistically significant differences in the number of identifications in either the global protein or phosphopeptide analyses (Fig. 3B and C) and the variability in the phosphoproteomics data of the RapiGestTM SF treated samples was significantly higher than that of the other methods. Therefore, with respect to numbers and variability, there is no benefit from investing in this expensive detergent.

Our GO analysis suggests that RapiGestTM SF at 0.1% succeeds in extracting proteins typical of detergent-containing buffers. Both IGEPAL TCA and RapiGestTM SF analyses enriched terms related to photosynthesis and membranes, compared to their respective controls.

4.2. The nonionic detergent IGEPAL extracts more membrane- and chloroplast-related proteins

Our method specificity analysis (Fig. 1E) suggests that with methods based on a nonionic detergent and TCA/acetone precipitation, a set of phosphoproteins is identified that was not found when extracting without detergent. According to our GO analysis, more photosynthesis/chloroplast related as well as membrane-related proteins are extracted by the IGEPAL TCA method than the detergent-free control, while the detergent-free approach should be preferred if ribosome-related proteins are of particular interest. A likely explanation is that the detergent solubilizes both the cell membrane and the plastid membranes, releasing the proteins inside the organelle. Abundant cytoplasmic proteins such as ribosome components will thereby be diluted. An enrichment of transmembrane proteins by the IGEPAL TCA method is also supported by an analysis using the TMHMM Server (v. 2.0; <http://www.cbs.dtu.dk/services/TMHMM/>; Krogh, Larsson, von Heijne, & Sonnhammer, 2001), which predicted significantly more transmembrane proteins in the IGEPAL TCA global protein results than the no detergent results ($p=0028$, χ^2 test of independence).

Phosphopeptides from proteins related to the circadian clock or phosphorylation that were consistently found in the IGEPAL extracted samples but not the “No Detergent” method were, for example, from MAP kinase 2, calcium-dependent kinase 19, casein kinase 1-like protein 2, BR-signaling kinase 1, and trehalose phosphatase/synthase 7.

The list of phosphopeptides that are likely to be found with the detergent-free method but not the IGEPAL TCA methods included three transcription factors. Lists of proteins and peptides found by each method will be available on the PRIDE database upon publication.

4.3. Fractionation by density using PEG is not superior to increasing replicate number

Mass spectrometric analysis of multiple sample fractions can be tedious for well-replicated circadian time series. We tested a separation into two fractions, which doubles the number of samples but only adds about one more working day to the procedure (due to doubling the number of digests to be phosphoenriched). Aryal et al. (2012) developed a fractionation procedure based on several centrifugation steps with PEG at different concentrations, which results in a pellet and a supernatant in each centrifugation step. They demonstrated that in the supernatant fraction, the RuBisCO subunits are depleted and 23% more peptides were identified than in an unfractionated sample. We streamlined this procedure (see Section 2) to one centrifugation, yielding two fractions, and precipitated with TCA/acetone instead of phenol/chloroform. We confirmed that the RuBisCO subunits are pelleted by PEG (Fig. 2B; small subunit not shown). Regarding clock-related proteins, the supernatant enriches proteins of CCR1 (Fig. 2B center) and PATHOGEN AND CIRCADIAN CONTROLLED 1 (data not shown), so if these particular proteins are of interest, the supernatant fraction is useful. There is no significant difference in the abundance of a TIC phosphopeptide between the pellet and supernatant fractions (Fig. 2B bottom).

Using the fractionation approach, we did not identify more proteins or phosphopeptides in the supernatant sample compared to the unfractionated control or the pellet. Therefore, our

effort to deplete the most abundant protein(s) did not increase protein identifications significantly. Moreover, this additional step increases the variability of quantification (Fig. 2E).

Although the number of identified peptides can be quite low per sample run on LC-MS, merging the different replicates and experiment significantly increases the number of identified peptides. We therefore asked whether we can increase the number of identifications by analyzing both fractions, rather than analyzing two replicates of an unfractionated sample. Assuming that two different fractions of a same sample analyzed by LC-MS and the identification output combined will likely result in more identifications than one, we compared pairs of runs of the same method with supernatant-pellet pairs. The pellet-supernatant combination performs better than a supernatant-pair and a pellet pair performs better than a supernatant pair, but combining one of each fraction does not perform better than two unfractionated replicates.

We conclude that analysis of both protein fractions does not result in more identifications than analysis of two different replicates. Workload may be smaller for fractionation than for more replication. However, we predict that this fractionation does not justify carrying out half the number of replicates because data analysis will become more unreliable for fractions compared to the same number of replicates.

Therefore, we recommend a reasonable number of replicates rather than fractionation with PEG, unless a subset of proteins enriched by one of the fractions is of particular interest.

4.4. Alternative strategies

There are other conceivable strategies for acceptable (phospho)proteomic analysis of a circadian time series. General strategies for proteomics in plants have been reviewed by Koroleva and Bindschedler (2011), and phosphoproteomics methods by Slade et al. (2014).

A simple fractionation approach is removal of chloroplasts, which will deplete the remaining fraction of several highly abundant proteins (Hiltbrunner et al., 2001). However, this requires careful and consistent homogenization without rupturing chloroplasts prior to separation. This in turn requires specialized equipment and is time consuming, making it less attractive for a large set of samples.

Plant samples can be depleted of RuBisCO using commercial antibody-based strategies (Cellar et al., 2008). The cost of these bead-coupled anti-bodies is unattractive for circadian time series. The same applies to equalizer beads (Righetti, Boschetti, Lomas, & Citterio, 2006), which can be used to reduce the range of protein abundance independent of protein identity. One potential issue with flattening of the dynamic range is reducing the sensitivity of relative protein quantitation.

The FASP principle, in which washing and digestion take place on a filter (Wisniewski, Zougman, Nagaraj, & Mann, 2009), is appealing to conveniently remove detergent within the digestion process, however the success of this method seems to depend very much on the specific experiment conducted with possible loss for low abundant proteins.

In-gel digestion is very commonly used in proteomics, is efficient at removing detergent, and offers the possibility of fractionating. However, the extract volumes required here exceed gel well volumes by far. Using multiple wells and several gel slices per well will lead to longer working time.

A strategy involving long LC separation during analysis could represent a possibility to increase the number of identifications.

5. CONCLUSIONS

We conclude that extraction with a nonionic detergent such as OG or IGEPAL and precipitation with TCA/acetone without fractionation is a successful, reliable, and feasible strategy for conducting phosphoproteomic time series. Time-consuming fractionation or expensive acid-labile detergents are not beneficial. We expect that these results can be transferred to proteomic analysis of other posttranslational modifications.

ACKNOWLEDGMENTS

We thank Lisa Imrie and Katalin Kis for expert technical support.

Funding: This work was supported by the Wellcome Trust [096995/Z/11/Z] and BBSRC, and EPSRC awards [BB/D019621] and [BB/J009423].

ABBREVIATIONS

OG	<i>n</i> -octyl- β -D-glucoside
PEG	polyethylene glycol
SDS	sodium dodecyl sulfate
TCA	trichloroacetate

REFERENCES

- Akimoto H, Wu C, Kinumi T, Ohmiya Y. Biological rhythmicity in expressed proteins of the marine dinoflagellate *Lingulodinium polyedrum* demonstrated by chronological proteomics. *Biochemical and Biophysical Research Communications*. 2004; 315:306–312. [PubMed: 14766208]
- Alexa, A.; Rahnenführer, J. topGO: Enrichment analysis for Gene Ontology. R package version 2.16.0. 2010.
- Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*. 2006; 22:1600–1607. [PubMed: 16606683]
- Arachea BT, Sun Z, Potente N, Malik R, Isailovic D, Viola RE. Detergent selection for enhanced extraction of membrane proteins. *Protein Expression and Purification*. 2012; 86:12–20. [PubMed: 22963795]
- Aryal UK, Krochko JE, Ross ARS. Identification of phosphoproteins in *Arabidopsis thaliana* leaves using polyethylene glycol fractionation, immobilized metal-ion affinity chromatography, two-dimensional gel electrophoresis and mass spectrometry. *Journal of Proteome Research*. 2012; 11:425–437. [PubMed: 22092075]
- Baker CL, Kettenbach AN, Loros JJ, Gerber SA, Dunlap JC. Quantitative proteomics reveals a dynamic interactome and phase-specific phosphorylation in the *Neurospora* circadian clock. *Molecular Cell*. 2009; 34:354–363. [PubMed: 19450533]

- Cellar NA, Kuppanan K, Langhorst ML, Ni W, Xu P, Young SA. Cross species applicability of abundant protein depletion columns for ribulose-1,5-bisphosphate carboxylase/oxygenase. *Journal of Chromatography B, Analytical Technologies in the Biomedical and Life Sciences*. 2008; 861:29–39.
- Chen EI, Cociorva D, Norris JL, Yates JR III. Optimization of mass spectrometry compatible surfactants for shotgun proteomics. *Journal of Proteome Research*. 2008; 6:2529–2538. [PubMed: 17530876]
- Covington MF, Harmer SL. The circadian clock regulates auxin signaling and responses in *Arabidopsis*. *PLoS Biology*. 2007; 5:1773–1784.
- Deery MJ, Maywood ES, Chesham JE, Sládek M, Karp N, Green EW, et al. Proteomic analysis reveals the role of synaptic vesicle cycling in sustaining the suprachiasmatic circadian clock. *Current Biology*. 2009; 19:2031–2036. [PubMed: 19913422]
- Dunn JD, Reid GE, Bruening ML. Techniques for phosphopeptide enrichment prior to analysis by mass spectrometry. *Mass Spectrometry Reviews*. 2010; 29:29–54. [PubMed: 19263479]
- Edwards KD, Anderson PE, Hall A, Salathia NS, Locke JCW, Lynn JR, et al. FLOWERING LOCUS C mediates natural variation in the high-temperature response of the *Arabidopsis* circadian clock. *Plant Cell*. 2006; 18:639–650. [PubMed: 16473970]
- Facette MR, Shen Z, Björnsdóttir FR, Briggs SP, Smith LG. Parallel proteomic and phosphoproteomic analyses of successive stages of maize leaf development. *Plant Cell*. 2013; 25:2798–2812. [PubMed: 23933881]
- Farré EM, Kay SA. PRR7 protein levels are regulated by light and the circadian clock in *Arabidopsis*. *The Plant Journal*. 2007; 52:548–560. [PubMed: 17877705]
- Guerreiro ACL, Benevento M, Lehmann R, van Breukelen B, Post H, Giansanti P, et al. Daily rhythms in the cyanobacterium *Synechococcus elongatus* probed by high-resolution mass spectrometry-based proteomics reveals a small-defined set of cyclic proteins. *Molecular and Cellular Proteomics*. 2014; 13:2042–2055. [PubMed: 24677030]
- Hiltbrunner A, Bauer J, Vidi P-A, Infanger S, Weibel P, Hohwy M, et al. Targeting of an abundant cytosolic form of the protein import receptor at Toc159 to the outer chloroplast membrane. *The Journal of Cell Biology*. 2001; 154:309–316. [PubMed: 11470820]
- Hindle MM, Le Bihan T, Krahmer J, Martin SF, Noordally ZB, Simpson TI, et al. qPmerge: Galaxy tools to reduce motif redundancy in proteomics. in review.
- Hwang H, Cho M-H, Hahn B-S, Lim H, Kwon Y-K, Hahn T-R, et al. Proteomic identification of rhythmic proteins in rice seedlings. *Biochimica et Biophysica Acta*. 2011; 1814:470–479. [PubMed: 21300183]
- Koroleva, OA.; Bindschedler, LV. Sample preparation in biological mass spectrometry. Ivanov, AR.; Lazarev, AV., editors. Springer; Netherlands: 2011. p. 381–409.
- Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology*. 2001; 305:567–580. [PubMed: 11152613]
- Kusakina J, Dodd AN. Phosphorylation in the plant circadian system. *Trends in Plant Science*. 2012; 17:575–583. [PubMed: 22784827]
- Mann M, Ong S, Grønborg M, Steen H, Jensen ON, Pandey A. Analysis of protein phosphorylation using mass spectrometry: Deciphering the phosphoproteome. *Trends in Biotechnology*. 2002; 20:261–268. [PubMed: 12007495]
- Mauvoisin D, Wang J, Jouffe C, Martin E, Atger F, Waridel P, et al. Circadian clock-dependent and -independent rhythmic proteomes implement distinct diurnal functions in mouse liver. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111:167–172. [PubMed: 24344304]
- Reddy AB, Karp NA, Maywood ES, Sage EA, Deery M, O'Neill JS, et al. Circadian orchestration of the hepatic proteome. *Current Biology*. 2006; 16:1107–1115. [PubMed: 16753565]
- Reiland S, Messerli G, Baerenfaller K, Gerrits B, Endler A, Grossmann J, et al. Large-scale *Arabidopsis* phosphoproteome profiling reveals novel chloroplast kinase substrates and phosphorylation networks. *Plant Physiology*. 2009; 150:889–903. [PubMed: 19376835]

- Righetti PG, Boschetti E, Lomas L, Citterio A. Protein Equalizer Technology: The quest for a “democratic proteome”. *Proteomics*. 2006; 6:3980–3992. [PubMed: 16800034]
- Robles, MS.; Mann, M. Handbook experimental pharmacology. Kramer, A.; Mellow, M., editors. Vol. 217. Springer; Berlin Heidelberg: 2013. p. 389-407.
- Russell WK, Park ZY, Russell DH. Proteolysis in mixed organic-aqueous solvent systems: Applications for peptide mass mapping using mass spectrometry. *Analytical Chemistry*. 2001; 73:2682–2685. [PubMed: 11403317]
- Slade, WO.; Werth, EG.; Chao, A.; Hicks, LM. Phosphoproteomics in photosynthetic organisms. *Electrophoresis*. 2014. <http://dx.doi.org/10.1002/elps.201400154>
- Sugano S, Andronis C, Ong MS, Green RM, Tobin EM. The protein kinase CK2 is involved in regulation of circadian rhythms in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*. 1999; 96:12362–12366. [PubMed: 10535927]
- Tsuji T, Hirota T, Takemori N, Komori N, Yoshitane H, Fukuda M, et al. Circadian proteomics of the mouse retina. *Proteomics*. 2007; 7:3500–3508. [PubMed: 17726681]
- Vizcaíno JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Ríos D, et al. ProteomeXchange provides globally co-ordinated proteomics data submission and dissemination. *Nature Biotechnology*. 2014; 30(3):223–226.
- Wagner V, Fiedler M, Markert C, Hippler M, Mittag M. Functional proteomics of circadian expressed proteins from *Chlamydomonas reinhardtii*. *FEBS Letters*. 2004; 559:129–135. [PubMed: 14960320]
- Wagner, V.; Mittag, M. Probing Circadian Rhythms in *Chlamydomonas reinhardtii* by Functional Proteomics. In: Pfannschmidt, T., editor. *Plant Signal Transduction*. Vol. 479. Humana Press; 2009. p. 173-188.
- Wang P, Xue L, Batelli G, Lee S, Hou Y-J, Van Oosten MJ, et al. Quantitative phosphoproteomics identifies SnRK2 protein kinase substrates and reveals the effectors of abscisic acid action. *Proceedings of the National Academy of Sciences of the United States of America*. 2013; 110:11205–11210. [PubMed: 23776212]
- Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nature Methods*. 2009; 6:3–7. [PubMed: 19116609]
- Yeung Y-G, Stanley ER. Rapid detergent removal from peptide samples with ethyl acetate for mass spectrometry analysis. *Current Protocols in Protein Science*. 2010; 59:16.12:16.12.1–16.12.5.

Figure 1. Comparison of protein extraction methods. Protein was extracted from *Arabidopsis* rosette tissue using several detergents and detergent removal methods. After digestion with trypsin, the resulting peptides were analyzed by high-resolution mass spectrometry, with or without phosphopeptide enrichment on TiO₂ resin. Replicates were five except for sample OG TCA and SDS TCA (three replicates). For all samples, the same pool of ground tissue was used except for OG TCA, which was extracted from rosettes that were 1-week younger. (A) Workflow for all extraction methods, names of resulting samples and approximate duration

of each step (red (light gray in the print version)) when processing 30 samples. (B) Average number of proteins identified by each method in the global peptide analysis. (C) Average number of phosphopeptides and nonphosphorylated peptides identified by each method in the phosphoenriched samples. Lower error bars: standard error of the mean (SEM) of number of phosphopeptides. Upper error bars: SEM of the number of total detected peptides. (D) Boxplot illustrating the coefficient of variation (CV) of protein abundance for the four best performing methods in the global peptide analysis (top) and the abundance of phosphopeptides in the phosphoenriched samples (bottom). The data were arcsinh transformed before calculating the CV and only entries with at least three nonzero values were used. Otherwise, zero values were ignored for determining the CV. Center of notch: median; notch: 95% confidence interval of median; box: first to third quartile (Q1–Q3); whiskers: $Q1 - 1.5 \times$ interquartile range or smallest data point (bottom whiskers); $Q3 + 1.5 \times$ interquartile range or largest data point (top whiskers). (E) Pairwise assessment of method specificity: each cell shows the number of phosphopeptides that were identified in all replicates of one method (top row, “present in”) but in none of the replicates of the respective other method (left column, “absent in”). Cells were shaded for ease of reading: the lighter the shade, the higher the number of method-specific phosphopeptides. * $p < 0.05$ and ** $p < 0.01$ in t -test against “No Detergent.”

Figure 2.

Protein fractionation by centrifugation with PEG. Protein was separated into two fractions (“Pellet” and “Supernatant”) by centrifugation with PEG, along with a control without fractionation (“Unfractionated”) and a control using the previously described method “IGEPAAL TCA.” Replicates were five for “Unfractionated” and “Pellet,” three for the “IGEPAAL TCA” control and four for “Supernatant.” (A) Workflow for all methods, names of resulting samples and approximate duration of each step (red (light gray in the print version)) when processing 30 samples. (B) Examples: normalized abundance of the phosphopeptide AGSFRDpSPEEEGPVELPEAAR from the clock protein TIC (top) found in the analysis of phosphoenriched peptides; normalized abundance of the clock output protein CCR1 (middle) and the RuBisCO large subunit protein (bottom) from the global peptide analysis. (C) Average number of proteins identified by each sample type in the global peptide analysis; error bars: SEM. (D) Average number of phosphopeptides and nonphosphorylated peptides identified by each method in the phosphoenriched peptide

analysis. Lower error bars: SEM of the number of phosphopeptides. Upper error bars: SEM of the number of total detected peptides. (E) Boxplot illustrating the CV of protein abundance in the global peptide analysis (top) and the CV of phosphopeptide abundance in the phosphoenriched sample (bottom). See legend of Fig. 1D for transformation of data and boxplot details. (F) Average number of phosphopeptides identified by any possible combination of two mass spec analyses, of either the same workflow or one pellet and one supernatant; error bars: SEM.

Figure 3.

Protein extraction using the acid-labile detergent RapiGestTM SF. Protein was extracted with the acid-labile detergent RapiGestTM (sample "RapiGest"). (A) Workflow and names of resulting samples and approximate duration of each step (red (light gray in the print version)) for processing 30 samples. A control was treated the same but without addition of RapiGestTM. Another control was performed in the same way as "No Detergent" in Fig. 1A. (B) Average number of proteins identified by each method in the global peptide analysis; error bars: SEM. (C) Average number of phosphopeptides and nonphosphorylated peptides identified by each method in the phosphoenriched peptide analysis. Lower error bars: SEM of number of phosphopeptides. Upper error bars: SEM of the number of total detected peptides. (D and E) Boxplots illustrating the CV of (D) protein abundance in the global peptide analysis and (E) the abundance of phosphopeptides in the phosphoenriched samples. See legend of Fig. 1D for transformation of data and boxplot details.

Table 1

Gene ontology analysis

Enriched in IGEPAL TCA over No Detergent		Enriched in No Detergent over IGEPAL TCA	
Global	Phospho-enriched	Global	Phospho-enriched
Chloroplast thylakoid membrane	Membrane	Cellular component	None
Membrane part		Ribosomal subunit	
Protein complex		Cytosolic ribosome	
Chloroplast stroma		Intracellular nonmembrane-bounded organelle	
Photosynthetic membrane		Intracellular organelle part	
Thylakoid membrane		Ribosome	
Plastid thylakoid membrane		Ribonucleoprotein complex	
Intracellular membrane-bounded organelle		Nonmembrane-bounded organelle	
Thylakoid		Organelle part	
Chloroplast		Cytosolic part	
Thylakoid part			
Cytoplasmic part			
Chloroplast part			
Membrane			
Chloroplast thylakoid			
Plastid			
Plastid stroma			
Intracellular organelle part			
Plastid thylakoid			
Organelle subcompartment			
Membrane-bounded organelle			
Organelle part			
Plastid part			
Plastid part			

Methods “IGEPAL TCA” and “No Detergent” were compared in a GO analysis. Terms enriched by one method relative to the other are shown for the global analysis as well as the phospho-enriched peptide analysis. Only the “cellular component” GO category is shown.