Published in final edited form as:

Methods. 2011 August ; 54(4): 379–386. doi:10.1016/j.ymeth.2011.03.008.

A solid phase extraction-based platform for rapid phosphoproteomic analysis

Noah Dephoure^{*} and Steven P. Gygi^{*}

Abstract

Protein phosphorylation is among the most common and intensely studied post-translational protein modification. It plays crucial roles in virtually all cellular processes and has been implicated in numerous human diseases, including cancer. Traditional biochemical and genetic methods for identifying and monitoring sites of phosphorylation are laborious and slow and in recent years have largely been replaced by mass spectrometric analysis. Improved methods for phosphopeptide enrichment coupled with faster and more sensitive mass spectrometers have led to an explosion in the size of phosphoproteomic datasets. However, wider application of these methods is limited by equipment costs and the resultant high demand for instrument time as well as by a technology gap between biologists and mass spectrometrists. Here we describe a modified two-step enrichment strategy that employs lysC digestion and step elution from self-packed strong cation exchange (SCX) solid phase extraction (SPE) columns followed by immobilized metal ion affinity chromatography (IMAC) and LC-MS/MS analysis using a hybrid LTQ Orbitrap Velos mass spectrometer. The SCX procedure does not require an HPLC system, demands little expertise, and because multiple samples can be processed in parallel, can provide a large savings of time and labor. We demonstrate this method in conjunction with stable isotope labeling to quantify peptides harboring >8,000 unique phosphorylation sites in yeast in 12 hours of instrument analysis time and examine the impact of enzyme choice and instrument platform.

Keywords

Proteomics; Phosphorylation; Chromatography/methods; Mass Spectrometry/methods; Phosphoproteomics

1. Introduction

Reversible protein phosphorylation modulates protein activity, localization, stability, and protein-protein interactions, and in so doing regulates a diverse array of cellular events that govern nearly every aspect of cellular growth and division[1]. Traditional methods for studying protein phosphorylation, such as in vitro kinase assays, metabolic labeling with radioactive phosphate, or the use of phospho-specific antibodies are laborious and slow and often require prior knowledge or educated guesses about the sites under study. In the wake of whole genome sequencing, mass spectrometric analysis has rapidly emerged as a

© 2011 Elsevier Inc. All rights reserved.

^{*}Correspondence to: Noah Dephoure (noah_dephoure@hms.harvard.edu) or Steven P. Gygi (sgygi@hms.harvard.edu), Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

powerful tool for unbiased protein analysis and has proven especially well-suited to the analysis of protein phosphorylation.

Search algorithms such as SEQUEST[2] and Mascot[3] match experimental peptide mass spectra with theoretical spectra generated from protein database sequences. In conjunction with a target-decoy searching strategy[4], they can accurately and reliably identify thousands of peptides and proteins from complex mixtures without prior knowledge of the sample content. Phosphopeptides are identified by considering the 79.9963 dalton mass difference of a phosphate group in each spectral match. In many cases, specific MS/MS fragment ions can uniquely localize the site of phosphorylation within the peptide. In some of the largest datasets, >10,000 unique phosphorylation sites have been identified [5–12].

The success of mass spectrometric phosphorylation analysis was only achieved after overcoming a number of complicating factors. Phosphate groups are particularly labile. Upon fragmentation for MS/MS, phosphopeptides often simply lose phosphoric acid, resulting in a neutral loss of 98 daltons, and may fail to produce sufficient fragmentation of the peptide backbone for identification. A number of strategies such as reselection of the neutral loss ion for additional fragmentation and MS³[13] analysis, additional activation and pseudo MS³[14] analysis, or alternate fragmentation methods, such as ETD[15], have been attempted with some success. More recent work suggests that this hurdle has largely been overcome by faster and more sensitive instrumentation[16].

A much more difficult problem is the low stoichiometry of most phosphorylation sites. It is still nearly impossible to detect phosphopeptides in complex mixtures without enrichment. One effective strategy for enrichment is the use of phospho-specific antibodies. This has been most successful in the analysis of tyrosine phosphorylation [17], due to the high selectivity of existing antibodies. Similar quality antibodies for phospho-serine and phospho-threonine have not been described.

The success of unbiased large-scale phosphoproteomic analysis has been enabled by two major enrichment techniques: 1) phospho-affinity purification, using immobilized metal ion affinity chromatography[18] (IMAC) or titanium dioxide[19] (TiO₂), and 2) charge-based separation using strong cation exchange chromatography (SCX). Early applications of each demonstrated their effectiveness, but resulted in quite modest datasets by current standards. In a breakthrough report in 2001 using IMAC, Ficarro et al. [20] reported 383 phosphorylation sites from *S. cerevisiae*, the largest such dataset at the time. In 2004, 2,002 sites enriched by SCX from HeLa cell nuclear extracts were reported [21], again a then unprecedented number.

Strong cation exchange chromatography separates peptides by their solution charge state. When performed at a pH below the pK_a for the acidic amino acids, only phosphorylated amino acids can contribute a negative charge, and a given phosphopeptide will elute earlier than its unphosphorylated form, concentrating phosphopeptides in the earliest fractions. It was only when IMAC or TiO₂ was combined with SCX, that effective phosphopeptide enrichment for all but the earliest eluting SCX fractions was achieved. This added two-step enrichment strategy, along with the introduction of faster mass spectrometers such as the LTQ, LTQ Orbitrap, and now the LTQ Orbitrap Velos, has led to an explosion in the number of reported sites. At the time of publication, the Phosphosite Database[22] (www.phosphosite.org) contained >95,000 unique protein phosphorylation sites from 21 different organisms, up from 13,018 in early 2008.

The first successful application of the two-step phosphopeptide enrichment was published in 2005 [13], describing the identification of 729 sites from yeast. By early 2007 a number of additional reports had surfaced [23–25], the largest of which identified 6,600 sites from

HeLa cells and 5,635 sites from mouse liver. Since these early reports, this strategy and variations of it have been successfully applied in different experimental systems[26].

Despite the apparent robustness of these methods, almost all of these reports come from only a handful of experienced mass spectrometry labs. In the current report we have ported the HPLC-based SCX separation to a solid-phase extraction system using vacuum flow and step elution, providing a platform for processing many samples simultaneously. In addition to streamlining the procedure, our method relieves the need for access to, and expertise in, running an HPLC system, allowing nearly any laboratory to prepare highly enriched phosphopeptide samples ready for mass spectrometric analysis. We describe the method and demonstrate its power by performing a quantitative phosphoproteomic analysis of a 1:1 mixture of two metabolically labeled wild-type budding yeast cultures. We identified more than 8,000 unique phosphorylation sites using only seven 95-min runs on a LTQ Orbitrap Velos mass spectrometer. We examine two additional factors in phosphoproteomic analysis, instrument platform and enzyme choice. We show that analysis of the same samples on the Velos yielded 27% more unique sites in 25% less time than analysis on a LTQ Orbitrap Discovery and demonstrate that lysC digestion generated 21% more unique phosphorylation sites than a parallel analysis using trypsin.

2. Rationale

As the phosphoproteomic platform in our lab matured, increases in the number of samples, led to worsening bottle-necks at both the SCX peptide separation and at the mass spectrometer. To increase sample preparation throughput, we began experimenting with solid phase extraction separation, allowing us to process many samples in parallel. We reasoned that because we were collecting relatively broad 4-min fractions from the HPLC, that sufficient separation could be achieved with discrete salt elution steps from a SPE column of the same material. Instrumentation costs preclude any similar increase in the number of samples that can be simultaneously analyzed in the mass spectrometer. However, the introduction of a new instrument, the LTQ Orbitrap Velos, with improvements in both sensitivity and speed[27], allowed us to reduce both the number of samples analyzed and the length of each analysis without sacrifices in the size or quality of the dataset.

3. Materials and Methods

The work-flow of a SILAC-based quantitative phosphoproteomic experiment is depicted in Fig. 1. We have previously described a detailed protocol for SCX-IMAC phosphopeptide enrichment and analysis[28]. Though the overall work-flow is unchanged, we describe a number of modifications to that method; namely - the replacement of 1) trypsin with lysC, 2) HPLC based SCX fractionation with a hand-packed SPE column and step elution, and 3) the LTQ Orbitrap with the LTQ Orbitrap Velos - and demonstrate the advantages of these changes. We address each step, and considerations for each, in turn, providing step-by-step protocols at select points, and detail how the described dataset was collected and analyzed.

3.1. Starting material

Most protein phosphorylation is sub-stoichiometric [29], i.e. for every unique amino acid sequence there are on average far more unphosphorylated molecules than phosphorylated. Successful phosphopeptide detection thus requires the enrichment of phosphopeptides and removal of the much more abundant unphosphorylated peptides. Because phosphopeptides make up such a small fraction of the total peptide pool, it is also necessary to start with a large amount of material in order to obtain detectable amounts of phosphopeptides from less abundant proteins. We routinely work with ~10 mg of total protein, using columns packed with 1 g SCX resin. Larger amounts of peptide can be processed by parallel separation on

multiple columns. Where sample material is limited, valuable, though smaller datasets may still be achieved. We have identified nearly 5,000 phosphorylation sites from 3 mg of total yeast protein. While we anticipate that even smaller amounts of starting material can still provide thousands of phosphorylation sites, this protocol was designed for large-scale analysis, and may not be suitable for smaller amounts of starting material.

3.2. SILAC Labeling and Cell Growth

The included quantitative data were collected using stable isotope labeling of amino acids in cell culture[30] (SILAC). The method is, however, compatible with almost any quantitative approach as well as non-quantitative analysis.

SILAC is commonly performed by dual labeling with isotopically labeled lysine and arginine, followed by tryptic digestion. We favor a SILAC protocol that omits labeled arginine, using only isotopically labeled lysine, because it reduces the cost and complexity of experiments. Labeled arginine can be metabolized to proline, visible as satellite peaks in the mass spectra of proline containing peptides, impairing the accuracy of quantification[31]. A lysine-only protocol also reduces the need for genetic manipulation required to render some organisms, such as *S. cerevisiae*, amenable to arginine labeling. It is not, however, compatible with tryptic digestion, as nearly half the resulting peptides will contain no lysines and thus cannot be quantified. Digestion with lysyl endopeptidase (lysC, see below), which cuts at the c-terminal side of lysine residues, overcomes this limitation.

3.3 Cell Lysis

For SILAC samples, we recommend mixing equal cell numbers before lysis to generate the most carefully controlled conditions. Alternatively, cells can be lysed separately and equal amounts of total protein mixed after a simple protein assay. Any lysis method and digestion protocol that preserves protein phosphorylation and yields a clean, detergent-free peptide pool should be compatible with this method. We routinely lyse both yeast and mammalian cells in 8 M urea buffer including both protease and phosphatase inhibitors as a precaution against residual endogenous activity of both classes of enzyme. If detergent containing lysis buffers are used, it is essential to remove the detergent before mass spectrometric analysis. This is most commonly done either by precipitating the proteins and extensive washing, or through the use of various filtration devices that allow buffer exchanges prior to enzymatic digestion.

In the presented experiments, yeast cells (strain S288C) were grown to mid-log phase in synthetic complete media containing either $^{13}\text{C}_6^{15}\text{N}_2$ -lysine (Cambridge Isotopes, Cambridge, MA, #CNLM-281-H) ("heavy") or naturally occurring lysine ("light"). Cells were collected by centrifugation at 4° C, and equal cell numbers of "heavy" and "light" cultures were combined, frozen in liquid nitrogen and stored at -80° C. After thawing on ice, cells were resuspended in 700 μl urea lysis buffer including protease and phosphatase inhibitors (8 M urea, 75 mM NaCl, 50 mM Tris pH 8.0, 1 mM AEBSF, 2.5 mM benzamidine, 1 $\mu\text{g/ml}$ leupeptin, 1 $\mu\text{g/ml}$ pepstatin, 100 mM beta glycerol-phosphate, 10 mM sodium fluoride, 100 nM calyculin A). 1.5–2.0 ml acid-washed glass beads (Biospec, Bartlesville, OK) were added and the cells were ruptured by bead-beating on a Biospec mini-beadbeater using 4 $\times$ 60 sec. cycles with 3 min. rests on ice in-between. Lysates were transferred to fresh tubes and cleared by centrifugation at 21k $\times$ g for 10 min at 4° C. The supernatants were removed to fresh tubes and the protein concentration was measured by dye binding assay. Disulfide bonds were reduced by adding dithiothreitol to a final concentration of 2.5 mM and incubating at 56°C for 40 min. The extract was allowed to cool to room temperature and the reduced cysteines were alkylated by adding iodoacetamide to

7.5 mM and incubating for 40 min in the dark at room temperature. Alkylation was quenched with an additional 5 mM dithiothreitol.

3.4 Enzymatic Digestion with LysC

We initially began using lysyl endopeptidase (LysC) (Wako USA, #12-02541) in a lysine-only SILAC protocol for cells that could not be arginine labeled. We quickly discovered that this was a reliable and effective approach with a number of advantages: LysC digestion i) produces markedly fewer missed cleavages, ii) can be performed in higher concentrations of urea, and iii) in our hands, leads to more unique protein identifications and more unique phosphopeptides and phosphorylation sites (Table 1).

LysC Digestion Protocol

1. Reduce the urea concentration to ≤ 4 M and bring to pH 8.8 with 25 mM (final concentration) Tris-HCl, pH 8.8. We routinely dilute our samples 2.5-fold.
2. Add LysC to 10 ng/ μ l. This usually results in an enzyme to substrate ratio ~1:50–200.
3. Digestion proceeds overnight with gentle agitation at room temperature.
4. Acidify the digest by the addition of FA to 1% (final concentration). Depending on the amount of sample, a visible lipid rich precipitate will form.
5. Centrifuge 2 min at $\geq 14k \times g$ to pellet this precipitate and remove the peptide containing supernatant to a new tube, or load directly onto a prepared solid-phase extraction cartridge for desalting.

3.5. Desalting

Desalting steps are required at three points in the protocol; after protein digestion, after SCX separation, and after IMAC elution. In these experiments we have simplified the buffer systems for all desalting steps, using only 1% FA as the aqueous loading and wash buffer (dsLWB), and 70% ACN/1% FA as the organic, elution buffer (dsEB). Desalting of the digested whole-cell lysate and SCX fractions is done on tC18, reverse-phase SPE cartridge (Waters, Sep-pak) using a 200 mg cartridge for the unseparated digest and 50 mg cartridges for the SCX fractions (if starting with >10 mg of total protein, use a larger capacity cartridge for the digest). After IMAC elution, only a tiny fraction of the original peptide pool remains. The final desalting step is performed on C18-reverse phase micro-columns, using self-prepared STAGE-tips[32]. Volumes are given for all three formats from largest to smallest. To minimize sample handling, we recommend a protocol that combines the final desalting step with IMAC enrichment (detailed below). During sample loading and elution steps, take care to limit the flow rates to ~ 0.5 ml/min for Sep-paks and ~ 10 μ l/min for STAGE-tips.

Desalting Buffers

- Desalting loading and wash buffer dsLWB: 1% FA
- Desalting elution buffer dsEB: 70% ACN, 1% FA

The general desalting protocol is:

1. Wet the column with 6 ml, 2 ml, or 50 μ l 100% methanol.
2. Wash with 6 ml, 2 ml, or 50 μ l dsEB
3. Equilibrate with 6 ml, 2 ml, or 50 μ l dsLWB

4. Load sample in aqueous solution, if dried, resuspend in 1–2 ml, 200–500 μ l, or 20–50 μ l dsLWB
5. Wash with 6 ml, 2 ml, or 50 μ l dsLWB
6. Elute with 1 ml, 500 μ l, or 50 μ l dsEB
7. Dry in a speed-vac. Dried peptides can be stored at -20° C.

3.6. SPE SCX Columns

Materials

- 6 ml syringe barrel style columns (Applied Separations, #2416)
- 20 micron PE frit (Applied Separations, #2426)
- 20 micron particle, 300 angstrom pore size, PolySULFOETHYL A bulk packing material (Nest Group, #BMSE2003)

Fabrication – A schematic of the column composition is shown in Fig. 1B

1. For each column, make a slurry of 1 g resin and 2 ml 30% ACN and transfer it to the column
2. Allow the resin bed to settle and excess buffer to flow through by gravity or gentle applied positive or vacuum pressure
3. Seal the bed with a second frit.

For all experiments, we used a vacuum manifold (Waters, #WAT200609) to drive the flow. Alternatively, flow can be controlled by positive pressure or centrifugation with the SPE cartridge fitted into plastic culture tubes for collection. In all cases the flow rate should be monitored. We have also had good success with similar columns packed with 12 μ m resin. However, its slower flow rates make it more difficult to work with.

3.7. Performing SCX by SPE

The salt cuts used in this study (10 mM, 25 mM, 40 mM, 60 mM, 90 mM, 150 mM) were chosen to provide greater separation at the low-salt concentrations where phosphopeptides concentrate and broader separation elsewhere. Though they were arrived at empirically, SPE lacks the precise reproducibility of HPLC separation and we have observed run to run variations in the peptide solution charge distribution. As presented, the protocol has proven robust. However, we recommend tracking this distribution in all experiments and modifying the protocol as needed.

SCX Buffers

- SCX Buffer A: 7 mM KH_2PO_4 , pH 2.65, 30% ACN (n.b. - adjust pH with H_3PO_4 before adding ACN)
- SCX Buffer B: 7 mM KH_2PO_4 , pH 2.65, 30% ACN, 350 mM KCl (n.b. - adjust pH with H_3PO_4 before adding ACN)

Prepare the column(s):

1. Wet with 6 ml 80% ACN.
2. Wash with 6 ml water.
3. Condition each column by performing a blank run as described below, using a single 100% SCX Buffer B elution step.

Resuspend the desalted peptide digest in 1–2 ml SCX Buffer A. Vigorous agitation and/or a sonicating water bath can be used to help resuspend stubborn pellets. Flow rates for all SCX steps should be ~1–2 ml/min. However, equilibration in SCX Buffer A before each run should proceed for ≥ 30 min either by decreasing the flow rate or by using additional SCX Buffer A. After equilibration and each elution step, bring the buffer level below the frit. We find that partially drying the column bed between steps has no unwanted consequences and may act to sharpen the transitions between salt elution steps.

SCX Separation Protocol - The step elution protocol is depicted in Fig. 1C

1. Equilibrate the column in ≥ 6 ml SCX Buffer A.
2. Prepare 6 ml of each salt step solution (10 mM, 25 mM, 40 mM, 60 mM, 90 mM, and 150 mM KCl) per column by mixing appropriate ratios of SCX Buffers A and B.
3. Load sample in 1–2 ml SCX Buffer A, collecting the flow-through as fraction 1.
4. Wash with 4–5 ml SCX Buffer A, pooling with the flow-through (6 ml total).
5. Elute with 6 ml of each salt-step in succession, collecting each fraction.
6. Freeze the fractions in liquid nitrogen and lyophilize overnight. Alternatively, fractions can be split to fit in 2 ml microcentrifuge tubes and dried in a speed-vac.
7. Dried peptides can be stored at -20°C
8. Resuspend peptides in 1 ml 1% FA in preparation for desalting.

3.8. IMAC

IMAC phosphopeptide enrichment is as described[28].

Materials

- IMAC resin: PHOS-select Iron Affinity Gel (Sigma, #P9740)

IMAC Buffers

- IMAC wash and equilibration buffer (WEB): 40% ACN, 1% FA
- IMAC elution buffer (EB): 500 mM K_2HPO_4 , pH = 7

Combined IMAC and STAGE-tip desalting protocol:

1. Resuspend dried peptides in 100 μl IMAC WEB.
2. Incubate with 10 μl of the original slurry of washed IMAC resin, with vigorous agitation for 60 min. at room temperature.
3. Transfer to prepared STAGE-tips (see desalting steps 1–3) and wash three times with 50 μl WEB.
4. Elute from the IMAC resin with 3 changes of 70 μl IMAC EB. Aqueous eluted peptides will be bound by the STAGE-tip.
5. Wash the stage tip with 50 μl dsLWB.
6. Elute in 40 μl dsEB directly into glass inserts for mass spec analysis.

3.9. Mass Spectrometric Analysis

Desalted phosphopeptides were resuspended in 10 μ l 5% FA. For all experiments, a single replicate of seven samples was analyzed on a LTQ Orbitrap Velos mass spectrometer equipped with an Agilent 1100 binary pump (#G1312A) or a LTQ Orbitrap Discovery equipped with a Thermo Surveyor MS Pump (ThermoFisher). Both systems use flow-splitting to achieve flow rates of 0.5 μ l/min across the column. Spray tips were hand-pulled using 100 μ m ID fused-silica tubing and packed to 15 cm with 3 μ m, 200 Å Magic C18AQ resin (Michrom Bioresources, Inc., PM3/61200/00). Peptides were loaded using a Famos autosampler (LC Packings) and separated using a 67-min linear gradient of 8–28% ACN in 0.125% FA on the Velos and a 99-min, 12–32% gradient on the Discovery. The different HPLC pumps behave quite differently as these gradients provide similar base peak profiles.

Peptides were detected in a hybrid dual-cell quadrupole linear ion trap – orbitrap mass spectrometer (LTQ Orbitrap Velos, ThermoFisher) using a data-dependent Top20 method, or the single-cell linear ion trap – orbitrap mass spectrometer (LTQ Orbitrap Discovery, ThermoFisher) using a Top10 method. For each cycle, one full MS scan of $m/z = 300$ –1500 was acquired in the Orbitrap at a resolution of 60,000 (Velos) or 30,000 (Discovery) at $m/z = 400$ with AGC target = 10^6 . Each full scan was followed by the selection of the most intense ions, up to 20 (Velos) or 10 (Discovery), for CID and MS/MS analysis in the LTQ. Selected ions were excluded from further analysis for 60 s. Ions with charge 1^+ or unassigned were also rejected. Maximum ion accumulation times were 1000 ms for each full MS scan and 150 ms for MS/MS scans. Lockmass, employing atmospheric polydimethylsiloxane ($m/z = 371.1012$) as an internal standard was used in all runs to normalize MS precursor masses.

MS/MS spectra were searched via the SEQUEST algorithm against a composite database containing all translated yeast open reading frames and their reversed complements using the following parameters: a precursor mass tolerance of ± 25 ppm; 1.0 Da product ion mass tolerance; fully tryptic or lysC digestion; up to two missed cleavages; a static modification of carbamidomethylation on cysteine (+57.0214); and dynamic modifications for phosphorylation on serine, threonine, and tyrosine (+79.9663), methionine oxidation (+15.9949), and $^{13}\text{C}_6^{15}\text{N}_2$ -lysine (+8.0142).

Peptide spectral matches were filtered to 1% FDR using the target-decoy strategy[4] combined with linear discriminant analysis (LDA) using several different parameters including Xcorr, $\Delta\text{Cn}'$, precursor mass error, and charge state[9]. Though consistent parameters were used throughout, linear discriminant models were calculated for each LC-MS/MS run using peptide matches using matches to forward and reversed protein sequences as positive and negative training data. Peptide spectral matches within each run were sorted in descending order by discriminant score and filtered to a 1% FDR as revealed by the number of decoy sequences remaining in the data set. The data were further filtered to control protein-level FDRs. Peptides from all fractions in each experiment were combined and assembled into proteins. Protein scores were derived from the product of all LDA peptide probabilities, sorted by rank, and filtered to 1% FDR as described for peptides. The resulting peptide FDRs fall dramatically after protein filtering. We used the Ascore algorithm[33] to assign phosphorylation site localization. Sites with $p < 0.05$ (Ascore > 13) were considered localized.

Peptide quantification was performed using the Vista program[34]. C-terminal peptides containing no lysine were removed from the dataset, and we required a signal-to-noise ratio ≥ 3 for both heavy and light species for quantification.

4. Results and Discussion

To demonstrate our platform, we performed a quantitative large-scale phosphoproteomic experiment comparing a 1:1 mixture of wild-type yeast grown in “heavy” media to wild-type yeast grown in “light” media. Cells were mixed in equal numbers, lysed, digested with lysC, enriched by the described SPE/SCX/IMAC protocol, and analyzed with a 95-min method on a LTQ Orbitrap Velos. Spectra were matched using SEQUEST[2] with a target-decoy strategy, and filtered first to a 1% FDR at the peptide level, and then to a 1% FDR at the protein level, resulting in a further reduction of the peptide FDR. Peptide quantification was performed using VISTA[34] and the probability of correct site localization was calculated using the Ascore algorithm[33].

To assess the quality of the SCX separation afforded by our method, we analyzed the solution charge state distribution for all peptides. Though compressed compared to our previous 12 fraction HPLC method, we observe the expected progression of low solution charge state peptides in the early fractions to higher charge peptides in later fractions (Fig. 2A). Another measure of the quality of separation is the peptide overlap between fractions. Nearly 80% of the observed peptides were found in only a single fraction (Fig. 2B), similar to our previously reported overlap using HPLC separation of tryptic peptides[28].

The dataset includes 29,623 phosphopeptides with a false discovery rate of 0.1 %, describing at least 8,183 unique phosphorylation sites on 1,938 proteins (Fig 3A). The complete list of phosphorylation sites and phosphopeptides are available as supplemental tables. Phosphopeptide enrichment was excellent (Fig. 3B). All but the last two fractions yielded more than 4,000 phosphopeptides. As expected, earlier fractions were more highly enriched for phosphopeptides. Total phosphopeptide enrichment was 77% and as high as 95% in fraction one. More than 7,000 unique sites were identified from fractions 1–5 alone (Fig. 3C). Where high-throughput is essential, an additional 25% savings in analysis time could be realized by simply omitting the analysis of fractions six and seven, with only ~11% data loss.

Though there is still great value in simply identifying cellular phosphoproteins and phosphorylation sites, probing the significance of these events requires the ability to monitor changes across experimental conditions. In these experiments, “heavy” lysine, incorporated during cell growth, allowed accurate relative quantification for more than 27,000 phosphopeptides in 12 hours of analysis (Fig. 3D). The distribution of $\log_2(\text{Heavy:Light})$ ratios is clustered tightly around the expected one-to-one ratio at zero, with standard deviation = 0.35.

We attribute much of the success of this method to the greater sensitivity and approximate 2-fold increase in scanning speed of the LTQ Orbitrap Velos[27, 35]. To more carefully assess its impact on our phosphoproteomic platform, we re-analyzed all seven same samples in a single replicate on a LTQ Orbitrap Discovery using a longer 125-min method. The results of this comparison are shown in Fig. 4A and summarized in Table 1. The complete list of phosphorylation sites and phosphopeptides is presented in supplemental tables 1 and 2 respectively. Though the LTQ Orbitrap Discovery analysis produced a strong dataset, the Velos collected 28% more MS/MS spectra leading to the identification of 27% more phosphorylation sites in three-fourths the time.

The last major change to our phosphopeptide analysis platform is the replacement of the proteomic workhorse enzyme, trypsin, with lysC. The original SCX phosphopeptide enrichment protocol [21] was developed using trypsin, which cleaves the peptide bond on the c-terminal side of the basic amino acids arginine and lysine. The majority of the resulting peptides contain only a single basic residue and have a solution charge = +2 at

acidic pH. Phospho-forms of the same peptides retain a negative charge on the phosphate group and an overall solution charge = +1. Minimizing the number of basic residues per peptide by using trypsin was important for concentrating phosphopeptides in the early, low-salt fractions of the SCX gradient. However, the addition of a second enrichment step, such as IMAC or TiO₂, provides efficient recovery of phosphopeptides from later higher solution charge state fractions. We reasoned that it might no longer be necessary to concentrate the phosphopeptides in the earliest fractions, and wondered how using lysC, which produces higher solution charge state peptides, would affect SCX/IMAC analysis. To address this question, we performed non-quantitative SPE/SCX/IMAC analysis using 35 mg of unlabeled tryptic or lysC digested yeast whole-cell extract proteins. In each case, seven IMAC-enriched salt fractions were analyzed using a 125-min method on a LTQ Orbitrap Discovery.

LysC digested samples led to the identification of 21% more phosphorylation sites (Fig. 4B). One possible reason for this is that our elution steps may be better optimized for the slightly higher charge lysC peptides (Fig. 2C). However, even in the earliest fractions, where we would expect more tryptic phosphopeptides, lysC performs as well or nearly as well as trypsin. We speculate that lysC, while still yielding tryptic-like peptides, may produce significantly less complex samples. In addition to cutting at fewer sites, it produces far fewer missed cleavages and fewer ragged ended peptides. The improved performance of lysC is compelling and deserves further study.

5. Conclusion

Given the crucial roles that phosphorylation plays in cellular processes and human pathologies, phosphoproteomic profiling holds great promise for understanding cellular disease states, as a possible diagnostic tool, and as a screening method for novel therapeutic compounds, such as kinase inhibitors. The simplicity of this method allows biologists interested in performing phosphoproteomic analysis that lack HPLC expertise to prepare samples for submission to a mass spectrometry facility; while its speed enhances the throughput of phosphorylation analysis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by in part by National Institutes of Health Grant HG3456 to SPG.

Abbreviations

ACN	acetonitrile
ETD	electron transfer dissociation
IMAC	immobilized metal ion affinity chromatography
SCX	strong cation exchange chromatography
m/z	mass to charge ratio
CID	collision induced dissociation
S/N	signal to noise ratio
SPE	solid phase extraction

FA	formic acid
HPLC	high performance liquid chromatography
AGC	automatic gain control
FDR	false discovery rate
SILAC	stable isotope labeling with amino acids in cell culture

References

1. Hunter T. Signaling--2000 and beyond. *Cell*. 2000; 100(1):113–127. [PubMed: 10647936]
2. Yates JR 3rd, Eng JK, McCormack AL, Schieltz D. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal Chem*. 1995; 67(8):1426–1436. [PubMed: 7741214]
3. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*. 1999; 20(18):3551–3567. [PubMed: 10612281]
4. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods*. 2007; 4(3):207–214. [PubMed: 17327847]
5. Chen SH, Albuquerque CP, Liang J, Suhandynata RT, Zhou H. A proteome-wide analysis of kinase-substrate network in the DNA damage response. *J Biol Chem*. 2010; 285(17):12803–12812. [PubMed: 20190278]
6. Dephoure N, Zhou C, Villen J, Beausoleil SA, Bakalarski CE, Elledge SJ, Gygi SP. A quantitative atlas of mitotic phosphorylation. *Proc Natl Acad Sci U S A*. 2008; 105(31):10762–10767. [PubMed: 18669648]
7. Hilger M, Bonaldi T, Gnäd F, Mann M. Systems-wide analysis of a phosphatase knockdown by quantitative proteomics and phosphoproteomics. *Mol Cell Proteomics*. 2009; 8(8):1908–1920. [PubMed: 19429919]
8. Holt LJ, Tuch BB, Villen J, Johnson AD, Gygi SP, Morgan DO. Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science*. 2009; 325(5948):1682–1686. [PubMed: 19779198]
9. Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, Beausoleil SA, Villen J, Haas W, Sowa ME, Gygi SP. A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell*. 2010; 143(7):1174–1189. [PubMed: 21183079]
10. Nagaraj N, D'Souza RC, Cox J, Olsen JV, Mann M. Feasibility of large-scale phosphoproteomics with higher energy collisional dissociation fragmentation. *J Proteome Res*. 2010; 9(12):6786–6794. [PubMed: 20873877]
11. Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, Jensen LJ, Gnäd F, Cox J, Jensen TS, Nigg EA, Brunak S, Mann M. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci Signal*. 2010; 3(104):ra3. [PubMed: 20068231]
12. Zhai B, Villen J, Beausoleil SA, Mintseris J, Gygi SP. Phosphoproteome analysis of *Drosophila melanogaster* embryos. *J Proteome Res*. 2008; 7(4):1675–1682. [PubMed: 18327897]
13. Gruhler A, Olsen JV, Mohammed S, Mortensen P, Faergeman NJ, Mann M, Jensen ON. Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol Cell Proteomics*. 2005; 4(3):310–327. [PubMed: 15665377]
14. Schroeder MJ, Shabanowitz J, Schwartz JC, Hunt DF, Coon JJ. A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. *Anal Chem*. 2004; 76(13):3590–3598. [PubMed: 15228329]
15. Swaney DL, McAlister GC, Coon JJ. Decision tree-driven tandem mass spectrometry for shotgun proteomics. *Nat Methods*. 2008; 5(11):959–964. [PubMed: 18931669]
16. Villen J, Beausoleil SA, Gygi SP. Evaluation of the utility of neutral-loss-dependent MS3 strategies in large-scale phosphorylation analysis. *Proteomics*. 2008; 8(21):4444–4452. [PubMed: 18972524]

17. Rush J, Moritz A, Lee KA, Guo A, Goss VL, Spek EJ, Zhang H, Zha XM, Polakiewicz RD, Comb MJ. Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat Biotechnol.* 2005; 23(1):94–101. [PubMed: 15592455]
18. Andersson L, Porath J. Isolation of phosphoproteins by immobilized metal (Fe³⁺) affinity chromatography. *Anal Biochem.* 1986; 154(1):250–254. [PubMed: 3085541]
19. Pinkse MW, Uitto PM, Hilhorst MJ, Ooms B, Heck AJ. Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Anal Chem.* 2004; 76(14):3935–3943. [PubMed: 15253627]
20. Ficarro SB, McClelland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt DF, White FM. Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat Biotechnol.* 2002; 20(3):301–305. [PubMed: 11875433]
21. Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villen J, Li J, Cohn MA, Cantley LC, Gygi SP. Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci U S A.* 2004; 101(33):12130–12135. [PubMed: 15302935]
22. Hornbeck PV, Chabra I, Kornhauser JM, Skrzypek E, Zhang B. PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics.* 2004; 4(6):1551–1561. [PubMed: 15174125]
23. Olsen JV, Blagoev B, Gnäd F, Macek B, Kumar C, Mortensen P, Mann M. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell.* 2006; 127(3):635–648. [PubMed: 17081983]
24. Villen J, Beausoleil SA, Gerber SA, Gygi SP. Large-scale phosphorylation analysis of mouse liver. *Proc Natl Acad Sci U S A.* 2007; 104(5):1488–1493. [PubMed: 17242355]
25. Trinidad JC, Specht CG, Thalhammer A, Schoepfer R, Burlingame AL. Comprehensive identification of phosphorylation sites in postsynaptic density preparations. *Mol Cell Proteomics.* 2006; 5(5):914–922. [PubMed: 16452087]
26. Grimsrud PA, Swaney DL, Wenger CD, Beauchene NA, Coon JJ. Phosphoproteomics for the masses. *ACS Chem Biol.* 2010; 5(1):105–119. [PubMed: 20047291]
27. Second TP, Blethrow JD, Schwartz JC, Merrihew GE, MacCoss MJ, Swaney DL, Russell JD, Coon JJ, Zabrouskov V. Dual-pressure linear ion trap mass spectrometer improving the analysis of complex protein mixtures. *Anal Chem.* 2009; 81(18):7757–7765. [PubMed: 19689114]
28. Villen J, Gygi SP. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nat Protoc.* 2008; 3(10):1630–1638. [PubMed: 18833199]
29. Rogers LD, Foster LJ. Phosphoproteomics--finally fulfilling the promise? *Mol Biosyst.* 2009; 5(10):1122–1129. [PubMed: 19756301]
30. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics.* 2002; 1(5):376–386. [PubMed: 12118079]
31. Van Hoof D, Pinkse MW, Oostwaard DW, Mummery CL, Heck AJ, Krijgsvelde J. An experimental correction for arginine-to-proline conversion artifacts in SILAC-based quantitative proteomics. *Nat Methods.* 2007; 4(9):677–678. [PubMed: 17762871]
32. Rappsilber J, Ishihama Y, Mann M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem.* 2003; 75(3):663–670. [PubMed: 12585499]
33. Beausoleil SA, Villen J, Gerber SA, Rush J, Gygi SP. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat Biotechnol.* 2006; 24(10):1285–1292. [PubMed: 16964243]
34. Bakalarski CE, Elias JE, Villen J, Haas W, Gerber SA, Everley PA, Gygi SP. The impact of peptide abundance and dynamic range on stable-isotope-based quantitative proteomic analyses. *J Proteome Res.* 2008; 7(11):4756–4765. [PubMed: 18798661]
35. Olsen JV, Schwartz JC, Griep-Raming J, Nielsen ML, Damoc E, Denisov E, Lange O, Remes P, Taylor D, Splendore M, Wouters ER, Senko M, Makarov A, Mann M, Horning S. A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. *Mol Cell Proteomics.* 2009; 8(12):2759–2769. [PubMed: 19828875]

Fig. 1. Solid phase extraction-based strong cation exchange chromatography for phosphopeptide analysis

(A) An overview of a typical SILAC SCX/IMAC experiment showing all major sample-handling steps. (B) Hand-packing SPE/SCX cartridges. Empty 6 ml syringe barrel columns fitted with a bottom frit are filled with 1 g polySULFOETHYL A and sealed with a second frit. (C) A step gradient elution protocol for SPE/SCX peptide separation. Including the flow-through, seven discrete 6 ml salt elution steps are used to separate whole-cell extract derived peptides by charge state prior to IMAC phosphopeptide enrichment.

Fig. 2. Solid phase extraction-based strong cation exchange

(A) Shown is the solution charge state distribution of 38,487 phosphorylated and unphosphorylated peptides from lysC digested yeast whole-cell extract subjected to the described SPE/SCX/IMAC enrichment protocol analyzed by LC-MS/MS on a LTQ Orbitrap Velos. (B) The peptide overlap between SCX fractions is shown as the fraction of all peptides that were identified in 1, 2, 3, or more different fractions. (C) A comparison of the solution charge states of the identified phosphopeptides from trypsin and lysC digested yeast whole cell extracts.

Fig. 3. Quantitative phosphoproteomic analysis of a 1:1 mix of “heavy” and “light” labeled wild-type yeast

(A) A summary of identified peptides, phosphopeptides and proteins. Peptide FDR = 0.1%. Protein FDR = 1%. (B) Unphosphorylated (gray) and phosphorylated peptides (blue) per fraction. (C) Unique phosphopeptides per fraction (bars) and a cumulative count of unique phosphopeptides (solid line). (D) The distribution of log₂ peptide abundance ratios (Heavy:Light) for 27,754 phosphopeptides. SD = 0.35.

Fig. 4. The impact of instrument platform and enzyme choice on phosphoproteomic analysis

In both (A) and (B), phosphorylation site counts are shown for each analyzed fraction (bars) along with cumulative counts across fractions (solid lines). (A) A comparison of the same SPE/SCX/IMAC phosphopeptide analysis of lysC digested yeast whole-cell extracts shown in Fig. 3C performed on the LTQ Orbitrap Velos (shown in red) with analysis of the same samples run on a LTQ Orbitrap Discovery (blue). (B) Equal aliquots of yeast extract were digested with lysC (red) or trypsin (blue) and subject to SPE/SCX/IMAC analysis on a LTQ Orbitrap Discovery.

Table 1

A Comparison of SPE/SCX/IMAC analysis on two Orbitrap platforms and with two different enzymes.

	Expt. 1	Expt. 2	Expt. 3	Expt. 4
Enzyme	lysC	lysC	lysC	Trypsin
Instrument Platform	Velos	Discovery	Discovery	Discovery
Quantification	SILAC	SILAC	None	None
Run Length (min.)	95	125	125	125
# of Runs	7	7	7	7
Total MS/MS Spectra	91,753	71,498	93,288	91,038
Total Peptides	38,487	21,715	29,037	20,877
Peptide FDR	0.2%	0.2%	0.2%	0.1%
Total Phosphopeptides	29,623	17,632	19,225	15,283
Phosphopeptide FDR	0.1 %	0.2 %	0.2 %	0.1 %
Unique Phosphopeptides	8,211	6,607	6,735	6,440
Phosphoproteins	1,938	1,725	1,635	1,451
Phosphoprotein FDR	1.0 %	1.0 %	1.0 %	1.0 %
Unique Phosphorylation Sites	8,183	6,421	6,616	5,470