Published in final edited form as:

J Proteome Res. 2008 May ; 7(5): 2140–2150. doi:10.1021/pr800147u.

Motif-specific sampling of phosphoproteomes

Cristian I. Ruse, Daniel B. McClatchy, Bingwen Lu, Daniel Cociorva, Akira Motoyama, Sung Kyu Park, and John R. Yates III*

Department of Chemical Physiology/Cell Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, SR11, La Jolla, California 92037

Abstract

Phosphoproteomics, the targeted study of a subfraction of the proteome which is modified by phosphorylation, has become an indispensable tool to study cell signaling dynamics. We described a methodology that linked phosphoproteome and proteome analysis based on Ba^{2+} binding properties of amino acids. This technology selected motif-specific phosphopeptides independent of the system under analysis. MudPIT (Multidimensional Identification Technology) identified 1037 precipitated phosphopeptides from as little as 250 μg of proteins. To extend coverage of the phosphoproteome, we sampled the nuclear extract of HeLa cells with three values of Ba^{2+} ions molarity. The presence of more than 70% of identified phosphoproteins was further substantiated by their non-modified peptides. Upon isoproterenol stimulation of HEK cells, we identified an increasing number of phosphoproteins from MAPK cascades and AKAP signaling hubs. We quantified changes in both protein and phosphorylation levels of 197 phosphoproteins including a critical kinase, MAPK1. Integration of differential phosphorylation of MAPK1 with knowledge bases constructed modules that correlated well with its role as node in cross-talk of canonical pathways.

INTRODUCTION

Phosphorylation is a key event in cellular signaling for a host of physiological processes¹. Phosphorylation cascades propagate and amplify the signal in a highly regulated manner by reversible site-specific protein modification of serine, threonine and tyrosine residues. Mass spectrometry (MS) identifies phosphorylation of proteins either in purified complexes or in the entire cellular milieu. MS is a well established technique for elucidation of phosphorylation's role in the regulation of key cellular events².

The most inclusive layout of a phosphoproteome study integrates four components: enrichment of phosphoproteins or phosphopeptides, proteolytic digestion, chromatography coupled to tandem mass spectrometry (LC-MS/MS) and proteomic informatics. Enrichment strategies facilitate mapping of thousands of phosphopeptides^{3,4}. A comparison of methods for enrichment of phosphopeptides highlights the advantage of combining multiple approaches for comprehensive mapping of the phosphoproteome⁵. Reversed-phase (RP) chromatography is currently the method of choice for separating phosphopeptides prior to MS analysis. In addition to collision-activated dissociation MS/MS, scanning functions specific to phosphate neutral loss can be employed in acquiring spectra of phosphopeptides^{4,6,7}. Bioinformatics strategies validate results from database searches⁸ and increase the confidence of identified phosphorylated peptides^{8,9,10,11}. Localization of site-specific phosphorylation can be derived directly from phosphopeptide-spectrum matches¹² or through the use of specialized

*Corresponding author: John R. Yates III, The Scripps Research Institute, 10550 North Torrey Pines Rd., SR11, Department of Chemical Physiology/Cell Biology, La Jolla, CA 92037. E-mail, E-mail: jyates@scripps.edu; Phones, (858)784-8862, (858)784-8883.

bioinformatics algorithms^{9,10,13}. A variety of proteomic platforms that integrate all four components have been used to build the current knowledge of phosphoproteome¹⁴.

MudPIT (multidimensional protein identification technology) is a strategy for shotgun proteomics developed in this laboratory^{15,16}. MudPIT represents the continuum of sample preparation^{17,18}, multidimensional chromatography^{18–21}, MS²² and proteomic informatics^{23–25}. Improvements of each component are necessary to provide tailored solutions for pursuing biological questions. For example, MudPIT platform can characterize diverse systems such as cell-based cancer models²⁶ or *Caenorhabditis elegans*²⁷. Combination of methods such as isoelectric focusing and IMAC enriches for phosphopeptides prior to MudPIT analysis²⁸. MudPIT facilitates high sensitivity analysis of low copy number proteins in the yeast proteome from as little as 400µg proteins. In this paper, we seek to adapt components of the MudPIT platform for large scale analysis of phosphorylated peptides. We show that Ba²⁺/pH/acetone precipitation enables large scale identification of both phosphopeptides and associated segments of proteomes with MudPIT.

We describe a procedure for enrichment of phosphopeptides without a solid support. Absence of solid support increases the number of free coordination sites of metals available for binding peptides or solvent molecules²⁹. Ba²⁺/Ca²⁺ and ethanol precipitate tryptic casein phosphopeptides at pHs ranging from 3.5 to 8.0³⁰. We used single dimension chromatography and MS/MS to measure the binding phenomena between Ba²⁺ ions and phosphopeptides. We report for the first time that Ba²⁺ can precipitate phosphopeptides from complex mixtures. We use this protocol of Ba²⁺-based precipitations with MudPIT for analysis of nuclear extracts of HeLa cells. Our strategy selected for certain sequences of amino acids in phosphopeptides and therefore enriched for phosphorylation sites and differentiated specific phosphorylation motifs. Identification of motif-specific phosphopeptides together with unmodified peptides covered sequences of phosphoproteins with multiple peptides. Motif-specific sampling of the phosphoproteome determined both differential phosphorylation levels and protein expression levels in isoproterenol-stimulated SILAC-labeled HEK cells. Placing quantified phosphoproteins into knowledgebase networks highlighted the presence of canonical pathways associated with isoproterenol stimulation, including ERK/MAPK, cardiac beta-adrenergic and cAMP-mediated signaling.

RESULTS AND DISCUSSION

I. Use of Ba²⁺/acetone/pH precipitation for phosphopeptide identification in HeLa cells nuclear extract

We aimed to apply fractionation by affinities to Ba²⁺ ions to the analysis of phosphopeptides from complex mixtures. First, we measured the enrichment of casein phosphopeptides after precipitation with Ba²⁺/acetone at pH 8. Detection of signals from phosphopeptides over signals from unmodified peptides was optimal near 10 moles of Ba²⁺/mole of protein (Supplementary material). We estimated that optimal phosphopeptide detection occurs at 0.5 µmole of Ba²⁺ ions per mg of protein. Optimally detected phosphopeptides contained between one to five phosphate groups per peptide chain. In order to transfer Ba²⁺/acetone precipitations of phosphopeptides from simple casein mixture to complex cellular mixtures, we developed pH fractionation. Three pH values were chosen: 3.5, 4.6, and 8.0³⁰ (Figure 1). We then analyzed these individual fractions by MudPIT. We applied this approach to study protein phosphorylation in HeLa cell nuclear extracts. We found that pH fractionation of phosphopeptides from the nuclear extract by Ba²⁺/acetone depended on a delicate balance of two factors: i) an increased pool of chelators and ii) the Ba²⁺ binding capacity at each pH for phosphopeptides.

Phosphopeptides were fractionated from 1000 μ g of starting material. Three values of Ba²⁺ ions molarity were tested: 1, 6 and 7.5 μ mol Ba²⁺ ions per pH fraction. Ba²⁺/acetone precipitation was performed at pH 3.5, 4.6 and 8.0 as described in the Methods. Each pH fraction of Ba²⁺/acetone precipitates was analyzed by 12-step MudPIT as described in the Methods section. Identifications were accepted with stringent values for false positive rates after decoy database searches (Methods). At pH 3.5, the highest number of phosphopeptides (1644) was identified by precipitation with 7.5 μ mol Ba²⁺ ions (Table 1). Using 6 μ mol Ba²⁺ ions we identified 872 phosphopeptides. For both pH 4.6 and pH 8.0, the number of identified phosphopeptides saturated as molarities of Ba²⁺ ions increased. When matched MS/MS and phosphate neutral loss triggered MS3 (NL-MS3) were combined, precipitation with 7.5 μ mol Ba²⁺ ions identified the highest number of unique phosphopeptides. Therefore, the optimum for phosphopeptide detection might depend on the concentration for phosphopeptide enrichment at pH 3.5.

We estimated the enrichment of phosphopeptides from MudPIT sampling of Ba²⁺ precipitates. Relative enrichment was calculated as the ratio of copies of phosphopeptides to copies of all peptides identified only in MS/MS spectra for individual MudPIT analysis (Figure 2A). For the nuclear phosphoproteome, fractionation at pH 3.5 was considerably more efficient with 7.5 μ mol of Ba²⁺ ions than 1 or 6 μ mol of Ba²⁺ ions. At pH 4.6, the enrichment of phosphopeptides was within the same order of magnitude across three Ba²⁺ ions conditions. We observed a similar trend in enrichment of phosphopeptides at pH 8. We identified a total of 4962 phosphopeptides by adding phosphopeptide-spectrum matches in nine MudPIT experiments, as Table 1 summarizes. Of these, 8% (416) of phosphopeptides were reproducibly identified across three concentrations of Ba²⁺ ions (Figure 2B). Phosphopeptides identified in at least two Ba²⁺ ion conditions numbered 1387 (28%). Therefore, it might be advantageous to perform precipitation with multiple Ba²⁺ concentrations to increase identification of unique phosphopeptides.

Next, we tested the sensitivity of Ba²⁺/pH/acetone fractionation for identification of phosphopeptides. BaCl₂ solutions were added in increments that followed the increase in protein amounts. Respectively, we added 1.5 μ mol Ba²⁺ ions per pH fraction for 250 μ g of proteins, 3 μ mol Ba²⁺ ions per pH fraction for 500 μ g of proteins and 6 μ mol Ba²⁺ ions per pH fraction for 1000 μ g of proteins. A total of 1037 phosphopeptides were identified from 250 μ g of HeLa cells nuclear extract, 1582 phosphopeptides from 500 μ g of proteins and 1868 from 1000 μ g of proteins, as Table 2 indicates. Increasingly more phosphopeptides were identified for proportional increase of moles of Ba²⁺ ions and protein amounts. These were expected results as we simultaneously increased the available capacity of metal ions and the available chelators. We noted that the response was not entirely linear, for example doubling the amount of proteins to 500 μ g increased the number of identified phosphopeptides by only 50%.

I.2. Enrichment of motif specific phosphopeptides—Results from profiling experiments showed motif-specific sampling of the phosphoproteome. We described motif-specific sampling of the phosphoproteome using two parameters: localized phosphosites and spectral copies. We assigned phosphorylation sites with an algorithm that uses both Δ Cn score and sequence ions together with corresponding phosphate neutral losses³¹ (Supplementary Material). For SEQUEST identified phosphopeptides, a Δ Cn score greater than 0.1 translates to 90 % confidence of site localization¹². Then, we used aggregation of matched spectra per identical phosphopeptide as a semi-quantitative measure of identified peptide species³². Between 1.8 to 4.5 spectral copies per phosphopeptides were detected for 2723 phosphopeptides sampled 7.5 μ mol Ba²⁺ ions (Table 3). Sampling with 1 or 6 μ mol Ba²⁺ ions identified similar number copies per phosphopeptide match. These results suggested we consistently pooled down a segment of phosphoproteome. We further classified phosphopeptides by comparing identifications in each pH fraction. For example, 86% of the

phosphopeptides at pH 3.5 were unique, 70% at pH 4.6 and 56% at pH 8.0 from the pool of 2723 phosphopeptides (Table 4, Figure 3).

Sequence logos³³ were then used to determine if consensus sequences were common to localized phosphosites of phosphorylated peptides pooled at each pH. Binding pockets for barium ions were defined by sequences with at least five amino acids flanking the phosphosite. Figure 4A illustrates Ba²⁺-chelating phosphopeptides with a binding pocket lined with acidic (D and E) amino acids at pH 3.5. In comparison, at pH 4.6 we observed a less significant contribution of acidic residues (Figure 4B). Position +1 adjacent to phosphosite showed an increased preference towards proline residues from pH 4.6 to pH 8.0 (Figure 4C). All other surrounding amino acids were randomized at pH 8.0. At each pH, patterns for Ba²⁺-binding coincided to consensus sequences for specific kinases³³. Doubly phosphorylated peptides represented 31% of identified phosphopeptide species at pH 3.5, of which 317 had localized sites (Table 4). We considered sequence logos for three categories of adjacent phosphosites and 1 or 2 amino acids segregating phosphosites. This analysis revealed the frequency of acidic residues as well as of proline in doubly phosphorylated peptides precipitated at pH 3.5 (Figure 4D).

We further compared our data with the dataset generated by Gygi et al⁴ from HeLa cells nuclear extract and found 317 phosphopeptides in common (Supplementary Figure 2A). Sequence logos of their dataset⁴ showed proline at +1 position as the most informative amino acid (Supplementary Figure 2B). These authors' dataset⁴ resembles precipitated phosphopeptides at pH 4.6 and 8.0 (Figure 4B,C) rather than pH 3.5. Therefore, Ba²⁺/acetone precipitated phosphopeptides at pH 3.5 have unique patterns as compared with the SCX fractionation method⁴. Consequently, our approach sequentially enriched from the same source phosphopeptides with both unique as well as previously described patterns.

We speculate that binding of motif-specific phosphopeptides to barium ions depends both on the capacity of metal ions and properties on phosphate groups. Several intrinsic properties of phosphates in peptide sequences might influence cooperativity (Hill coefficient) of metal ion-binding: i) exposure of binding residues to solvent^{34,35}, ii) orientation of phosphate groups in clusters of phosphoamino acids^{36,37} and iii) pKa values of phosphate groups in their structural context^{38,39}. For example, intra or inter-chain interaction of phosphorylated amino acids with basic or acidic residues decreases or increases pKa of phosphoserine, respectively.³⁸ In addition, acidic side chains and backbone carbonyls might contribute to chelation of metal ions⁴⁰. The binding capacity of metal ions might influence the outcome of phosphopeptide precipitations. For example, barium ions have a coordination number of 9 as compared to 7 for calcium ions⁴¹. In this study, barium salts were used and not calcium salts as proposed for characterization of caseins³⁰. Last but not least, other pH values might be used to dissect motif-specific phosphoproteomes. In this study, we employed pH values determined in previous studies³⁰.

II. Analysis of pooled MudPIT datasets for more confident identification of phosphoproteins

Reversible addition of phosphate to proteins propagates cellular signals through formation of functional complexes in cellular structures. Changes in activity of signaling proteins are determined by comparing variation in phosphorylation signals relative to the abundance of total proteins. Identification of both phosphopeptides and peptides from phosphoproteins in a single analysis bridges the quantitative measurements of changes in signaling proteins. Therefore, we integrated identified phosphopeptides and peptides into proteins or group of proteins.

Identification of single phosphopeptide hits was instrumental in evaluating the efficacy and characteristics of Ba²⁺/pH/acetone precipitates to enrich phosphopeptides. For this purpose,

SEQUEST database searches and error-tolerant filtering of peptide-spectrum matches were performed for individual MudPIT runs. Beside identification of phosphopeptides, fractionation of complex mixtures with Ba^{2+} /pH/acetone identified a considerable number of unmodified peptides. To obtain more confident phosphoprotein identifications, we pooled three MudPIT datasets belonging to sampling with 1 μmol of Ba^{2+} ions. When filtering with at least one non-modified peptide match per protein which had a phosphopeptide match, we identified 812 phosphoproteins. Approximately 70% of phosphoproteins were identified by non-phosphorylated peptides, too. For example, sampling with 7.5 μmol of Ba^{2+} ions identified 2723 phosphopeptides corresponding to 1191 phosphoproteins. Analysis of corresponding MudPIT runs retained 872 phosphoproteins (73%) when requiring the presence of a non-phosphorylated peptide. Ba^{2+} /pH/acetone strategy enriches peptides with motif-specific sequences. For 50% of phosphopeptides there were identical peptide sequence matches in phosphoproteins identified by both peptides and phosphopeptides. This aspect of the process confirmed phosphopeptide identifications by identical sequences without phosphatase treatment.

Our MudPIT-based approach added confidence in identification of phosphoproteins by including both peptide and phosphopeptide matches. Thus, motif-specific enrichment coupled with MudPIT identified phosphoproteins more confidently than double enrichment techniques. In double enrichment strategies, fractionation of a pool of phosphopeptides is followed by almost complete isolation with a complementary enrichment method. Hence, double enrichment strategies for phosphopeptides lead to identifications of proteins that often rely on single phosphopeptide hits. Currently, identification of cell signaling events by proteomics relies on measuring either protein expression levels (protein immunoprecipitation or ICAT⁴² strategies) or changes in phosphosites (double enrichment techniques).

Reports of double enrichment techniques describe considerable numbers of phosphorylation sites: 6600 (from at least 15 mg of proteins)⁴³ and 5635 (from 10 mg of liver homogenate)⁴⁴. Versatile combinations of inhibitors of phosphatases with complementary isolation strategies for phosphopeptides establish the presence of 10,000 phosphorylation sites in a *Drosophila melanogaster* cell line (from 280mg of peptides)¹². Our strategy identified 2723 phosphopeptides containing more than 3000 phosphorylation sites using 1 mg of HeLa cells nuclear extract.

III. Application of Ba^{2+} -based enrichment to identification of phosphopeptides from whole cell lysates

Besides mapping homeostatic/basal phosphorylation of proteins, enrichment methods should ultimately lead to identification of known signaling phosphoproteins given a background of existing phosphorylation signals. For this purpose, we tested the capability of Ba^{2+} /acetone/pH fractionation protocol to map signaling phosphoproteins from whole cell lysates of 5 minute isoproterenol-stimulated HEK cells overexpressing beta 1 adrenergic receptor. Isoproterenol is an agonist of beta-adrenergic receptor. We applied Ba^{2+} /pH/acetone precipitation with 1.0 μmole of Ba^{2+} ions / mg of proteins for each pH fraction. As before, three pH fractions of Ba^{2+} /acetone precipitates were analyzed by individual MudPIT runs on a linear ion trap mass spectrometer.

Table 5 summarizes identification of phosphopeptides from individual MudPIT runs per each pH fraction treated with 1.0 μmole of Ba^{2+} ions / mg of proteins. Enrichment levels were similar as compared with the HeLa cell nuclear extract. Also, we obtained detailed site-localization analysis of monophosphorylated peptides from each individual pH fraction (Supplementary Figures 3A–C). Figure 5A–C illustrates sampling of motif-specific phosphoproteome by Ba^{2+} -chelating properties from 5 minute isoproterenol-stimulated HEK cells overexpressing beta 1 adrenergic receptor. We identified similar sequence motifs surrounding the

phosphorylation sites from HEK cells and HeLa cells nuclear extract at corresponding pH values. As before, sequential precipitations from the same source identified similar patterns for Ba²⁺-binding phosphopeptides. Therefore, we show that motif-specific sampling of the phosphoproteome is independent of the sample under analysis.

Description of phosphoproteomes of isoproterenol-stimulated and control HEK cells demonstrated the capability of Ba²⁺-based protocol for analyzing whole cell lysates. Cumulative individual analyses of MudPIT runs identified 823 phosphoproteins and 493 phosphoproteins in stimulated and control cells, respectively. We compared phosphorylated kinases and AKAPs (A-kinase anchoring proteins) between control and isoproterenol-stimulated HEK cells (Supplementary table 1). Following isoproterenol stimulation, four kinases associated with MAPK cascades were detected as compared to only one in the control sample. In addition, three phosphorylated AKAPs were detected after agonist stimulation as compared with only one in the control. AKAPs are prototypical examples of scaffold proteins that integrate multiprotein signaling networks: cAMP, calcium and MAPK^{45–47}. In cardiac myocytes, AKAPs regulate both substrate phosphorylation by PKA and cellular functions upon beta-adrenergic stimulation⁴⁸. MAPK1 activation has been shown to be mediated by activation of beta-adrenergic receptor^{49,50}. Furthermore, we tested our experimental platform for collection of quantitative measurements during isoproterenol stimulation of HEK cells.

Quantitative analysis of Ba²⁺/acetone precipitated fractions—We used SILAC for quantitative description of phosphoproteins after isoproterenol stimulation. SILAC (stable-isotope labeling of amino acids in cell culture)⁵¹ is one of the metabolic labeling techniques for protein labeling⁵². It involves growing two populations of cells in media that contain either “Light” or “Heavy” amino acids⁵³. Light and heavy amino acids differ only by the mass of the replaced isotope, i.e. ¹⁵N instead of ¹⁴N. In this work, we use labeled ¹³C₆-Lys (mass shift 6 Da) and ¹³C₆, ¹⁵N₄-Arg (mass shift 10 Da). We quantified the ratio of signals in mass spectra of SILAC-labeled peptides with a linear ion trap–orbitrap hybrid mass spectrometer⁵⁴. A previous study from our laboratory describes the high efficiency of peptide quantification when using a hybrid linear ion trap–orbitrap mass spectrometer⁵⁵. We combined Ba²⁺-based fractionation and MudPIT for quantitative measurements of SILAC-labeled peptides on a hybrid linear ion trap–orbitrap mass spectrometer.

Quantitative analyses of Ba²⁺/pH/acetone precipitated phosphopeptides and supernatant fractions were performed by eight-step MudPIT on a hybrid linear ion trap–orbitrap mass spectrometer (Methods). Two time points were compared in SILAC labeled HEK cells overexpressing beta-adrenergic receptor: 30 second and 2 minute isoproterenol stimulation. We used 500 µg of proteins per stimulation condition, i.e. a total of 1mg of proteins. By choosing to compare two non-zero time points we minimized the frequency of singletons, all-or-none events of phosphosite activation⁵⁶.

We quantified expression levels and phosphorylation levels for a total of 197 unique proteins. Our quantified peptides integrated two types of information: protein expression levels and phosphorylation levels. Protein expression levels were derived from 2066 peptide ratios. Phosphorylation levels were derived from 398 phosphopeptide ratios (Figure 6). Protein expression levels were calculated from both peptides and phosphopeptides signals. Up- and down-regulation were considered significant for protein expression levels with a ratio of 2 or 0.5 respectively. If phosphopeptide signals were to be subtracted from protein ratios, expression levels would distribute closer to a ratio of 1. Therefore, protein expression levels were essentially unchanged during 90 second stimulation with isoproterenol.

We imposed a different filter for detecting quantitative differences in phosphorylation of peptides. For a ratio more than 2 standard deviations away from the mean, we considered

peptides to be differentially phosphorylated (ratio < 0.33 and ratio > 2.66). Therefore, four proteins showed upregulation of phosphorylation levels while seven proteins had downregulated phosphorylation (Table 6). Of these, nine phosphoproteins were previously detected as phosphorylated at the sites highlighted in Table 6. The highest increase in phosphorylation levels (16 fold) belonged to phosphorylation of MAPK1. Thus, we detected changes in phosphorylation levels of MAPK1, a critical kinase for cross-talk with β -adrenergic pathway.

Proteomics experiments contribute to a system-wide perspective by substantiating scale-free networks and their composing modules^{57,58}. Conversely, overlaying quantified phosphoproteins onto pre-defined pathways and signaling modules decodes information from phosphoproteomics experiments⁵⁹. Ingenuity Knowledgebase analyzes such collections of identified proteins using curated information from pathways databases⁶⁰. Therefore, we interrogated the quantified phosphoproteins and their expression levels with Ingenuity Pathway Analysis (IPA). We added two phosphoproteins to the list of phosphoproteins with quantified expression levels: Splice Isoform 1C of Mitogen-activated protein kinase kinase 7 and Stromal interaction molecule 1 precursor. These two proteins were identified by both phosphopeptides and peptides, but were quantified only by signals from phosphopeptides. Using IPA analyses (Methods), we highlighted modules that might be activated in response to isoproterenol stimulation.

IPA analysis of quantified phosphoproteins identified 10 modules with 6 or more focused genes (Supplementary Data). Interestingly, within module “3” we distinguished molecules from the following canonical pathways: ERK/MAPK (8 molecules), cardiac beta-adrenergic (4), calcium signaling (4), insulin receptor signaling (4) and cAMP-mediated signaling (3) (Figure 7). These results further supported the established cross-talk between cAMP-mediated signaling and the ERK pathway⁶¹. The module centered around MAPK (“3”) is connected by few links to eight (from nine possible) other modules of phosphoproteins (Supplementary Data). This module was connected by a link to four other modules (“1”, “4”, “5”, “6”, “7”) and two links to three modules (“1”, “8” and “10”). The vital role of MAPK activation in the early response to agonist stimulation of beta adrenergic receptors has been shown in previous studies^{49,50}. Thus, quantification of signals from 500 μ g of proteins per each stimulation condition confirmed known phenomena of signaling crosstalk.

We showed that combination of Ba^{2+} -based fractionation and MudPIT is suitable for measuring quantitative changes of SILAC-labeled peptides and phosphopeptides. Protein expression levels and phosphorylation levels determined by our strategy were assembled in informative networks. The experimental outcomes of our platform contribute to the global study of systems such as signaling networks. Potentially activated phospho-networks constitute a starting point for biological validation of mechanisms regulating cell signaling.

CONCLUSIONS

Most large scale phosphoproteomic analyses require significant amounts of starting proteins. We identified 1037 phosphopeptides from as little as 250 μ g of HeLa cells nuclear extracts. We described an alternative experimental platform for data collection of motif-specific phosphoproteomes. Sequence logos showed the selection of phosphopeptides based on amino acid compositions. Independent of sample, Ba^{2+} -binding motifs preferentially precipitated at pH 3.5 matched those of potential substrates for acidophilic kinases. At pH 4.6 and 8.0, Ba^{2+} -binding motifs matched motifs of potential substrates for proline-directed kinases (Figure 4A–C). Most of phosphopeptides precipitated in Ba^{2+} -complexes at pH 3.5 (Figure 3). Fraction pH 3.5 contained a significant number of multiply phosphorylated peptides as well (Table 4). Fractionation of proteomes with a Ba^{2+} -based protocol complemented analysis of

phosphopeptides with identification of unmodified peptides. More than 70% of phosphoproteins initially identified only by phosphopeptides could be confirmed with unmodified peptides, independent of sample. Consequently, two thirds of identified phosphoproteins from Ba^{2+} /acetone precipitates could be confirmed by unmodified peptides.

We used SILAC-labeled HEK cells to quantify protein expression levels and phosphorylation levels in response to isoproterenol from a 1mg of proteins. Crosstalk of cAMP and MAPK signaling pathways⁶¹ and β AR-mediated activation of ERK 1/2^{49,50} are well documented facts. Our proteomic methodology highlighted the possibility of both outcomes over a 90 second stimulation interval (Table 6 and Figure 7).

In conclusion, Ba^{2+} /pH/acetone precipitation is a strategy for large-scale analysis of motif specific phosphoproteomes by MudPIT from low amounts of proteins. More than two thirds of identified phosphoproteomes were confirmed by unmodified peptides. Ba^{2+} -based fractionation with MudPIT allowed for simultaneous quantitative measurements of both protein expression levels and phosphorylation levels.

METHODS

Human embryonic kidney 293 cells overexpressing the human beta1-adrenergic (β 1AR-HEK293) cells were kindly provided by Robert J. Lefkowitz at Department of Biochemistry, Duke University. All cell culture reagents were from SILAC Protein ID and Quantitation Kit (Invitrogen, Carlsbad, CA) unless otherwise noted. All other materials were purchased from Sigma, St. Louis, MO unless otherwise noted.

Cell Culture

The β 1AR-HEK293 cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 5% dialyzed fetal bovine serum, 2mM glutamax, 1% penicillin/streptomycin, and 200ug/ml G418 in a 5% CO_2 37°C incubator. Cells were cultured for at least 14 days in 100mm² dishes in DMEM without lysine and arginine supplemented with either isotopically labeled ("Heavy") arginine ($^{13}\text{C}_6$ - $^{15}\text{N}_4$ -Arg) and lysine ($^{13}\text{C}_6$ -Lys) or normal ("Light") arginine and lysine. Experiments were performed when the plates were 80% confluent. Cells were incubated in DMEM without serum for 4 hours before the drug treatment, and then, 10 μ M isoproterenol (Sigma, St. Louis, MO) in DMEM was added to cells for 30 seconds (Light), 2 (Heavy) or 5 min (Light). The cells were then quickly washed with ice cold PBS with phosphatase inhibitor cocktail (Calbiochem, San Diego, CA). Trypsin solution containing 0.05% EDTA was briefly added to the cells, and then the cells were removed from the plate with 10ml of PBS with phosphatase inhibitors. The cells were centrifuged for 5 minutes at 500 $\times$ g. The cell pellet was resuspended in lysis buffer (10mM Tris, 0.1mM EDTA, 0.5% NP-40 with protease (Roche Applied Science, Indianapolis, IN) and phosphatase inhibitors. BCA protein assay (Pierce, Rockford, IL) was performed and samples were placed in -80°C.

Digestion and Ba^{2+} /acetone precipitation of complex mixtures

Nuclear extract from HeLa cells (BioVision Inc.) was extracted with chloroform/methanol and digested with trypsin (1:50) in 50% MeOH / 100mM TRIS pH 7.6, at 37C overnight. Digested peptides were clarified by centrifugation. For HEK cells, 1 mg of proteins from either control, 5 min stimulated HEK cells (Light) or a mixture of 1:1 of Light (0.5 min stimulated)/Heavy (2min stimulated) were used. In both cases, 1 mg of proteins from β 1AR HEK whole cell lysates were processed by chloroform/methanol extraction and digested as above.

A sequential protocol was developed to precipitate phosphopeptides by Ba^{2+} , where supernatant from one fraction was further used for phosphopeptide enrichment by increasing the pH. After drying, samples were resuspended in 100mM TRIS pH 7.6 and pH adjusted to 3.5. For nuclear extract samples, an aliquot of $x = (1, 6 \text{ or } 7.5) \mu\text{mol Ba}^{2+}$ ions was added per 1mg starting total protein and samples were mixed. Ice-cold acetone was added to 1:1 by volume and samples were mixed. After centrifugation at 14,000 rpm for 15 min at 4°C, pellet was stored for analysis and supernatant was dried. Dried supernatant was resuspended in 100mM TRIS pH 7.6 and pH was adjusted to 4.6. Precipitation with Ba^{2+} was performed as above and pellet saved for analysis while the supernatant was dried. After reconstitution in 100mM TRIS pH 7.6, pH of solution was adjusted to 8.0. Precipitation with Ba^{2+} was performed as above and pellet stored for analysis. A total of three pellets were obtained: from pH 3.5, pH 4.6 and pH 8.0. Same sequential protocol was applied to HEK cells with an aliquot of 1 $\mu\text{mol Ba}^{2+}$ ions added per each pH fraction.

Mass spectrometry analysis

MudPIT—The buffer solutions were 5% acetonitrile/0.1% formic acid (buffer A), 80% acetonitrile/0.1% formic acid (buffer B), and 500 mM ammonium acetate/5% acetonitrile/0.1% formic acid (buffer C).

Samples were pressure-loaded onto a 250- μm i.d. fused silica capillary column containing 3 cm of 5- μm Aqua C₁₈ material (Phenomenex, Ventura, CA) followed by 3 cm of 5- μm Partisphere strong cation exchanger (Whatman, Clifton, NJ) and capped with a 2 μm filtered union. The biphasic column was washed with buffer A. The biphasic column was then connected to an analytical column of a 100- μm i.d. capillary with a 5- μm pulled tip and packed with 12~13 cm of 3- μm Aqua C₁₈ material (Phenomenex, Ventura, CA).

The column was placed inline with an Agilent 1100 quaternary HPLC. Step 1 consisted of 15 min of 100% A followed by a gradient of 80 min from 0 to 55 % B, reversal to 100% A in 2 min and 3 min of re-equilibration with 100% A. For each salt step, we applied a 2 min delivery of X% buffer C, 10 min of 100% A and then peptides were separated by a gradient of 85 min from 0 to 55% B. The 2 min buffer C percentages (X) were 5, 10, 15, 20, 25, 30, 35, 40, 45 and 65%. The final step, the gradient contained: 3 min of 100% buffer A, 20 min of 100% buffer C, a 10 min gradient from 0–15% buffer B, and a 107 min gradient from 15–70% buffer B.

Peptides were electrosprayed directly into a LTQ mass spectrometer (ThermoFinnigan, Palo Alto, CA) and analyzed with a data dependent neutral loss routine.

For Ba^{2+} precipitated pH fractions of tryptic digested whole cell lysates of $\beta 1\text{AR}$ HEK (Light - 0.5 min / Heavy - 2 min), the entire split-column (biphasic column–union–analytical column) was placed inline with an Eksigent nano-flow HPLC pump (Eksigent) directly in front of the heated capillary opening of an LTQ-Orbitrap hybrid mass spectrometer (ThermoElectron). The flow rate was set at 300 nL/min (splitless). Eight salt steps were performed with following buffer C concentrations: 0%, 5%, 10%, 15%, 20 %, 25%, 45% and 100% at 5 min durations for each step. For each salt step, peptides were then separated using a linear gradient of 0.5% per min for 90 min.

Peptides eluted from the microcapillary fritless column were directly electrosprayed into an LTQ-Orbitrap hybrid spectrometer with the application of a distal 2.4 kV spray voltage. A cycle of one full-scan mass spectrum (400–1700 m/z) with $R = 60,000$ (by Orbitrap) followed by 8 data-dependent MS/MS spectra at a 35% normalized collision energy (by LTQ) was repeated continuously throughout each step of the multidimensional separation. LTQ-Orbitrap

mass spectrometer was programmed to conduct neutral loss triggered data-dependent MS/MS for the peptides that produced 49 or 98 neutral loss from their precursor ions.

Data analysis

Database searches were conducted with SEQUEST⁶² against IPI human database complemented with its reversed protein sequences (version 3.04_03-07-2005). Parameters for MS2 spectra searches were as follows: tryptic specificity, 2 missed cleavage sites, differential modifications STY = 80 (for linear ion trap data) and up to 3 modifications per peptide. Peptide mass tolerance was allowed to be 6 amu to account for effective database size⁶³ in phosphorylation searches⁹. When precursor mass was determined with high mass accuracy in full scan, MS2 spectra were searched against IPI human database complemented with its reversed protein sequences (version 3.04_03-07-2005) with no enzyme specificity and a peptide mass tolerance of 50 ppm. Differential modifications SEQUEST⁶² searches included STY = 79.966332 for linear ion trap-orbitrap data. Parameters for MS3 spectra searches were as follows: no enzyme specificity, differential modifications searches included: ST = -18 and up to 2 modifications per peptide and peptide mass tolerance of 6 amu.

Database searches were followed by linear discriminant analysis based on XCorr and DCn between forward and database hits, similar to published references⁶⁴. Peptide hits were accepted for 95% or higher confidence reported by linear discriminant analysis. In addition, we removed a) peptides shorter than 7 aa and b) outliers even if they pass 95% confidence. Finally, we requested Zscore to be higher than 3 for each peptide identification. Z-score is a more general measure of the cross-correlation score distribution⁶⁵. For analysis of orbitrap data we used mass deviation (less than 10 ppm) as an additional filtering parameter.

Quantification of phosphopeptides—In house developed software (Census, manuscript in preparation)⁶⁶ was used for peptide quantification. The correlation coefficient (r) - a measure of the closeness of fit between the data points of the unlabeled and labeled ion chromatograms - was set to 0.7²⁴. Standard deviations are calculated for all proteins using their respective peptide ratio measurements. We then use Grubbs test ($P < 0.05$) to remove outliers peptide ratios.

Network identification—We analyzed quantified phosphoproteins using Ingenuity Pathway Analysis (IPA)⁶⁷. IPI identifiers were used and mapped to corresponding gene objects in Ingenuity Pathways Knowledge Base. IPA algorithms generated networks (modules) from connectivity of these genes (focus genes)⁶⁸. A score of > 3 for a module (network) was considered significant ($p < 0.001$)⁶⁹.

Additional information—Proportional Venn Diagrams were generated with VennMaster. Sequence logos were plotted with MATLAB Bioinformatics Toolbox.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENT

The authors acknowledge support from NIH 5R01MH067880-02 and NIH RR 11823-10. CIR would like to thank the following colleagues: Dr. Emily Chen for helpful suggestions and critical reading of the manuscript, Dr. Greg Cantin for reading the manuscript and discussions on phosphoproteome analysis, Dr. Daniel Salomon and Dr. Sunil Kurian for suggestions on using Ingenuity Pathway Analysis and Dr. James Wohschlegel for help with preliminary experiments for yeast phosphoproteome quantification. We thank members of Yates laboratory for discussions.

References

1. Hunter T. Signaling--2000 and beyond. *Cell* 2000;100:113–127. [PubMed: 10647936]
2. Loyet KM, Stults JT, Arnott D. Mass spectrometric contributions to the practice of phosphorylation site mapping through 2003: a literature review. *Mol Cell Proteomics* 2005;4:235–245. [PubMed: 15640519]
3. Ficarro SB, et al. Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat Biotechnol* 2002;20:301–305. [PubMed: 11875433]
4. Beausoleil SA, et al. Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci U S A* 2004;101:12130–12135. [PubMed: 15302935]
5. Bodenmiller B, Mueller LN, Mueller M, Domon B, Aebersold R. Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat Methods* 2007;4:231–237. [PubMed: 17293869]
6. 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;76:3590–3598. [PubMed: 15228329]
7. Wolschin F, Lehmann U, Glinski M, Weckwerth W. An integrated strategy for identification and relative quantification of site-specific protein phosphorylation using liquid chromatography coupled to MS2/MS3. *Rapid Commun Mass Spectrom* 2005;19:3626–3632. [PubMed: 16287031]
8. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* 2007;4:207–214. [PubMed: 17327847]
9. 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;24:1285–1292. [PubMed: 16964243]
10. Bodenmiller B, et al. An integrated chemical, mass spectrometric and computational strategy for (quantitative) phosphoproteomics: application to *Drosophila melanogaster* Kc167 cells. *Mol Biosyst* 2007;3:275–286. [PubMed: 17372656]
11. Lu B, Ruse C, Xu T, Park SK, Yates J 3rd. Automatic validation of phosphopeptide identifications from tandem mass spectra. *Anal Chem* 2007;79:1301–1310. [PubMed: 17297928]
12. Bodenmiller B, et al. PhosphoPep--a phosphoproteome resource for systems biology research in *Drosophila* Kc167 cells. *Mol Syst Biol* 2007;3:139. [PubMed: 17940529]
13. Schlosser A, Vanselow JT, Kramer A. Comprehensive Phosphorylation Site Analysis of Individual Phosphoproteins Applying Scoring Schemes for MS/MS Data. *Anal Chem* 2007;79:7439–7449. [PubMed: 17718535]
14. Collins MO, Yu L, Choudhary JS. Analysis of protein phosphorylation on a proteome-scale. *Proteomics* 2007;7:2751–2768. [PubMed: 17703509]
15. Link AJ, et al. Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol* 1999;17:676–682. [PubMed: 10404161]
16. Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 2001;19:242–247. [PubMed: 11231557]
17. Wu CC, MacCoss MJ, Howell KE, Yates JR 3rd. A method for the comprehensive proteomic analysis of membrane proteins. *Nat Biotechnol* 2003;21:532–538. [PubMed: 12692561]
18. MacCoss MJ, et al. Shotgun identification of protein modifications from protein complexes and lens tissue. *Proc Natl Acad Sci U S A* 2002;99:7900–7905. [PubMed: 12060738]
19. McDonald WH, Ohi R, Miyamoto DT, Mitchison TJ, Yates JR. Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single-dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT. *International Journal of Mass Spectrometry* 2002;219:245–251.
20. Motoyama A, Venable JD, Ruse CI, Yates JR 3rd. Automated ultrahigh-pressure multidimensional protein identification technology (UHP-MudPIT) for improved peptide identification of proteomic samples. *Anal Chem* 2006;78:5109–5118. [PubMed: 16841936]
21. Motoyama A, Xu T, Ruse CI, Wohlschlegel JA, Yates JR 3rd. Anion and Cation Mixed-Bed Ion Exchange for Enhanced Multidimensional Separations of Peptides and Phosphopeptides. *Anal Chem*. 2007

22. Venable JD, Dong MQ, Wohlschlegel J, Dillin A, Yates JR. Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat Methods* 2004;1:39–45. [PubMed: 15782151]
23. MacCoss MJ, Wu CC, Yates JR 3rd. Probability-based validation of protein identifications using a modified SEQUEST algorithm. *Anal Chem* (3rd) 2002;74:5593–5599. [PubMed: 12433093]
24. MacCoss MJ, Wu CC, Liu H, Sadygov R, Yates JR 3rd. A correlation algorithm for the automated quantitative analysis of shotgun proteomics data. *Anal Chem* (3rd) 2003;75:6912–6921. [PubMed: 14670053]
25. Sadygov RG, Liu H, Yates JR. Statistical models for protein validation using tandem mass spectral data and protein amino acid sequence databases. *Anal Chem* 2004;76:1664–1671. [PubMed: 15018565]
26. Chen EI, et al. Adaptation of energy metabolism in breast cancer brain metastases. *Cancer Res* 2007;67:1472–1486. [PubMed: 17308085]
27. Dong MQ, et al. Quantitative mass spectrometry identifies insulin signaling targets in *C. elegans*. *Science* 2007;317:660–663. [PubMed: 17673661]
28. Cantin GT, Venable JD, Cociorva D, Yates JR 3rd. Quantitative phosphoproteomic analysis of the tumor necrosis factor pathway. *J Proteome Res* 2006;5:127–134. [PubMed: 16396503]
29. Ueda EK, Gout PW, Morganti L. Current and prospective applications of metal ion-protein binding. *J Chromatogr A* 2003;988:1–23. [PubMed: 12647817]
30. Reynolds EC, Riley PF, Adamson NJ. A selective precipitation purification procedure for multiple phosphoserine-containing peptides and methods for their identification. *Anal Biochem* 1994;217:277–284. [PubMed: 8203756]
31. Lu, Bea. Automated assignment of phosphorylation sites from tandem mass spectra. (in preparation)
32. Liu H, Sadygov RG, Yates JR 3rd. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* 2004;76:4193–4201. [PubMed: 15253663]
33. Kreegipuu A, Blom N, Brunak S, Jarv J. Statistical analysis of protein kinase specificity determinants. *FEBS Lett* 1998;430:45–50. [PubMed: 9678592]
34. Finley N, et al. NMR analysis of cardiac troponin C-troponin I complexes: effects of phosphorylation. *FEBS Lett* 1999;453:107–112. [PubMed: 10403385]
35. Gaponenko V, et al. Effects of troponin I phosphorylation on conformational exchange in the regulatory domain of cardiac troponin C. *J Biol Chem* 1999;274:16681–16684. [PubMed: 10358006]
36. Keane NE, et al. The ordered phosphorylation of cardiac troponin I by the cAMP-dependent protein kinase--structural consequences and functional implications. *Eur J Biochem* 1997;248:329–337. [PubMed: 9346285]
37. Quirk PG, Patchell VB, Gao Y, Levine BA, Perry SV. Sequential phosphorylation of adjacent serine residues on the N-terminal region of cardiac troponin-I: structure-activity implications of ordered phosphorylation. *FEBS Lett* 1995;370:175–178. [PubMed: 7656971]
38. Jaquet K, Korte K, Schnackerz K, Vyska K, Heilmeyer LM Jr. Characterization of the cardiac troponin I phosphorylation domain by ³¹P nuclear magnetic resonance spectroscopy. *Biochemistry* 1993;32:13873–13878. [PubMed: 8268162]
39. Reiffert SU, Jaquet K, Heilmeyer LM Jr, Ritchie MD, Geeves MA. Bisphosphorylation of cardiac troponin I modulates the Ca(2+)-dependent binding of myosin subfragment S1 to reconstituted thin filaments. *FEBS Lett* 1996;384:43–47. [PubMed: 8797800]
40. Anderson PJ, Brooks CL, Berliner LJ. Functional identification of calcium binding residues in bovine alpha-lactalbumin. *Biochemistry* 1997;36:11648–11654. [PubMed: 9305954]
41. Murugavel R, Karambelkar VV, Anantharaman G, Walawalkar MG. Synthesis, spectral characterization, and structural studies of 2-aminobenzoate complexes of divalent alkaline earth metal ions: X-ray crystal structures of [Ca(2-aba)(2)(OH2)(3)](infinity), [{Sr(2-aba)(2)(OH2)(2)} center dot H2O](infinity), and [Ba(2-aba)(2)(OH2)](infinity) (2-abaH=2-NH2C6H4COOH). *Inorganic Chemistry* 2000;39:1381–1390. [PubMed: 12526440]
42. Gygi SP, et al. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 1999;17:994–999. [PubMed: 10504701]
43. Olsen JV, et al. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 2006;127:635–648. [PubMed: 17081983]

44. Villen J, Beausoleil SA, Gerber SA, Gygi SP. Large-scale phosphorylation analysis of mouse liver. *Proc Natl Acad Sci U S A* 2007;104:1488–1493. [PubMed: 17242355]
45. Ruehr ML, Russell MA, Bond M. A-kinase anchoring protein targeting of protein kinase A in the heart. *J Mol Cell Cardiol* 2004;37:653–665. [PubMed: 15350838]
46. Wong W, Scott JD. AKAP signalling complexes: focal points in space and time. *Nat Rev Mol Cell Biol* 2004;5:959–970. [PubMed: 15573134]
47. Dodge-Kafka KL, Kapiloff MS. The mAKAP signaling complex: integration of cAMP, calcium, and MAP kinase signaling pathways. *Eur J Cell Biol* 2006;85:593–602. [PubMed: 16460834]
48. Fink MA, et al. AKAP-mediated targeting of protein kinase a regulates contractility in cardiac myocytes. *Circ Res* 2001;88:291–297. [PubMed: 11179196]
49. Martin NP, Whalen EJ, Zamah MA, Pierce KL, Lefkowitz RJ. PKA-mediated phosphorylation of the beta1-adrenergic receptor promotes Gs/Gi switching. *Cell Signal* 2004;16:1397–1403. [PubMed: 15381255]
50. Shenoy SK, et al. beta-arrestin-dependent, G protein-independent ERK1/2 activation by the beta2 adrenergic receptor. *J Biol Chem* 2006;281:1261–1273. [PubMed: 16280323]
51. Mann M. Functional and quantitative proteomics using SILAC. *Nat Rev Mol Cell Biol* 2006;7:952–958. [PubMed: 17139335]
52. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* 2007;389:1017–1031. [PubMed: 17668192]
53. Ong SE, Mann M. Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* 2005;1:252–262. [PubMed: 16408053]
54. Lau KW, Jones AR, Swainston N, Siepen JA, Hubbard SJ. Capture and analysis of quantitative proteomic data. *Proteomics* 2007;7:2787–2799. [PubMed: 17640002]
55. Venable JD, Wohlschlegel J, McClatchy DB, Park SK, Yates JR 3rd. Relative quantification of stable isotope labeled peptides using a linear ion trap-Orbitrap hybrid mass spectrometer. *Anal Chem* 2007;79:3056–3064. [PubMed: 17367114]
56. Gruhler A, et al. Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol Cell Proteomics* 2005;4:310–327. [PubMed: 15665377]
57. Rual JF, et al. Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 2005;437:1173–1178. [PubMed: 16189514]
58. Drake TA, Ping P. Thematic review series: systems biology approaches to metabolic and cardiovascular disorders. Proteomics approaches to the systems biology of cardiovascular diseases. *J Lipid Res* 2007;48:1–8. [PubMed: 17065662]
59. Matsuoka S, et al. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* 2007;316:1160–1166. [PubMed: 17525332]
60. Calvano SE, et al. A network-based analysis of systemic inflammation in humans. *Nature* 2005;437:1032–1037. [PubMed: 16136080]
61. Stork PJ, Schmitt JM. Crosstalk between cAMP and MAP kinase signaling in the regulation of cell proliferation. *Trends Cell Biol* 2002;12:258–266. [PubMed: 12074885]
62. Eng JK, McCormack AL, Yates JR. An Approach to Correlate Tandem Mass-Spectral Data of Peptides with Amino-Acid-Sequences in a Protein Database. *Journal of the American Society for Mass Spectrometry* 1994;5:976–989.
63. Moore RE, Young MK, Lee TD. Qscore: an algorithm for evaluating SEQUEST database search results. *J Am Soc Mass Spectrom* 2002;13:378–386. [PubMed: 11951976]
64. Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 2002;74:5383–5392. [PubMed: 12403597]
65. Xu T, et al. ProLuCID, a fast and sensitive tandem mass spectra-based protein identification program. *Molecular & Cellular Proteomics* 2006;5:S174–S174.
66. Park SK, Venable JD, Xu T, Liao L, Yates JR. A tool for quantitative analysis of high-throughput mass spectrometry data. *Molecular & Cellular Proteomics* 2006;5:S199–S199.
67. 6741986, U.S.p. Method and system for performing information extraction and quality control for a knowledge base

68. Li RW, Capuco AV. Canonical pathways and networks regulated by estrogen in the bovine mammary gland. *Funct Integr Genomics*. 2007
69. Raponi M, et al. Microarray analysis reveals genetic pathways modulated by tipifarnib in acute myeloid leukemia. *BMC Cancer* 2004;4:56. [PubMed: 15329151]

Figure 1.

Solution pH is initially adjusted to 3.5 and upon addition of Ba²⁺/acetone, a first set of phosphopeptides is precipitated. Dried supernatant is solubilized and adjusted to pH 4.6 and precipitated again. The remaining supernatant is dried, solubilized, solution pH adjusted to 8.0 and precipitated again. Individual fractions are independently analyzed using MudPIT. Data were acquired using our MudPIT platform on a linear ion trap mass spectrometer or a hybrid linear ion trap orbitrap mass spectrometer.

Figure 2A

Figure 2B

Figure 2. Identification of phosphoproteome sampled with different Ba²⁺ ions concentrations and analyzed MudPIT. A. Enrichment of phosphopeptides per pH fraction for three different Ba²⁺ molarities. B. Overlapping segments of phosphoproteome precipitated by different Ba²⁺ molarities: 7.5 μmol (yellow), 6 μmol (red) and 1 μmol (blue).

Figure 3. Proportional Venn diagram describing overlap of identified phosphopeptides among pH 3.5, 4.6 and 8.0 fraction.

Figure 4A - pH 3.5

Figure 4B - pH 4.6

Figure 4C – pH 8.0

Figure 4D

Figure 4.

Sequence logos of phosphopeptides from nuclear extract of HeLa cells: pH 3.5 (A), pH 4.6 (B) and pH 8.0 (C). Sequence logos were built for amino acids lining of Ba^{2+} ions binding pockets (−5 to +5 AAs). Sequence logos statistically evaluate the binding properties of a population of phosphopeptides by measuring the information (in bits) required at each position around the phosphosite for interaction with barium ions. D. Sequence logo profiles of doubly phosphorylated peptides identified in pH 3.5 Ba^{2+} /acetone fraction of HeLa cell nuclear extract.

Figure 5A – pH 3.5

Figure 5B – pH 4.6

Figure 5C – pH 8.0

Figure 5.

Sequence logos of phosphopeptides from whole cell lysates of 5minute *iso*-stimulated HEK cells overexpressing β 1AR: pH 3.5 (A), pH 4.6 (B) and pH 8.0 (C). Sequence logos were built for amino acids lining of Ba^{2+} ions binding pockets (–5 to +5 AAs).

Figure 6. Log2 ratio of protein expression levels (2 min stimulation / 30 sec stimulation) and their corresponding phosphorylation levels.

Figure 7.

Network centered on MAPK1. Canonical pathways are highlighted. Legend: arrows – proteins that act upon another protein, lines – protein-protein interactions, numbers below shapes – protein expression levels, shapes – inverted triangle/kinase, diamond/enzyme, double circle/group or complex, circle/other. Filled shapes – dataset file genes, Non-filled shapes – Knowledge base genes. Red color represents up-regulation and green color down-regulation. While we kept the original IPA color codes, our majority of the proteins showed unchanged expression levels.

Table 1

Identified phosphopeptides from 1 mg of HeLa cell nuclear extract. Three different conditions of sampling with Ba²⁺ ions (moles) are presented for constant amount of proteins. Results are organized per pH fraction.

[Ba ²⁺] (μmol)	<i>pH 3.5</i>	<i>pH 4.6</i>	<i>pH 8.0</i>	Total [*]
7.5	1644	801	153	2723
6	872	821	177	1868
1	1217	416	295	2215

^{*} It includes MS3 identified phosphopeptides

Table 2

Identified phosphopeptides from nuclear extract of HeLa cells. Three different equivalent ratios total protein amount/ moles of Ba²⁺ ions were investigated.

MudPIT runs	3	3	3
Total protein (μg)	250	500	1000
Moles Ba ²⁺ per pH fraction	1.5	3	6
Identified phosphopeptides*	1037	1582	1868

* It includes MS3 identified phosphopeptides

Table 3

Spectral copies per identified phosphopeptide from HeLa cell nuclear extract. Results are organized per pH fraction and fractionation conditions (total protein amount, moles of Ba²⁺ ions).

[Ba ²⁺] (μmol)	<i>pH 3.5</i>	<i>pH 4.6</i>	<i>pH 8.0</i>
7.5	4.5	1.8	3.8
6	2.9	2.5	2.5
1	3.4	3.4	3.1

Table 4

Phosphopeptides identified from HeLa cell nuclear extract: 1 mg total protein and 7.5 μ mol per pH fraction. Results are organized per pH fraction.

Fraction	pH 3.5	pH 4.6	pH 8.0
<i>Total number of phosphopeptides MS2 (localized sites-phosphopeptides)</i>	1643 (1146)	796 (583)	153 (112)
<i>Phosphorylated sites per peptide: 1</i>	905	751	135
<i>Phosphorylated sites per peptide: 2</i>	505 (317) *	45	14
<i>Phosphorylated sites per peptide: 3</i>	233	0	4
<i>Total number of phosphopeptides MS3</i>	622	517	101
<i>Total phosphopeptides (MS2+ MS3)</i>	1816	1018	225

* localized phosphorylated site

Table 5

Analysis of HEK cells lysate with Ba²⁺/pH/acetone fractionation.

pH	MS2	Enrichment	Copies
3.5	801	15	7.6
4.6	375	3.6	4.8
8.0	80	0.2	1.5

Table 6

Protein expression and phosphorylation levels.

<u>Phosphoproteins with up-regulated phosphorylation</u>				
ID/Protein name/(Entrez Gene)	Phosphopeptide	Ratio	Protein Levels*	Networks
IPI00016513.3 Ras-related protein Rab-10 (RAB10)	K.TPVKEPNSENVDIS*SGGGVT*GWK.S	5.26	1.9	4
IPI00470498.1 Isoform 3 of Plasminogen activator inhibitor 1	K.S*EEAHAEDSVMDHHFR.K	2.94	0.76	4, 7
RNA-binding protein (SERBP1)				
IPI00221141.1 Mitogen-activated protein kinase 14 isoform1 (MAPK14)	R.HTDDEMTGY*VATR.W	2.7	1.27	3
IPI00376295.1 Mitogen-activated protein kinase 1 (MAPK1)	R.VADPDHDHTGFLTEY*VATR.W	16.7	0.85	3
<u>Phosphoproteins with down-regulated phosphorylation</u>				
IPI00019502.1 Myosin heavy chain, nonmuscle type A (MYH9)	R.KGAGDGS*DEEVDGK.A	0.24	0.92	2
IPI00470779.2 Alpha-taxilin (TXLNA)	K.SS*PGQPEAGPEGAQERPSQAAPAVEAEGPGSSQAPR.K	0.24	0.63	6
IPI00418458.1 Pinin, desmosome associated protein (PNN)	K.EIAIVHS*DAEKEQEEEEQKQEMEVK.M	0.29	0.54	1, 4
IPI00374770.1 Microtubule-associated protein 1B isoform 2 (MAP1B)	K.VQSLEGEKLS*PK.S	0.31	0.53	7
IPI00021831.1 cAMP-dependent protein kinase type I-alpha regulatory subunit (PRKAR1A)	R.TDSREDEIS*PPPPNPVVK.G	0.33	0.58	3
IPI00456832.2 Hypothetical protein LOC84726 (KIAA0515) GI:149192855	K.LKFS*DDEEEEEEVVK.D	0.33	0.55	
IPI00328293.2 Serine/arginine repetitive matrix 1 (SRRM1)	K.KET*ES*EAEDNLDDLEK.H	0.34	0.48	1