Published in final edited form as:

J Proteome Res. 2008 October ; 7(10): 4215–4224. doi:10.1021/pr7007785.

Quantitative Phosphoproteome Analysis of Lysophosphatidic Acid Induced Chemotaxis applying Dual-step ^{18}O Labeling Coupled with Immobilized Metal-ion Affinity Chromatography

Shi-Jian Ding^{1,#}, Yingchun Wang^{2,#}, Jon M. Jacobs¹, Wei-Jun Qian¹, Feng Yang¹, Aleksey V. Tolmachev¹, Xiuxia Du¹, Wei Wang², Ronald J. Moore¹, Matthew E. Monroe¹, Samuel O. Purvine¹, Katrina Waters¹, Tyler H. Heibeck¹, Joshua N. Adkins¹, David G. Camp II¹, Richard L. Klemke², and Richard D. Smith^{1,*}

¹Biological Sciences Division, Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA.

²Department of Pathology and Moores Cancer Center, University of California, San Diego, La Jolla, CA 92093.

Summary

Reversible protein phosphorylation is a central cellular regulatory mechanism in modulating protein activity and propagating signals within cellular pathways and networks. Development of more effective methods for the simultaneous identification of phosphorylation sites and quantification of temporal changes in protein phosphorylation could provide important insights into molecular signaling mechanisms in various cellular processes. Here we present an integrated quantitative phosphoproteomics approach and its application for comparative analysis of Cos-7 cells in response to lysophosphatidic acid (LPA) gradient stimulation. The approach combines trypsin-catalyzed $^{16}\text{O}/^{18}\text{O}$ labeling plus $^{16}\text{O}/^{18}\text{O}$ -methanol esterification labeling for quantitation, a macro-Immobilized Metal-ion Affinity Chromatography trap for phosphopeptide enrichment, and LC-MS/MS analysis. LC separation and MS/MS are followed by neutral loss-dependent MS/MS/MS for phosphopeptide identification using a linear ion trap (LTQ)-FT mass spectrometer. A variety of phosphorylated proteins were identified and quantified including receptors, kinases, proteins associated with small GTPases, and cytoskeleton proteins. A number of hypothetical proteins were also identified as differentially expressed followed by LPA stimulation, and we have shown evidence of pseudopodia subcellular localization of one of these candidate proteins. These results demonstrate the efficiency of this quantitative phosphoproteomics approach and its application for rapid discovery of phosphorylation events associated with LPA gradient sensing and cell chemotaxis.

Keywords

Protein Phosphorylation; Liquid Chromatography; Mass Spectrometry; ^{18}O Labeling; Accurate Mass; Quantitation; Phosphoproteomics

*To whom correspondence should be addressed at: Dr. Richard D. Smith, Biological Sciences Division, Pacific Northwest National Laboratory, P. O. Box 999, MSIN: K8-98, Richland, WA 99352. Phone: 509-376-0723; Fax: +509-376-0323; Email: rds@pnl.gov.

[#]These authors contributed equally to the work.

Introduction

Protein post-translational modifications (PTMs) generate tremendous diversity, complexity and heterogeneity of gene products and their characterization remains one of the major challenges in proteomics¹. Among more than 300 different types of PTMs, protein phosphorylation plays an extensive and important role in eukaryotic signal transduction, gene regulation, and metabolic control in cells^{1,2}. Conversely, abnormal phosphorylation is a cause or consequence of multiple diseases². Accordingly, protein kinases and their substrates are becoming the most important group of drug targets and could be invaluable cancer markers in molecular classification of cancers and predicting clinical outcomes and responses to therapy^{3,4}. Therefore, the development of sensitive and comprehensive analytical methods for quantitative phosphoproteomics analysis is an area of high interest in proteomics research.

Due to the relatively low abundance of phosphorylation within the global proteome, successful strategies have employed specific enrichment approaches for those proteins/peptides containing active phosphorylation sites. Several approaches have been explored to date for their selective isolation; the most notable of these include either affinity- or chemical derivatization-based methodologies. Among these approaches, anti-phosphotyrosine antibodies have been used successfully for the isolation of phosphoproteins⁵ and more recently, phosphopeptides⁶. Chemical tagging of pSer/pThr residues via beta-elimination has been extensively reported⁷⁻⁹. However, chemical modifications of phosphate groups often are accompanied by significant sample losses and side reactions, and to date only limited applications for the analysis of phosphopeptides within complex mixtures have been reported.

Immobilized metal-ion affinity chromatography (IMAC), strong cation exchange chromatography (SCX) and titanium dioxide chromatography have recently been demonstrated as effective methods for enrichment of phosphopeptides from peptide mixtures¹⁰⁻¹². After enrichment the phosphate remains attached to the peptide backbone, but the metastability of the phosphate group presents a significant challenge for peptide sequencing and phosphorylation site identification by tandem mass spectrometry. This is due to the exposure of this labile modification, especially for pSer/pThr-containing peptides, to collision-induced dissociation (CID) which often results in extensive neutral loss of phosphoric acid, and generates spectra containing few sequence-informative ions¹¹. An alternative approach to overcome this problem is to introduce additional stages of tandem MS¹⁰. In ion traps the abundant neutral loss product ions can be further isolated, fragmented, and analyzed, by a process termed MS/MS/MS (MS³).

Elucidation of signaling networks also requires quantification of the dynamic changes of protein phosphorylation. Several recent studies have reported quantitative analysis of protein phosphorylation by coupling stable isotope labeling with phosphopeptide enrichment and MS analysis¹³⁻¹⁶. Stable isotope labeling by amino acids in cell culture has enabled quantification of pheromone-induced changes in a protein's general state of phosphorylation¹⁵; however, this method involves *in vivo* cell culture labeling which is not applicable to tissue and biological fluids. Moran and co-workers¹⁶ have applied methyl esterification with normal or deuterated methanol to determine changes in the epidermal growth factor-induced phosphorylation in human tumors in the presence or absence of a chemical inhibitor. A limitation of deuterium-labeling is that the light and heavy isotopic labeled peptide pairs do not co-elute during LC separation, resulting in inaccuracies when attempting to quantitate the results¹⁷. Other approaches include N-terminal labeling using commercially available iTRAQ reagents coupled with immunoprecipitation for quantitative analysis of tyrosine phosphorylation, but its application currently is expensive for large sample sets and is limited to specific MS instrumentation¹⁴.

In this report, we present an alternative method for quantitative analysis of protein phosphorylation. This work extends the previously reported ^{18}O labeling method by coupling an ^{18}O -methanol esterification reaction with trypsin-catalyzed ^{18}O labeling followed by IMAC enrichment for quantitative analysis of phosphopeptides¹⁸. As a demonstration, we applied this dual-step ^{18}O labeling method to the study of phosphorylation events involved in directed cell migration (chemotaxis) induced by an LPA gradient. Chemotaxis is a critical process in many diverse physiological and pathological settings, such as immune response, embryogenesis, angiogenesis, and tumor metastasis¹⁹. Chemoattractants, such as LPA, bind to a variety of cell-surface receptors as ligands and thereby activate multiple signal transduction pathways, including those initiated by the small GTPases and Erk/MAPK signaling pathway, leading to the rearrangement of the actin and other cytoskeletal components which drives cell locomotion by the extension of pseudopods¹⁹. In this study the regulation by site-specific phosphorylation of a wide range of molecules involved in the LPA-induced signaling pathways was revealed including proteins associated with small GTPases, such as Rho-GTPase activating protein 6, Rho guanine nucleotide exchange factor 11, and Rho/Rac guanine nucleotide exchange factor 2; the upstream and downstream players of the Erk/MAPK signaling pathway such as dual-specificity mitogen-activated protein kinase kinase 2 and dual-specificity tyrosine-phosphorylation regulated kinase 1. Many cytoskeleton proteins and previously uncharacterized hypothetical proteins were also identified. These results provide a novel insight of the phosphorylation-based LPA signaling mechanism as well as new cellular targets regulated by LPA in chemotaxis. Accurate quantitation is facilitated by this approach since $^{16}\text{O}/^{18}\text{O}$ labeled peptides largely co-elute during chromatographic separations and the approach is applicable to any digestible protein-based sample (tissue, biofluid, cell culture).

EXPERIMENTAL PROCEDURES

Cell culture and sample preparation

Cos-7 cells (ATCC, Rockville, MD) were cultured in high glucose DMEM (GIBCO-BRL, Life Technologies, Carlsbad, CA) supplemented with 10% fetal calf serum (Hyclone, Logan, UT). The low passage cells were grown to 80% confluence and harvested by treating with 0.25% trypsin. Each 1.5×10^6 cells were re-plated on fibronectin-coated, 3 μm -pore polycarbonate membranes in Costar Transwell insert. LPA was added to the lower chamber for 10 min to create a diffusion gradient. The inserts were washed in PBS. Cells were harvested into the lysis buffer (8 M urea, 40 mM Tris, pH 8.4, protease inhibitor cocktail, 2 mM Vanadate and 50 nM calyculin) and centrifuged at $45,000 \times g$ for 1 h at 4 °C to pellet cell debris. Multiple patches of samples were pooled together and the total protein samples were reduced with 10 mM DTT for 30 min at 37 °C. Protein cysteinyl residues were alkylated with 40 mM iodoacetamide for 90 min at room temperature. The protein concentrations were measured using a Coomassie protein assay (Pierce, Rockford, IL). The samples were then diluted 10 times by using 100 mM NH_4HCO_3 (pH 8.3) and digested into peptides using sequencing grade trypsin (Promega, Madison, WI) overnight at 37 °C with a 1:50 (w/w) trypsin-to-protein ratio. The peptides were desalted by using a SPE C18 column (Supelco, Bellefonte, PA) and lyophilized. Peptide concentration was determined using BCA assay (Pierce). Residual trypsin activity was quenched by boiling the samples for 10 min and immediately placing the samples on ice.

Trypsin-catalyzed $^{16}\text{O}/^{18}\text{O}$ labeling

Trypsin-catalyzed $^{16}\text{O}/^{18}\text{O}$ labeling was performed as previously described¹⁸. Briefly, 500 μg dried peptides from control and LPA stimulated cells were dissolved with 50 mM NH_4HCO_3 , 10 mM CaCl_2 in 250 μL normal ^{16}O water and 95% ^{18}O water (Isotec, Miamisburg, OH), respectively, and were incubated with 12.5 μL immobilized trypsin (Applied Biosystems, Foster City, CA) for 24 h at 30 °C. Following labeling, supernatant from $^{16}\text{O}/^{18}\text{O}$ -labeled

sample was collected after centrifuging the samples at $15,000 \times g$ for 5 min and then lyophilized.

Esterification of peptides using $^{16}\text{O}/^{18}\text{O}$ -methanol

Peptides were converted to corresponding methyl esters according to the previously reported procedure¹¹. Briefly, dried peptides from the corresponding ^{16}O - and ^{18}O -labeled samples were reconstituted in 2 M ^{16}O -methanolic and ^{18}O -methanolic HCl, respectively, and allowed to stand at room temperature for 2 h under nitrogen. ^{16}O -Methanolic or ^{18}O -methanolic HCl was removed in a SpeedVac centrifuge (Savant, Hicksville, NY), and the sample was stored at -80°C until phosphopeptides were enriched by IMAC.

IMAC enrichment

Phosphopeptides were enriched by IMAC as previously described with some modification¹¹. A 3 mm $\times$ 8 mm length Macro trap (Microchrom, Auburn, CA) was custom packed with POROS 20 MC resin (Applied Biosystems, Foster city, CA). Columns were activated with 300 μL 0.1 M FeCl_3 (Aldrich, Milwaukee, WI). Excess metal ions were removed by 300 μL of 0.1% acetic acid. Dried dual-step $^{16}\text{O}/^{18}\text{O}$ labeled peptides from each sample were dissolved in equal parts of methanol, acetonitrile, and 0.1% acetic acid and combined for a total of 250 μg of peptide loaded onto the IMAC column. Non-specific binding peptides were removed by washing the column with 300 μL 100 mM NaCl, 1% acetic acid, 25% methanol. Phosphopeptides were eluted by 150 μL 50 mM Na_2HPO_4 (pH 9.0) and were acidified with acetic acid immediately after elution.

LC-MS/MS and MS³ Analysis

One-fifth (30 μL) of the IMAC eluate was loaded on a fused-silica pre-column (4 cm $\times$ 100 μm i.d.) packed with 5- μm C18 particles (YMC, Wilmington, NC). The pre-column was then rinsed with 0.1% acetic acid to remove Na_2HPO_4 . The pre-column was connected to the in-house analytical monolithic capillary LC column (40 cm $\times$ 360 μm o.d. $\times$ 50 μm i.d. fused silica) prepared as previously described²⁰ with 2 cm of Teflon tubing. Samples were analyzed by nanoflow HPLC-micro-electrospray ionization using a linear ion trap-FT (LTQ-FT) mass spectrometer (ThermoFinnigan, San Jose, CA). The HPLC gradient (A = 100 mM acetic acid in water, B = 70% acetonitrile/100 mM acetic acid in water), 0-5% B for the first 5 min, 5-60% B for 215 min, 60-100% B for 20 min, 100-0% B for 20 min, and 0% B for the last 60 min. Briefly, the mass spectrometer was operated in the data-dependent mode to automatically switch between MS, MS/MS (MS^2), and neutral loss-dependent MS^3 acquisition, as described previously, with some modification¹³. MS spectra from m/z 300-1575 were acquired by FTICR with resolution $r = 25,000$ at m/z 400 (after accumulation to a target value of $5\text{E}6$ in the linear ion trap). The top 3 most intense ions were sequentially isolated for a selected ion monitor (SIM) scan that consisted of a 20- m/z mass range and then simultaneously fragmented in the linear ion trap using collision-induced dissociation at a target value of 10,000. The isolation window for MS/MS is ± 1.5 Da. An MS^3 scan was triggered when a neutral loss of 98, 49, or 32.7 m/z change was detected among the 10 most intense fragment ions. The LC-MS analysis for each sample was repeated 4 times. Initial LC-MS analyses of dual-step ^{18}O -labeling efficiency were performed on a fully automated custom built capillary LC system coupled to an Apex III 9.4-tesla FTICR mass spectrometer (Bruker Daltonics, Billerica, MA) using an in-house manufactured ESI interface¹⁸.

Data Analysis

All MS^2 and MS^3 spectra were searched against the human International Protein Index data base (consisting of 49,161 protein entries, Version 3.05, April, 2005; available online at www.ebi.ac.uk/IPI) using SEQUEST (Version v.27, Rev 12, ThermoFisher Scientific,

Waltham MA) and tryptic rules. Methyl esterification of aspartic acid, glutamic acid, and peptide C-terminus, in addition to carboxamidomethylation of cysteine were selected as static modifications and phosphorylation of serine, threonine, and tyrosine as dynamic modifications. MS³ spectra were extracted and searched with an additional dynamic modification of loss of water on serine and threonine. All the data were searched twice using the static modification of either ¹⁶O- or ¹⁸O-methyl esterification. Parent mass tolerance of ± 2 Da was used by SEQUEST search considering peptides 1 Da and 2 Da away from the parent ion. The measured monoisotopic m/z value, charge state, as well as intensity information of peptide ions from SIM scan were extracted using in-house software²¹. Accurate masses calculated from SIM scan were used to match the theoretical masses of both light and heavy species of peptide sequences given by SEQUEST from the next MS² scan. >96% of all identified peptides achieved a mass accuracy of less than 2 ppm after mass correction using the described algorithm²¹ with a minimal Xcorr cutoff of 1.5 for SEQUEST results utilized for further processing. A probability score was automatically calculated for every identified phosphorylation site using an in-house program and described algorithm²².

The isotopic ratios of light and heavy isotope labeled pairs from multiple observations of the same peptide over consecutive MS cycles for the duration of their respective LC-MS peaks were averaged to give one ratio per observed peptide pair and normalized according to mean ratio. To address the issue of incomplete incorporation of ¹⁸O during the dual-step labeling, mainly due to the ~95% pure ¹⁸O water and ¹⁸O methanol, if a peptide contains no aspartic acid (D) or glutamic acid (E), the relative abundance ratios (¹⁸O/¹⁶O) of the identified phosphopeptide pairs were accurately computed using the equation as previously reported¹⁸. If the peptide contains one or more residues of D and/or E, the following equation was used to calculate the ratio.

$$H_0'' = H_0 - L_0 \frac{M_{2n}}{M_0}, \quad H_2'' = H_2 - H_0'' \frac{M_2}{M_0} - L_0 \frac{M_{4n}}{M_2}$$

$$R\left(\frac{^{18}\text{O}}{^{16}\text{O}}\right) = \frac{H_0'' + H_2'' + H_4 - H_2'' \frac{M_2}{M_0}}{L_0}$$

Where n equals the number of glutamic acid and/or aspartic acid residues present in the peptide sequence; L_0 is the measured intensity for the monoisotopic peak for a peptide without ¹⁸O labeling, H_4 , H_2 , and H_0 are the measured intensities for the monoisotopic peak for the labeled peptide with complete ¹⁸O incorporation, the peaks with a mass 2 Da and 4 Da lower than the completely labeled monoisotopic peak, respectively. M_0 , M_2 , and M_4 are the predicted relative abundances for the monoisotopic peak for a peptide, the peaks with a mass 2 Da and 4 Da higher than the monoisotopic peak, respectively. The M_2/M_0 and M_4/M_0 ratios are estimated using equations as previously reported¹⁸.

Immunofluorescence Analysis

The full length cDNA sequence of hypothetical protein FLJ38771 was retrieved from NCBI database (<http://www.ncbi.nlm.nih.gov>). The cDNAs were amplified by RT-PCR using total RNA from human MDA-MB-231 cell line as the template and sequenced by the DNA facility of the Scripps Research Institute. The amplified cDNAs were cloned into the modified pTriEx vector (EMD Biosciences, San Diego, CA) and fused with GFP at the N-terminal. Cos-7 cells were transfected with the plasmids expressing GFP fused hypothetical protein FLJ38771 using the Effectene transfection reagent (Qiagen), and allowed to express the fusion protein for 36 h post-transfection. The cells were then suspended with trypsin and re-plated on fibronectin (1 $\mu\text{g/mL}$ in PBS) coated glass coverslips to attach and spread for 1.5 h. The attached cells were then washed with PBS for two times and fixed with 4% paraformaldehyde in PBS. The fixed cells were permeabilized with 0.1% Triton-X100 and stained with Alex Fluor 546 phalloidin (Molecular Probes) overnight. The coverslips were mounted with ProLong Gold antifade

reagent (Invitrogen), and the images of the cells were scanned alternatively in red (560 nm) and green channels (488 nm) by Nikon CSi confocal microscope using 60x objective and merged by the software EZ-C1 (Nikon).

Results

Strategy overview

The objective of this study was to develop methods for quantitative measurements of differential protein phosphorylation occurring between different cell states, treatments, or time course studies. To accomplish this goal, an integrated strategy that combines dual-step ^{18}O labeling, IMAC, and LC-MS analysis has been designed (Figure 1). Protein samples extracted from cells stimulated with an optimal LPA gradient were digested with trypsin and then the C-termini of the peptides were labeled with $^{16}\text{O}/^{18}\text{O}$ by incubation in either ^{16}O or ^{18}O water with immobilized trypsin, as previously described¹⁸. To enable simultaneous esterification and quantitative labeling, $^{16}\text{O}/^{18}\text{O}$ -methanol was used for the esterification reaction to give a $4 + (2 \times n)$ mass difference for the ^{18}O -labeled peptide pairs (n equals to the number of aspartic acid and glutamic acid residues that are present in the peptide sequence). To improve the robustness and capacity of phosphopeptide enrichment, an IMAC macro trap with a 3 mm inner diameter and 8 mm length was performed off-line¹¹. For the LC-MS analysis of phosphopeptides enriched from the IMAC column, a $50\ \mu\text{m i.d.} \times 40\ \text{cm}$ long monolithic column with an integrated electrospray emitter was used, which provided an increase in robustness, and high sensitivity for comparative quantitative analysis²⁰. A hybrid instrument LTQ-FT mass spectrometer was used to acquire the data and provided accurate mass measurement for the quantification and identification including MS^3 for improved fragmentation of phosphopeptides.

Quantitation approach using dual-step ^{18}O -labeling

With the dual-step ^{18}O -labeling strategy, we anticipated that the light and heavy labeled isotopic pairs would have a $4 + (2 \times n)$ Da mass shift (n equals to the number of D and/or E residues existing in the peptide sequence). Ideally the incorporation of ^{18}O atoms would proceed to 100%, but in practice labeling is limited by the ~95% pure ^{18}O water and ^{18}O methanol. In addition, the efficiency of the esterification step monitored by a parallel experiment using standard proteins was found to be ~95%, in agreement with a previous report¹¹. Despite the issue of incomplete incorporation, the relative abundance ratios could be determined using the described equation (See Experimental Procedures Section) to correct for the effect of incomplete ^{18}O incorporation.

To assess the quantitative accuracy using the dual-step ^{18}O labeling, a control experiment starting with two identical standard peptide samples was performed. Figure 2A shows a histogram of the relative abundance ratios observed from this experiment. The abundance ratios for the ^{18}O - versus ^{16}O -labeled versions of each peptide were calculated using the described equations. As expected, the $\log_2$ abundance ratios for the detected peptide pairs displayed a Gaussian-like distribution centered at 1.0. The average ratio for all peptide pairs is 1.03 ± 0.27 , which is comparable to previous results (1.02 ± 0.23 .) reported for single-step ^{18}O -labeling¹⁸, demonstrating the accuracy for relative abundance measurements in a complex, proteome-wide sample. The distribution of $\log_2$ ratios for control vs LPA experiments is shown in Figure 2B, which displays a broader range of relative abundance ratios with a subset of phosphorylation sites that show significant changes in phosphorylation levels in response to LPA gradient stimulation. The average ratio for the entire set is 1.61 ± 3.89 . To examine the reproducibility for relative abundance ratio measurements, we performed four replicate analyses of the same sample, the average CV for all quantified peptide across the technical

replicates is ~19%. The log2 distribution of ratios between the technical replicate experiments is shown as Supplemental Figure 1.

The identification of phosphopeptides using isotopic pairing and accurate mass

Phosphopeptides are notorious for their uninformative fragmentation modes using collisional dissociation. In order to obtain more sequence informative ions for phosphopeptides, we performed neutral loss-driven MS³ analysis from the most intense ions using the hybrid instrument LTQ-FT MS. Figure 3 depicts the experimental process. Each phosphopeptide pair was subjected to a narrow m/z bandwidth analysis (so called “zoom” scan) using the built-in single-ion-monitoring (SIM) capability in the FTICR stage. Upon successful observation with the zoom scan approach, paired signals were subjected to MS² and MS³ in the linear ion trap stage. In total, we recorded 21,552 MS² fragmentation events with neutral loss-triggered acquisition of 4173 MS³ spectra. From the 3566 phosphopeptides identified, many peptides were sequenced multiple times due to ¹⁸O/¹⁶O pairs, MS2/MS3 redundancy, and different precursor charge states, which collapsed into a total of 497 unique phosphopeptides (Supplemental Table 1).

We found that the mass difference between the labeled peptide pairs using this dual-step ¹⁸O labeling method can be utilized to help distinguish between true positive and false positive peptide identifications. For example, as shown in Figure 3, the peptide sequence ASAVSELS*PR contains an aspartic acid, so the expected mass difference between the light and heavy forms of the peptide pairs should be $4 + (2 \times 1) = 6$ Da, which is in agreement with the observed mass from the SIM scan. To improve the confidence and throughput of phosphopeptide identifications, parameters utilizing accurate mass and reverse database approaches were applied^{13, 23}. Figure 4 shows the effect of mass accuracy for phosphopeptide identification. Typically less than 2 ppm mass accuracy was achieved after mass correction using the described algorithm²¹.

After database searching, accurate mass information in combination with isotopic pairing helped confirm correct phosphopeptide identifications, which are more challenging due to the reduced MS/MS spectral quality caused by the significant neutral loss of phosphoric acid and water typically from pSer/pThr-containing peptides¹¹. A total of 497 unique phosphopeptides were identified, and 384 out of 497 provided quantitative data (Supplemental Table 2). 113 out of 497 peptides did not provide quantitative data mostly due to low signal-to-noise ratios of the MS level precursor data. In some cases, the S/N ratio of the identified peak is good, but no pairing peaks could be found. This could be due to the on- and off-regulation of the phosphorylation of this peptide in the two experimental conditions. Of 497 unique phosphopeptides, 312 had a confident site location (P, 0.05; Ascore > 19), 425 (71.19%) contained pSer, 134 (22.45%) contained pThr, and 38 (6.37%) contained pTyr (Supplemental Figure 2). 411 were monophosphorylated (82.7%), 75 were diphosphorylated (15.1%), 3 were triphosphorylated (0.6%), and 2 were tetraphosphorylated (0.04%). Comparing our total identified phosphorylation sites with the phosphor ELM database and two other large scale phosphoproteomics studies^{10,15} using PhosphoBlast²⁴ resulted in a 9%, 31%, and 12% overlap respectively (Supplemental Table S3).

Phosphorylation events induced by LPA gradient stimulation

LPA is a lipid mediator with a wide variety of biological actions, particularly as an inducer of cell migration, proliferation and survival²⁵. LPA binds to specific G-protein-coupled receptors and thereby activates multiple signal transduction pathways, including those initiated by the small GTPases Ras, Rho, and Rac, which result in stress fiber formation, assembly of focal adhesion, neurite retraction, and modulation of gene transcription²⁵. In this study, from the total of 384 quantified phosphopeptides, 55 were observed with more than 2-fold enrichment

in LPA stimulated cells (Table 1). A number of proteins associated with small GTPases were identified including Rho-GTPase-Activating Protein 6 (ARHGAP6, 5.54), Rho Guanine Nucleotide Exchange Factor 11 (ARHGEF11, 3.10), Molecule Interacting With Rab13 (MICAL-L1, -0.64), TBC1 Domain Family Member 1 (TBC1D1, 0.78), Rho/Rac Guanine Nucleotide Exchange Factor 2 (ARHGEF2, -1.15), and GTPase Activating Rap/Ran-Gap Domain-Like 1 (-1.34). ARHGAP6 has two independent functions: one as a GAP with specificity for RhoA and the other as a cytoskeletal protein that promotes actin remodeling²⁵. ARHGEF10 is a member of the family of Rho guanine nucleotide exchange factors (GEFs), which are implicated in neural morphogenesis and connectivity and regulate the activity of Rho GTPases by catalyzing the exchange of bound GDP by GTP²⁶. Rab family is the largest branch of the Ras superfamily of GTPases and plays a key role in the regulation of membrane trafficking and receptor localization in eukaryotic cells²⁷. In this study, two proteins related with Rab were identified with increased expression in pseudopodium: TBC1 domain family member 4 (Rab GAP4, Phosphorylated by AKT1; insulin-induced) and the protein interacting with Rab13. Kamei et al. proposed that some Rab family members may play a cooperative role with Rho in cell adhesion and migration²⁸. In addition, Rab13 may be involved in the regulation of PKA signaling during tight junction assembly²⁹.

ERK plays an important role in regulating cell motility^{30,31,32}. Motif analysis revealed that 32 phosphopeptides in our dataset (Table 2) contain Erk kinase substrate motif suggesting that Erk may be implicated in LPA-induced phosphorylation events. Among these motifs, some are known Erk substrates including the known site pTyr271 of dual-specificity tyrosine-phosphorylation regulated kinase 1b³³, Cortactin, and α -Parvin^{34,35,36}. Moreover, Erk itself is activated through phosphorylation by its upstream kinase, the dual specificity mitogen-activated protein kinase kinase 2 (MEK2), which is also identified to be phosphorylated at the site Thr394. Cytoskeletal proteins are the final downstream effectors of various signaling events that control cell morphology and motility. Phosphorylation sites from a variety of cytoskeleton proteins were revealed in this study, including plectin, vimentin, synaptopodin, catenin (cadherin-associated protein), delta 1, microtubule-associated protein 4, microtubule-associated protein tau, kinesin light chain 1, filamin c, alpha-parvin, dynamin 2: microtubule-associated protein, beta-catenin, kinesin-like protein kif1c, dynein light chain-a, drebrin, cortactin, and AHNAK. Actin filaments form a branching “dendritic” network and numerous other proteins play supporting roles in the “dendritic” network. Plectin has been shown to function as a general cross-linker of the cytoskeleton, interacting with all three major groups of cytoskeletal proteins³⁷. The interactions between plectin and other cytoskeletal proteins, including vimentin and lamin B, are regulated by protein phosphorylation catalyzed by CDK1, PKC or PKA, CaMK-II³⁸. Alpha-Parvin is a 42 kDa focal adhesion protein, related to the α -actinin superfamily which crosslinks actin filaments into tight bundles or meshworks of looser arrangement, and also connects them to the plasma membrane³⁹. Alpha-Parvin could be phosphorylated by integrin-like kinase and regulates cell adhesion and spreading⁴⁰. AHNAK is a giant phosphoprotein whose carboxyl-terminal domain was found to induce actin bundling. Cortactin, whose known phosphorylation site S418 was identified in this study, is an SH3 domain-containing protein that contributes to the formation of dynamic cortical actin-associated structures, such as lamellipodia and membrane ruffles⁴¹. A recent study showed that phosphorylation of Cortactin by Fer tyrosine kinase regulates N-Cadherin mobility and intercellular adhesion strength⁴². Vimentin is shown to play an important role in vital mechanical and biological functions such as cell contractility and migration⁴³.

Besides the proteins with known relation to cell migration, we also detected regulated phosphoproteins falling into different functional categories. Protein network analysis using MetaCore™ revealed 14 identified phosphoproteins interacting with cell division cycle 2 protein isoform 2 (CDK1). Supplemental Table 4 shows phosphorylation sites of CDK1 and its substrates/interacting proteins characterized in our data set. Of the 24 identified

phosphorylation sites, 16 were previously identified in the Swiss-Prot database. The majority of these sites were unambiguously assigned in this study except the peptide from lamin A/C, which may be phosphorylated either on S407 or S411. The inability to localize the exact site of phosphorylation is due to the lack of definite evidence from the mass spectrum. These results, including activation of the Erk/MAPK signaling pathway and a relative large number of proteins interacting with CDK1, are in agreement with the mitogenic and chemoattractant activity of LPA²⁵.

In addition to the proteins discussed above, a number of hypothetical proteins were also identified as differentially expressed in control vs LPA gradient stimulated cells. These proteins are potentially important kinase substrates and signaling proteins involved in cell polarization and migration. For an initial investigation, one of the identified hypothetical phosphoprotein genes (FLJ38771, IPI00298373) that display differential phosphorylation in the LPA gradient stimulated cells (Figure 5A) was cloned and then fused with GFP to visualize cellular localization. Results showed that FLJ38771 strongly localizes to the leading front of polarized cells (pseudopodial projections) and actin-rich membrane ruffles at the cell's edge (Figure 5B). This protein may bind actin and regulate the cytoskeleton of migratory cells. Functional predictions based on domain and motif structures revealed ERK docking sites in FLJ38771, suggesting that it has potential roles in ERK-mediated pseudopodium formation. Recent evidence suggests that FLJ38771 is an amyotrophic lateral sclerosis 2 disease related protein, although its precise role is not yet known (Figure 5A)⁴⁴. This and other to be characterized hypothetical proteins are strong candidates for key players in cell polarity and will be examined in future studies.

Discussion

Significant progress has recently been made in phosphoproteomics in terms of enrichment of phosphopeptides from complex mixtures and identification of phosphopeptides using mass spectrometry^{4,7,38}. However, there are few literature reports describing global comparative quantitative phosphoproteomics studies. Here we describe an integrated quantitative phosphoproteomics approach that has been applied for comparative analysis of Cos-7 cells in response to LPA gradient stimulation using IMAC coupled with trypsin-catalyzed ¹⁶O/¹⁸O labeling plus ¹⁶O/¹⁸O-methanol esterification. A key advantage of using dual-step ¹⁸O labeling is a minimal chromatographic shift thus facilitating accurate quantitation compared to ⁴D-methanol labeling. Also, comparing to SILAC or any metabolic labeling method, ¹⁸O labeling is broadly applicable. Additionally, the added mass difference can assist in accurately quantifying large peptide pair masses containing multiple aspartic acids and/or glutamic acids which normally would be difficult to distinguish using trypsin-catalyzed ¹⁸O labeling alone. Moreover, the isotopic pairing information can be applied to orthogonal filtering criteria where the number of false positive identifications can be lowered substantially.

Despite the advantages of our current approach, there are some remaining issues. The number of phosphorylation sites identified in this study is not as extensive as those recently reported using additional peptide fractionation preparation steps^{15,22}. Coupling protein or peptide level fractionation with the current strategy should similarly increase the phosphopeptide coverage significantly. Methyl esterification could complicate the annotation of spectrum resulting in a number of unannotated major peaks in MS/MS spectra due to unexpected side reactions and the additional methyl esterification step leads to the loss of peptides presumably due to the additional drying of peptides in the organic buffer. It has been reported that a lower pH during elution improves peptide losses in the methyl esterification step⁴⁵. However, the best conditions for IMAC enrichment of phosphopeptides still involve methyl esterification of digests⁴⁵.

Notably, only a few multiply phosphorylated peptides were identified reflecting a bias toward monophosphorylated peptides by LC-ESI-MS/MS compared to MALDI-MS¹². Previously, Ficarro *et al.* showed that 68% of total identified phosphopeptides are doubly phosphorylated peptides¹¹, which may be attributed to the sample overload of capillary IMAC column⁴⁶, or the difficulty in interpreting and assigning multiply phosphorylated peptides. In addition to the high confidence phosphopeptide identifications, we could not conclusively assign some phosphorylation sites and sequences due to insufficient MS/MS data although they showed a neutral loss of phosphate and partial fragmentation information. Electron capture dissociation or electron transfer dissociation MS is likely an effective approach to address this issue since in this method, all ions in the fragments tend to retain their modifications^{47,48}.

In conclusion, we have developed an integrated approach for global quantitative phosphoproteomics studies. A large variety of intracellular signaling molecules have been identified in this study, including small GTPases, MAPK cascades, Ser/Thr and Tyr kinases and scaffold proteins. The regulated phosphopeptides in our dataset reflect LPA-induced signaling events that are critical for the initiation of morphological polarization of migrating cells. Recently, we have reported the global profiling of the signaling polarity in chemotactic cells, and this quantitative approach was key in obtaining chemotactic signaling information to elucidate the spatial relationship of the protein networks³². Augmenting the previous study with this current dataset should greatly facilitate the comprehensive understanding of the signaling events regulating directional cell migration.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank the NIH (grants RR018522 to R.D.S., GM068487 and CA097022 to R.L.K.), Susan G. Komen Foundation (PDF0503999, to Y.W.) and the Laboratory Directed Research Development program at Pacific Northwest National Laboratory (PNNL) for support of this research and the Environmental Molecular Sciences Laboratory (EMSL) for use of the instrumentation applied in this research. EMSL is a U.S. Department of Energy (DOE) national scientific user facility located at PNNL in Richland, Washington. PNNL is a multi-program national laboratory operated by Battelle Memorial Institute for the DOE under Contract DE-AC05-76RL01830.

Abbreviations

LTQ	linear ion trap
SIM	selected ion monitoring
LC-MS	liquid chromatography coupled with mass spectrometry
PTMs	post-translational modifications
pSer	phosphoserine
pThr	phosphothreonine
pTyr	phosphotyrosine
IMAC	immobilized metal-ion affinity chromatography
LPA	lysophosphatidic acid

References

1. Jensen ON. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr Opin Chem Biol* 2004;8:33–41. [PubMed: 15036154]

2. Hunter T. Signaling-2000 and beyond. *Cell* 2000;100:113–127. [PubMed: 10647936]
3. Lim YP. Mining the tumor phosphoproteome for cancer markers. *Clin Cancer Res* 2005;11:3163–3169. [PubMed: 15867208]
4. Cohen P. Protein kinases--the major drug targets of the twenty-first century? *Nat Rev Drug Discov* 2002;1:309–315. [PubMed: 12120282]
5. Salomon AR, Ficarro SB, Brill LM, Brinker A, Phung QT, Ericson C, Sauer K, Brock A, Horn DM, Schultz PG, Peters EC. Profiling of tyrosine phosphorylation pathways in human cells using mass spectrometry. *Proc Natl Acad Sci U S A* 2003;100:443–448. [PubMed: 12522270]
6. 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;23:94–101. [PubMed: 15592455]
7. Goshe MB, Veenstra TD, Panisko EA, Conrads TP, Angell NH, Smith RD. Phosphoprotein isotope-coded affinity tags: Application to the enrichment and identification of low-abundance phosphoproteins. *Analytical Chemistry* 2002;74(3):607–616. [PubMed: 11838682]
8. Oda Y, Nagasu T, Chait B. Enrichment Analysis of Phosphorylated Proteins as a Tool for Probing the Phosphoproteome. *Nature Biotechnology* 2001;19(4):379–382.
9. Zhou H, Watts JD, Aebersold R. A Systematic Approach to the Analysis of Protein Phosphorylation. *Nat. Biotechnol* 2001;19:375–378. [PubMed: 11283598]
10. Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villen J, Jiazu L, Cohn MA, Cantley LC, Gygi SP. Large-scale characterization of HeLa cell nuclear phosphoproteins. *PNAS* 2004;101:12130–12135. [PubMed: 15302935]
11. 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;20:301–305. [PubMed: 11875433]
12. Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, Jorgensen TJ. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics* 2005;4:873–886. [PubMed: 15858219]
13. Gruhler A, Olsen JV, Mohammed S, Mortensen P, Faergeman NJ, Mann M, Jensen ON. Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Molecular & Cellular Proteomics* 2005;4(3):310–327. [PubMed: 15665377]
14. Zhang Y, Wolf-Yadlin A, Ross PL, Pappin DJ, Rush J, Lauffenburger DA, White FM. Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol Cell Proteomics* 2005;4:1240–1250. [PubMed: 15951569]
15. Olsen JV, Blagoev B, Gnadt F, Macek B, Kumar C, Mortensen P, Mann M. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 2006;127:635–648. [PubMed: 17081983]
16. Stover DR, Caldwell J, Marto J, Root R, Mestan J, Stumm M, Ornatsky O, Orsi C, Radosevic N, Liao L, Fabbro D, Moran MF. Differential phosphoproteomes of EGF and EGFR kinase inhibitor-treated human tumor cells and mouse xenografts. *Clin Proteomics* 2004;1:69–80.
17. Zhang R, Sioma CS, Thompson RA, Xiong L, Regnier FE. Controlling deuterium isotope effects in comparative proteomics. *Anal Chem* 2002;74:3662–3669. [PubMed: 12175151]
18. Qian WJ, Monroe ME, Liu T, Jacobs JM, Anderson GA, Shen Y, Moore RJ, Anderson DJ, Zhang R, Calvano SE, Lowry SF, Xiao W, Moldawer LL, Davis RW, Tompkins RG, Camp DG, Smith RD. Quantitative Proteome Analysis of Human Plasma following in Vivo Lipopolysaccharide Administration Using ¹⁶O/¹⁸O Labeling and the Accurate Mass and Time Tag Approach. *Mol. Cell. Proteomics* 2005;4(5):700–709. [PubMed: 15753121]
19. Firtel RA, Chung CY. The molecular genetics of chemotaxis: sensing and responding to chemoattractant gradients. *Bioessays* 2000;22:603–615. [PubMed: 10878573]
20. Luo Q, Shen Y, Hixson KK, Zhao R, Yang F, Moore RJ, Mottaz HM, Smith RD. Preparation of 20-microm-i.d. silica-based monolithic columns and their performance for proteomics analyses. *Anal Chem* 2005;77:5028–5035. [PubMed: 16053318]
21. Tolmachev AV, Monroe ME, Jaitly N, Petyuk VA, Adkins JN, Smith RD. Mass measurement accuracy in analyses of highly complex mixtures based upon multidimensional recalibration. *Anal Chem* 2006;78:8374–8385. [PubMed: 17165830]

22. 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]
23. Qian WJ, Liu T, Monroe ME, Strittmatter EF, Jacobs JM, Kangas LJ, Petritis K, Camp DG, Smith RD. Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: The human proteome. *J. Proteome Res* 2005;4:53–62. [PubMed: 15707357]
24. Wang Y, Klemke RL. Phosphoblast: A computational tool for comparing phosphoprotein signatures among large datasets. *Mol Cell Proteomics*. 2007Epub
25. Swarthout JT, Walling HW. Lysophosphatidic acid: receptors, signaling and survival. *Cell Mol Life Sci* 2000;57:1978–1985. [PubMed: 11215523]
26. Verhoeven K, De Jonghe P, Van de Putte T, Nelis E, Zwijsen A, Verpoorten N, De Vriendt E, Jacobs A, Van Gerwen V, Francis A, Ceuterick C, Huylebroeck D, Timmerman V. Slowed conduction and thin myelination of peripheral nerves associated with mutant rho Guanine-nucleotide exchange factor 10. *Am J Hum Genet* 2003;73:926–932. [PubMed: 14508709]
27. Pereira-Leal JB, Seabra MC. The mammalian Rab family of small GTPases: definition of family and subfamily sequence motifs suggests a mechanism for functional specificity in the Ras superfamily. *J Mol Biol* 2000;301:1077–1087. [PubMed: 10966806]
28. Kamei T, Matozaki T, Takai Y. [Mechanisms of cell adhesion and migration]. *Gan To Kagaku Ryoho* 1999;26:1359–1366. [PubMed: 10478193]
29. Kane S, Sano H, Liu SC, Asara JM, Lane WS, Garner CC, Lienhard GE. A method to identify serine kinase substrates. Akt phosphorylates a novel adipocyte protein with a Rab GTPase-activating protein (GAP) domain. *J Biol Chem* 2002;277:22115–22118. [PubMed: 11994271]
30. Brahmabhatt AA, Klemke RL. ERK and RhoA differentially regulate pseudopodia growth and retraction during chemotaxis. *J Biol Chem* 2003;278:13016–13025. [PubMed: 12571246]
31. Klemke RL, Cai S, Giannini AL, Gallagher PJ, de Lanerolle P, Cheresh DA. Regulation of cell motility by mitogen-activated protein kinase. *J Cell Biol* 1997;137:481–492. [PubMed: 9128257]
32. Wang Y, Ding SJ, Wang W, Jacobs JM, Qian WJ, Moore RJ, Yang F, Camp DG, Smith RD, Klemke RL. Profiling Signaling Polarity in Chemotactic Cells. *Proc Natl Acad Sci U S A* 2007;104:8328–8333. [PubMed: 17494752]
33. Zheng CF, Guan KL. Cloning and characterization of two distinct human extracellular signal-regulated kinase activator kinases, MEK1 and MEK2. *J Biol Chem* 1993;268:11435–11439. [PubMed: 8388392]
34. Campbell DH, Sutherland RL, Daly RJ. Signaling pathways and structural domains required for phosphorylation of EMS1/cortactin. *Cancer Research* 1999;59:5376–5385. [PubMed: 10537323]
35. Clarke DM, Brown MC, LaLonde DP, Turner CE. Phosphorylation of actopaxin regulates cell spreading and migration 2004;166:901–912.
36. Martinez-Quiles N, Ho HY, Kirschner MW, Ramesh N, Geha RS. Erk/Src phosphorylation of cortactin acts as a switch on-switch off mechanism that controls its ability to activate N-WASP 2004;24:5269–5280.
37. Foisner R, Bohn W, Mannweiler K, Wiche GD. a. u. o. p. a. i. s. o. r. g. C. c. d. i. i. f. p. v. e. J. S. B., 304-317, Distribution and ultrastructure of plectin arrays in subclones of rat glioma C6 cells differing in intermediate filament protein (vimentin) expression. *J Struct Biol* 1995;115:304–317. [PubMed: 8573472]
38. Foisner R, Traub P, Wiche GP. k. A.-a. p. k. C.-r. i. o. p. w. l. B. a. v. P. N. A. S. U. S. A., 3812-3816, Protein kinase A- and protein kinase C-regulated interaction of plectin with lamin B and vimentin. *Proc Natl Acad Sci U S A* 1991;88:3812–3816. [PubMed: 2023931]
39. Olski TM, Noegel AA, Korenbaum E. Parvin, a 42 kDa focal adhesion protein, related to the alpha-actinin superfamily. *J Cell Sci* 2001;114:525–538. [PubMed: 11171322]
40. Tu Y, Huang Y, Zhang Y, Hua Y, Wu C. A new focal adhesion protein that interacts with integrin-linked kinase and regulates cell adhesion and spreading. *J Cell Biol* 2001;153:585–598. [PubMed: 11331308]
41. Bryce NS, Clark ES, Leysath JL, Currie JD, Webb DJ, Weaver AM. Cortactin promotes cell motility by enhancing lamellipodial persistence. *Curr Biol* 2005;15:1276–1285. [PubMed: 16051170]

42. El Sayegh TY, Arora PD, Fan L, Laschinger CA, Greer PA, McCulloch CA, Kapus AP. o. N.-c.-a. c. b. F. k. r. N.-c. m. a. i. a. s. M. B. C., 5514-5527., Phosphorylation of N-cadherin-associated cortactin by Fer kinase regulates N-cadherin mobility and intercellular adhesion strength. *Mol Biol Cell* 2005;16:5514-5527. [PubMed: 16176974]
43. Gilles C, Polette M, Zahm JM, Tournier JM, Volders L, Foidart JM, Birembaut P. Vimentin contributes to human mammary epithelial cell migration. *J Cell Sci* 1999;112:4615-4625. [PubMed: 10574710]
44. Hadano S, Hand CK, Osuga H, Yanagisawa Y, Otomo A, Devon RS, Miyamoto N, Showguchi-Miyata J, Okada Y, Singaraja R, Figlewicz DA, Kwiatkowski T, Hosler BA, Sagie T, Skaug J, Nasir J, Brown RH, Scherer SW, Rouleau GA, Hayden MR, Ikeda JE. A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. *Nat Genet* 2001;29:166-173. [PubMed: 11586298]
45. Haydon CE, Eysers PA, Aveline-Wolf LD, Resing KA, Maller JL, Ahn NG. Identification of novel phosphorylation sites on *Xenopus laevis* Aurora A and analysis of phosphopeptide enrichment by immobilized metal-affinity chromatography. *Mol Cell Proteomics* 2003;2:1055-1067. [PubMed: 12885952]
46. Moser K, White FM. Phosphoproteomic analysis of rat liver by high capacity IMAC and LC-MS/MS. *J Proteome Res* 2006;5:98-104. [PubMed: 16396499]
47. Shi SD-H, Hemling ME, Carr SA, Horn DM, Lindh I, McLafferty FW. Phosphopeptide/Phosphoprotein Mapping by Electron Capture Dissociation Mass Spectrometry. *Anal. Chem* 2001;73(1):19-22. [PubMed: 11195502]
48. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A* 2004;101:9528-9533. [PubMed: 15210983]

Figure 1.

Schematic overview of the strategy used for quantitative phosphoproteomic analysis of Cos-7 cells in response to an LPA gradient stimulus. Protein samples from control and LPA stimulated cells were digested into peptides with trypsin and then labeled by trypsin-catalyzed ¹⁶O/¹⁸O labeling and ¹⁶O/¹⁸O-methanol esterification, respectively. Phosphopeptides enriched by IMAC from mixed ¹⁶O/¹⁸O labeled peptides were analyzed by nanoscale LC-MS²-MS³.

Figure 2.

The ratio distribution of 1:1 labeled control samples and control vs LPA stimulated samples using dual-step ^{18}O labeling. (A) The ratio distribution of 312 detected peptide pairs from 1:1 labeled control samples. (B) The ratio distribution of 384 detected phosphopeptide pairs from control vs LPA stimulated samples.

Figure 3.

Two examples reviewing phosphopeptide identification and quantification by LTQ-FT analysis. Each phosphopeptide is sequenced and consequently identified by MS² and MS³. ASAVSELS*PR was identified by the heavy form of dual ¹⁸O labeled peptide, while HKMS*PPPSGFGER was identified by the light form of dual ¹⁸O labeled peptide. The symbol * labels the monoisotopic peak for a peptide without ¹⁸O labeling, the peak with ¹⁸O labeling and the peak with a mass 2 Da lower than the monoisotopic peak with ¹⁸O-labeling; Dha, dehydroalanine.

Figure 4.

Estimation of false discovery rate (FDR) of phosphopeptide identifications using two different approaches. (A) Peptide ion mass accuracy distribution from SIM scans before (blue) and after mass correction (red). Mass accuracy typically less than 2 ppm from the SIM scans was achieved after mass correction. It was observed that the probability density of the random matches is uniform in the range ± 30 ppm, which makes it possible to estimate the portion of the histogram peak that corresponds to false identifications as a rectangular area within ± 2 ppm and the amplitude equal to the average height of the histogram wings. FDR obtained from the histogram Fig. 4a is ~ 0.01 , corresponding to $\sim 99\%$ confidence of identifications. (B) The effect of mass accuracy and Xcorr from SEQUEST analysis of an LC-MS/MS analysis using

a forward/reverse database approach. Peptide hits derived from the reverse database are shown in red, and those derived from the forward database are shown in blue. The number of false identifications obtained from the reverse database results closely agrees with the random hits evaluation based on the mass accuracy histogram.

Figure 5.

Structure and subcellular localization of a hypothetical phosphoprotein. (A) The representative motifs and functional domains of hypothetical protein H5 (IPI00298373) were predicted by ScanSite. The identified phosphorylation sites are shown in red with the corresponding phosphorylation ratio in parenthesis. (B) Confocal images of Cos-7 cells expressing hypothetical proteins fused with GFP (green) and co-stained with rhodamine phalloidin to visualize the F-actin cytoskeleton (red). The merged images of both channels illustrate the co-localization (yellow) of hypothetical proteins with F-actin. Arrows indicate the localization of H5 to actin-rich membrane ruffles at the leading edge of the pseudopodia.

Selected list of phosphopeptides identified with increased expression in LPA stimulated cells in this study.

Table 1

Protein ID	Description	Phosphorylation sites ¹⁾	Xcorr	MA ²⁾	Ratio ³⁾	SD ⁴⁾
IPI:IP100419724.1	Semaphorin C.	R.GY*QS* ¹ LSDSPGARVFTSEKR.P	2.51	0.80	5.57	0.0000
IPI:IP100158615.5	Tho Complex 2.	K.ERCT*ALQDKLLEEK.K	2.55	0.96	5.52	N/A
IPI:IP100030917.2	Growth Factor Receptor-Bound Protein 14.	R.RSGLYESTKGTGS*K.E	1.61	1.20	5.26	0.0003
IPI:IP100299517.2	Kiaa0157 Protein.	K.DIRAIQVY*NALQEK.V	1.72	0.37	5.21	0.0000
IPI:IP100171749.2	Conserved Hypothetical Protein.	R.VERPERS*GREVS*GHSVR.G	1.88	1.33	4.87	0.0000
IPI:IP100033022.2	Dynamin 2.	K.GIS*PVPINLR.V	1.59	0.98	4.68	0.0000
IPI:IP100010063.4	Kiaa1210 Protein.	K.DDMGRNAGIDFGS*RK.A	1.82	1.40	4.67	0.0000
IPI:IP100232047.3	Protein Tyrosine Phosphatase	R.Y*QVDLVPSDFVITR.D	1.63	0.83	4.63	N/A
IPI:IP100221121.1	Rho-GTPase-Activating Protein 6.	R.S*VPIQSLSELER.A	1.71	1.69	4.54	0.0001
IPI:IP100384218.1	Poi Protein.	R.QGERS*LGSR.K	1.65	0.93	4.06	0.0022
IPI:IP100152642.1	Ad158.	K.NSLS*VLSPK.I	1.61	0.39	4.01	0.0002
IPI:IP100395907.1	Fli32884 Protein.	K.LGSGQSPQTGT*PK.K	1.51	2.17	3.92	0.0003
IPI:IP100302742.6	Fli11588 Protein.	R.KPS*EVAHK.S	1.50	1.21	3.60	0.0003
IPI:IP100384286.1	Alanine-Glyoxylate Aminotransferase	R.QSQT*GSGRGPAPR.P	1.62	1.31	3.60	N/A
IPI:IP100220365.2	Homolog.	R.PS*QPEGLRKAASLT*EDR.D	1.98	1.71	3.36	0.0000
IPI:IP100397125.2	Eukaryotic Translation Initiation Factor 4 Gamma	R.RS*RLAS*VGPAAPAR.P	1.91	0.06	3.30	0.0001
IPI:IP100400878.2	Hypothetical Protein Xp_374104.	R.NTIHEFKS*AKTLIK.I	2.31	0.50	3.12	0.0000
IPI:IP100171678.2	Hepatocyte Growth Factor Precursor.	K.VVT*VLVRDRE	1.55	0.33	3.02	0.3075
IPI:IP100216704.2	Dopamine Beta-Monooxygenase Precursor.	K.PTT*LELKER.Q	1.59	0.59	2.81	0.0000
IPI:IP100374244.2	Spectrin Beta Chain	R.IAS*SHSGRVLKE	1.54	0.87	2.52	N/A
IPI:IP100028031.1	Cg12206-Pa.	K.LWIS*NGGLADIFT*VFAK.T	2.09	1.01	2.49	0.0012
IPI:IP100418277.2	Acyl-CoA Dehydrogenase	R.DGRPGS*S*HNGSGDGAAPS*AR.P	1.78	0.65	2.44	0.0000
IPI:IP100454788.1	Chondroitin Sulfate Synthase 3.	R.NDTRVT*HQNVPRL	1.85	1.38	2.43	0.0000
IPI:IP100181265.1	Hypothetical Protein Xp_499130.	R.NSVASPTSPT*R.S	1.52	0.77	2.13	0.0002
IPI:IP100377085.1	Jm11 Protein.	R.NLAEDAS*STEAAGGYKVVVRK.A	2.20	0.55	2.10	0.0000
IPI:IP100183349.6	Ahrgef11 Protein.	K.S*GCNKNS*VLVKKP.K	1.80	2.15	2.07	0.0004
IPI:IP100295437.2	R31155_1.					
IPI:IP100552892.1	Laminin Beta-4 Chain.	K.GNLS*LERLK.Q	1.63	1.06	2.01	0.0006
IPI:IP100019226.1	114 Kda Protein.	R.LDFGS*AVPLRTR.V	1.83	0.02	1.95	N/A
IPI:IP100032955.1	Bromodomain-Containing Protein 8.	R.PMDLTS*LKRNLKGR.I	1.67	1.48	1.81	N/A
IPI:IP100293375.3	Zinc Finger Protein 313.	K.Y*QNY*MEGVKATIK.D	2.09	2.13	1.64	0.0000
IPI:IP100411901.2	Hypothetical Protein Fli13150.	R.KAGAPRCS*RK.A	1.56	1.42	1.58	0.0000
IPI:IP100063160.1	Orthump0000060196.	K.RPOT*PKIDYLLPGPGPAHS*PQPSKR.A	2.70	0.27	1.55	0.1666
IPI:IP100006444.1	Hypothetical Protein Mgc15763.	R.IEAASRLTSLT*TLR.H	1.91	0.83	1.46	0.0006
IPI:IP100550920.2	Sodium/Potassium/Calcium Exchanger 2	K.YYDT*MTEGRFREKASILHK.I	2.74	0.94	1.42	N/A
IPI:IP100044683.1	Coactivator Activator.	R.DRS*PLRRS*PPRASYVAPLTAQPATYR.A	3.08	0.20	1.41	0.2087
IPI:IP100215965.1	Als2cr4 Protein.	R.T*KNTPAS*ASLEGLAQTAGRR	1.66	1.04	1.39	N/A
IPI:IP100220038.1	Heterogeneous Nuclear Ribonucleoprotein A1	K.SES*PKEPEQLRKL	2.56	0.01	1.35	0.4076
IPI:IP100012750.3	Arsenite-Resistance Protein 2.	R.TQLWASEPGT*PPLPTSLPSQNLIK.N	2.96	0.65	1.20	N/A
IPI:IP100478694.1	40s Ribosomal Protein S25.	K.DPVNKS*GGKAKK.K	2.52	0.32	1.20	N/A
IPI:IP100418385.3	14 Kda Protein.	K.DPVNKS*GGKAKK.K	2.52	0.32	1.20	N/A
IPI:IP10022145.5	13 Kda Protein.	R.LSSRRVY*ER.M	1.58	0.79	1.17	0.0005
	Nuclear Ubiquitous Casein And Cyclin-Dependent Kinases Substrate.	K.ATVTPS*PVK.G	1.69	0.05	1.17	N/A

IP1:IP100304692.1	Heterogeneous Nuclear Ribonucleoprotein G.	R.DVYLS*PR.D	1.69	0.89	1.10	0.0002
IP1:IP100010700.1	Large Proline-Rich Protein Bat2.	K.LIPGPLS*PVAR.G	1.90	0.48	1.01	0.0002

1) “*” labels the phosphorylated amino acid in a peptide sequence

2) MA refers to mass accuracy, ppm

3) The presented ratios is averaged log2 Ratio (LPA/control) across four technical replicates

4) SD refers to standard deviation for the observed log2 ratio across the technical replicates

Table 2

Motif analysis of our phosphorylation dataset

Kinase	Motif	Number
Proline Directed	(S/T)*P	204
Basophilic2	(R K).(S/T)*	61
14-3-3 sub1	R.{3}S*	16
GSK3	(S/T)*.{3}(S/T)*	7
CK2	(S/T)*.{2}(D E)	43
Basophilic1	(R K)(S/T)*	62
Basophilic3	(R K).{2}(S/T)*	37
CDK1/2/5	(S/T)*P.(R K H)	50
Akt/RSK	(R K).(R K).{2}(S/T)*	1
PDK1	(S/T)*(Y F)	15
CK1	D.{2}(S/T)*	18
PDK2	(Y F)(S/T)*(Y F)	1
14-3-3 sub2	R.{2}S*	16
Erk1/2	P.(S/T)*P	32