

## Combining Protein-Based IMAC, Peptide-Based IMAC, and MudPIT for Efficient Phosphoproteomic Analysis

Greg T. Cantin,<sup>†,§,△</sup> Wei Yi,<sup>‡,§</sup> Bingwen Lu,<sup>†</sup> Sung Kyu Park,<sup>†</sup> Tao Xu,<sup>†</sup> Jiing-Dwan Lee,<sup>\*,‡</sup> and John R. Yates III<sup>\*,†</sup>

*Departments of Cell Biology and Immunology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037*

Received August 18, 2007

**Abstract:** Immobilized metal affinity chromatography (IMAC) is a common strategy used for the enrichment of phosphopeptides from digested protein mixtures. However, this strategy by itself is inefficient when analyzing complex protein mixtures. Here, we assess the effectiveness of using protein-based IMAC as a pre-enrichment step prior to peptide-based IMAC. Ultimately, we couple the two IMAC-based enrichments and MudPIT in a quantitative phosphoproteomic analysis of the epidermal growth factor pathway in mammalian cells identifying 4470 unique phosphopeptides containing 4729 phosphorylation sites.

**Keywords:** phosphoproteome • IMAC • SILAC • EGF

### Introduction

Protein phosphorylation is a dynamic and important event occurring within the cell. It has been shown that phosphorylation can affect a proteins enzymatic activity, its ability to interact with other proteins and molecules, its subcellular localization, and its degradation rate in a cell. Because of these affects, it is clear how phosphorylation can be a key component in many signal transduction cascades of eukaryotic cells.<sup>1</sup> One rapidly developing area of research is global phosphorylation (or phosphoproteomic) analysis of cells, tissues, and organisms using mass spectrometry. From these pursuits, hundreds to thousands of phosphorylation sites can now routinely be identified, quantified, and assigned to a particular cellular state or process.

Liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) is a powerful method utilized in proteomics research. In a typical proteomic analysis, an extracted protein mixture is often digested with a sequence-specific protease and subjected to LC-MS/MS, resulting in the identification of up to thousands of proteins in a single analysis. Our laboratory utilizes multidimensional liquid chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS), also termed MudPIT, to enhance the number of proteins that

can be identified in a single analysis.<sup>2</sup> However, these LC-MS strategies by themselves are not sufficient for carrying out a global cellular analysis of phosphorylation. This is primarily due to the enormous number and dynamic range of proteins in a cell, which exacerbates the limitations in the resolving capabilities of chromatographic methods as well as the speed and sensitivity of the mass spectrometer. Thus, most phosphoproteomic strategies utilize some form of phosphoprotein/phosphopeptide enrichment step prior to LC-MS analysis in order to enhance the detection of phosphorylation sites within a complex protein mixture.<sup>3,4</sup>

Many different enrichment methods that target the phosphate moiety have been utilized in phosphoproteomic studies. One approach includes the enrichment of phosphoproteins using immobilized metal affinity chromatography (IMAC)<sup>5,6</sup> or phosphorylation-specific antibodies;<sup>7,8</sup> subsequently, protein digestion and LC-MS analysis is carried out. A second type of method includes the enrichment of phosphopeptides after proteolytic digestion of a complex protein mixture. In this second approach, a variety of separation techniques have been used, such as ion exchange chromatography,<sup>9,10</sup> isoelectric focusing,<sup>11</sup> IMAC,<sup>12,13</sup> TiO<sub>2</sub>,<sup>14</sup> and immunoaffinity.<sup>15</sup> Many recent studies have utilized multiple phosphoprotein/phosphopeptide enrichment strategies in succession in order to further enhance the detection of phosphorylation sites in complex protein mixtures.<sup>6,8,10,11,16,17</sup>

Here, we report on a phosphoproteomic method consisting of IMAC-based phosphoprotein enrichment followed by IMAC-based phosphopeptide enrichment, generating a highly enriched pool of phosphopeptides that is analyzed by MudPIT. This method is applied to the quantitative phosphoproteomic analysis of the EGF pathway in HeLa cells using stable isotope labeling by amino acids in cell culture (SILAC).<sup>18,19</sup> The number of phosphorylation sites identified and the amount of time in which the analysis can be carried out demonstrate the effectiveness and efficiency of the method relative to other recently published studies.

### Experimental Section

**Cell Culture and SILAC-Labeling.** GIBCO SILAC D-MEM basal cell culture medium (Invitrogen) containing 2 mM L-Glutamine, 10% dialyzed fetal bovine serum (FBS), and 100 U/mL penicillin and streptomycin was supplemented with 100 mg/L L-lysine and 20 mg/L L-arginine or 100 mg/L [U-<sup>13</sup>C<sub>6</sub>]-L-lysine and 20 mg/L [U-<sup>13</sup>C<sub>6</sub>, <sup>15</sup>N<sub>4</sub>]-L-arginine (Invitrogen) to

\* To whom correspondence should be addressed by e-mail: (J.R.Y.) jyates@scripps.edu, (J.-D.L.) jdlee@scripps.edu.

<sup>†</sup> Department of Cell Biology, The Scripps Research Institute.

<sup>§</sup> These authors contributed equally to this work.

<sup>△</sup> Present address: The Dow Chemical Company, 5501 Oberlin Dr., San Diego, CA, 92121.

<sup>‡</sup> Department of Immunology, The Scripps Research Institute.

make the “Light” SILAC or “Heavy” SILAC culture media, respectively. HeLa cells were obtained from ATCC and were propagated in SILAC medium for six generations to ensure nearly 100% incorporation of labeled amino acids before an experiment is carried out. Prior to cell treatment, both heavy and light cells were cultured overnight (O/N) in their corresponding SILAC media without FBS. For EGF-treatment, “Light” cells were treated for 15 min with 10 ng/mL of EGF (R&D Systems), using a stock solution of 20  $\mu\text{g}/\text{mL}$  in 10 mM acetic acid/0.1% bovine serum albumin. “Heavy” cells were left untreated.

**Protein Lysate Preparation, Protein-Based IMAC, Digestion, and Peptide-Based IMAC.** HeLa cells were harvested in an Extraction/Loading buffer from Clontech (TALON PMAC Phosphoprotein Enrichment Kit) with cocktail protease inhibitors (Roche) and 10 mM sodium fluoride (as a phosphatase inhibitor) as described in the user manual (PT3731-1, Clontech). Briefly, cells were washed with PBS and lysed by resuspending in Loading buffer (30  $\mu\text{L}/\text{mg}$  of cells). After lysis, the sample was centrifuged at 10 000g for 20 min at 4 °C to produce a clarified lysate. “Light” SILAC HeLa lysates and “Heavy” SILAC HeLa lysate were mixed at 1:1 [w/w] ratio, determined by Coomassie Plus protein assay (Pierce). In total, 13 mg of the 1:1 mixed protein lysate was loaded onto one phosphoprotein affinity column (Talon PMAC, 1 mL phosphorylation metal affinity chromatography resin, Clontech), capped, and rotated for 20 min at 4 °C. The column was then attached to a ring stand at room temperature, and the flow through was discarded. The column was washed four times with 5 mL of loading buffer and eluted with 5 mL of Elution buffer (20 mM sodium phosphate and 0.5 M sodium chloride). Approximately 14% (1.8 mg) of the loaded protein was eluted from the metal chelate resin (as expected for a HeLa cell lysate, based on the Clontech manual). The eluate was reduced by adding DTT to 5 mM and incubating at 56 °C for 45 min, and then alkylated by adding iodoacetamide to 25 mM and incubating at 37 °C for 30 min.

The resulting solution was dialyzed with membrane tubing (3.5 kDa MWCO) against 1 M urea/100 mM  $\text{NH}_4\text{HCO}_3$  at 4 °C O/N. To the dialyzed sample, 30  $\mu\text{g}$  of trypsin (Promega, sequencing grade) was added (~1:60, enzyme/substrate) and incubated at 37 °C O/N. The digestion was subjected to solid-phase extraction, where it was loaded onto a 500 mg Extract-clean SPE C18 column (Alltech), washed with 0.1% TFA, eluted in 70% acetonitrile/0.1%TFA, and dried via speed-vac.

The lyophilized peptide sample was then reconstituted in 100  $\mu\text{L}$  of 0.1% formic acid/2% acetonitrile and combined with an equal volume of peptide IMAC binding buffer (5% acetic acid). The final 200  $\mu\text{L}$  peptide solution was split into four 50  $\mu\text{L}$  aliquots and loaded onto four IMAC spin columns, containing 25  $\mu\text{L}$  of gallium-chelated resin (Phosphopeptide Isolation Kit, Pierce), and mixed intermittently for ~5 min, and the flow through was collected by centrifugation at ~720 g for 1 min. The resin was then washed two times by the addition of 50  $\mu\text{L}$  of Wash Buffer 1 (0.1% acetic acid) and followed by centrifugation. The resin was then washed two times by the addition of 50  $\mu\text{L}$  of Wash Buffer 2 (0.1% acetic acid and 10% acetonitrile) and subsequent centrifugation. The resin was then washed once by the addition of 75  $\mu\text{L}$  of water and followed by centrifugation. The IMAC-retained peptides were then subjected to three replicate 40  $\mu\text{L}$  elutions of 100 mM  $\text{NH}_4\text{HCO}_3$ , pH 9, and 5% acetonitrile for each spin column. Eluates from all four spin columns were combined (~480  $\mu\text{L}$ ). Formic acid

was added to the phosphopeptide-enriched sample to a final concentration of 4% and was subsequently centrifuged to remove any particulate matter prior to analysis by MudPIT.

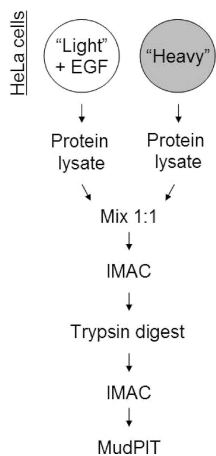
**LC–MS Analysis.** Initial analysis and optimization of the phosphoprotein and phosphopeptide-enriched samples were analyzed by LC–MS/MS and MudPIT using an LCQ Deca ion trap mass spectrometer (Thermo Electron) as described previously.<sup>11</sup> Final SILAC experiments were performed on an LTQ-Orbitrap mass spectrometer.

The following was done in the analysis of the SILAC experiment. For the back-end of the MudPIT column, ~3.5 cm of strong cation exchange resin (SCX Luna, 5  $\mu\text{m}$ , 100 Å, Phenomenex), followed by ~3.5 cm of C18 resin (Aqua, 5  $\mu\text{m}$ , 125 Å, Phenomenex) was packed within a Kasil-fritted piece of 250  $\mu\text{m}$  (i.d.) fused silica tubing (Agilent) using a pressure-loading device. This column was equilibrated with buffer A (5% acetonitrile and 0.1% formic acid), loaded with sample, and then washed with buffer A. With the use of a zero-dead-volume union (Upchurch), the loaded back-end column was then connected to the analytical column, consisting of ~11 cm of packed C18 (Aqua, 3  $\mu\text{m}$ , 125 Å, Phenomenex) in 100  $\mu\text{m}$  (i.d.) fused silica (Polymicro) with a laser-pulled tip (5–15  $\mu\text{m}$ ). The combined column was subjected to LC/LC–MS analysis essentially as previously described,<sup>20</sup> using an HPLC (Nano LC-2D, Eksigent) online with the hybrid LTQ-Orbitrap mass spectrometer (Thermo Electron). However, the composition of buffers in this study are slightly different for buffer A (5% acetonitrile, 0.1% formic acid), B (80% acetonitrile, 0.1% formic acid), and C (500 mM ammonium acetate, 5% acetonitrile, and 0.1% formic acid). A 9-step LC/LC–MS analysis was carried out. The first step elutes all peptides bound to the back-end C18 resin onto the SCX resin via a 40 min reversed-phase gradient of 2–98% buffer B. Steps 2–9 include salt steps of 5, 10, 20, 30, 40, 50, 98, and 98% buffer C (5 min, 500 nL/min), followed by a reversed-phase gradient of 2–35% buffer B over 130 min, then 35–75% buffer B over 15 min at a flow rate of 300 nL/min.

The instrument method consisted of a full MS scan in the Orbitrap from 400 to 1800  $m/z$ , AGC target of 500 000, max. ion injection time of 500 ms, with 1  $\mu\text{scan}$ , a resolution of 60 000, and with preview scan on. Data-dependent MS/MS (MS2) and MS/MS/MS (MS3) were performed in the linear ion trap on the five most intense ions. AGC targets of 30 000 and 10 000 and max. ion injection times of 50 and 100 ms were used for MS and MS<sup>n</sup> scans, respectively; and each with 1  $\mu\text{scan}$ . MS3s were triggered for precursor ions undergoing a neutral-loss of either 98, 49, or 32.7  $m/z$  upon MS/MS.<sup>9,10</sup> The minimum signals required for data-dependent MS2 and MS3 were 1000 and 100, respectively; and an isolation width of 3  $m/z$  and normalized collision energy of 35% was used for both. Additionally, dynamic exclusion was used with an exclusion list of 100 and exclusion duration of 80 s.

RAW files were split into MS1, MS2, and MS3 files using the program Raw Extract (written by John Venable, T.S.R.I.). MS2 files were filtered using the previously described program,<sup>21</sup> and MS3 files were filtered using the program filter\_ms2 (written by Daniel Cociorva, T.S.R.I.). The MS2 and MS3 files were then separately subjected to SEQUEST searches. To construct the database used in these searches, the EBI-IPI human database (version 3.17 released 05–09–2006) was first attached to a list of frequently occurring contaminants (e.g., proteases and keratins); then, this database was combined with its reversed version. For MS2 searches, serine, threonine, and tyrosine differential phosphorylation was considered, a maxi-

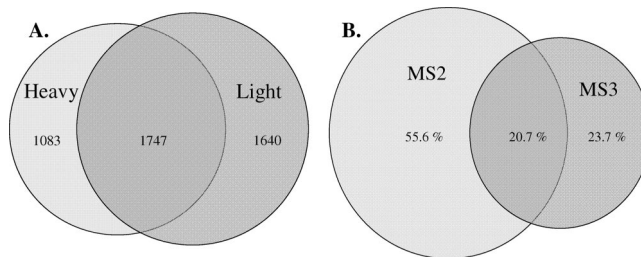
Scheme 1



num of three phosphorylations per peptide was allowed, only tryptic peptides were considered, and a 50 ppm mass tolerance was used for candidate precursors. For MS3 searches, a loss of 18 Da on serine and threonine and phosphorylation (+80) on tyrosine were searched for differentially; a maximum of two modification types per peptide was allowed, no enzyme specificity was used, and a 3 amu tolerance was used for candidate precursors. Two separate searches for each MS2 and MS3 data file were carried out; light and heavy versions of both lysine and arginine were searched for separately using fixed mass differences for the two amino acids. Cysteine was considered to have a static modification of +57.0246 mass units.

In filtering the search results, a feature in DTASelect (DTA2.0) was used that reports the DeltaCN based on the best peptide-spectral match of the highest scoring peptide with an unrelated amino acid sequence compared to the top scoring modified peptide (as previously described<sup>11</sup>). The SEQUEST identifications were additionally filtered using a false-positive cutoff based on XCorr and DeltaCN scores and mass accuracy (MS2 search only) using a reversed database approach.<sup>22</sup> First, SEQUEST identifications were grouped into categories by charge state (+1, +2, and +3), and subcategories of modification status, and tryptic status (for the MS3 search only). For each category, the identifications were filtered at a false-positive rate of 1% (based on a combination of XCorr, DeltaCN, and mass accuracy (MS2 search only) for spectra matching the reversed database). For the final reported numbers, identifications from the reversed database and the contaminants list were removed. The delta mass distribution of phosphorylated peptides identified in the MS2 searches had a mean of -1.198 ppm and standard deviation of 2.074 ppm and is displayed in Supplementary Figure 1 in Supporting Information.

**Determination of Percent Isotope Enrichment.** We investigated the degree of heavy-labeled amino acid incorporation into proteins in cells grown in the “Heavy” SILAC media. To “Heavy” labeled cells, cultured as described above, we prepared a protein extract as detailed above and subsequently precipitated the protein content via addition of 8 vol of acetone. The dried protein precipitate was digested with trypsin as above, and 50 μg of this was subjected to MudPIT analysis as described above on an LTQ-Orbitrap. Upon manual inspection of the raw data, the degree of enrichment of “heavy” arginine and lysine was ≥97% (similar to that for the original atomic percent enrichment of each supplemented amino acid). Additionally,



**Figure 1.** Summary of phosphopeptides identified in the EGF pathway analysis. (A) Proportional Venn diagram of the unique phosphopeptides identified from “heavy”, “light”, or both “heavy and light” SEQUEST searches. “Light” phosphopeptides originated from EGF-treated cells, and “heavy” phosphopeptides originated from untreated cells. MS2 and MS3 identifications are combined for this representation. Multiple charge states of the same phosphopeptide sequence are counted as one peptide. (B) Comparison of MS2 and MS3 phosphopeptide identifications. Heavy and light phosphopeptide identifications were combined for both MS2 and MS3 spectra. A proportional Venn diagram is used to illustrate the average distribution of phosphopeptide identifications arising from MS2, MS3, or both MS2 and MS3 spectra.

the incorporation of “heavy” proline (+6 amu), due to metabolic conversion of arginine to proline, was virtually undetectable (data not shown). These observations are similar to what has previously been observed in a study that included a similar labeling strategy.<sup>23</sup>

**Quantification Analysis.** All identified phosphopeptides were subjected to relative quantification analysis using the program Census (Park, S. K. et al., manuscript submitted for publication). This program quantifies relative abundances of light and heavy versions of precursor peptides identified by MS2 and MS3 spectra using linear regression and does so in a similar manner as the previously used program RelEx.<sup>24</sup>

**Results and Discussion**

Our goal was to develop a method that would generate a highly enriched phosphopeptide sample from a mammalian whole cell lysate. Analyzing a whole cell lysate by tryptic digestion and LC/LC–MS/MS (MudPIT) yields relatively few phosphopeptide identifications (previous observations, data not shown). Alternatively, by subjecting a tryptic-digested whole cell lysate to IMAC, an increase in phosphopeptide identifications is achieved. However, the number of phosphopeptide identifications yielded from this one-step enrichment is inferior to other recently published methods (data not shown). As a result, we decided to assess the use of IMAC at the protein level as a pre-enrichment strategy prior to the use of peptide-based IMAC.

A whole cell lysate was prepared from HeLa cells stimulated with epidermal growth factor (EGF). The lysate was then subjected to protein-based IMAC, and the eluate was digested and analyzed by MudPIT. Only 11 phosphopeptides were identified in the sample, even though ~1400 proteins were identified (from ~4800 peptides), Supplementary Table 1 in Supporting Information (S-Table 1). Though this is a proven strategy for the enrichment of phosphoproteins, the IMAC eluate is still too complex to identify a significant number of phosphopeptides by any LC–MS strategy. We next assessed the effectiveness of coupling this phosphoprotein enrichment strategy with peptide-based IMAC, together referred to as double-IMAC. Eluate from the above protein-based IMAC was

**Table 1.** Automatic Validation of MS2-Derived Phosphoserine and Phosphothreonine-Containing Peptides Using the Program DeBunker<sup>a</sup>

	phosphopeptide copies	pY	pS and pT		
			pS and pT	pass DeBunker filter	percentage passing DeBunker
Heavy MS2	2352	29	2323	2224	95.7%
Light MS2	2947	47	2900	2792	96.3%

<sup>a</sup> Multiple charge states are included.

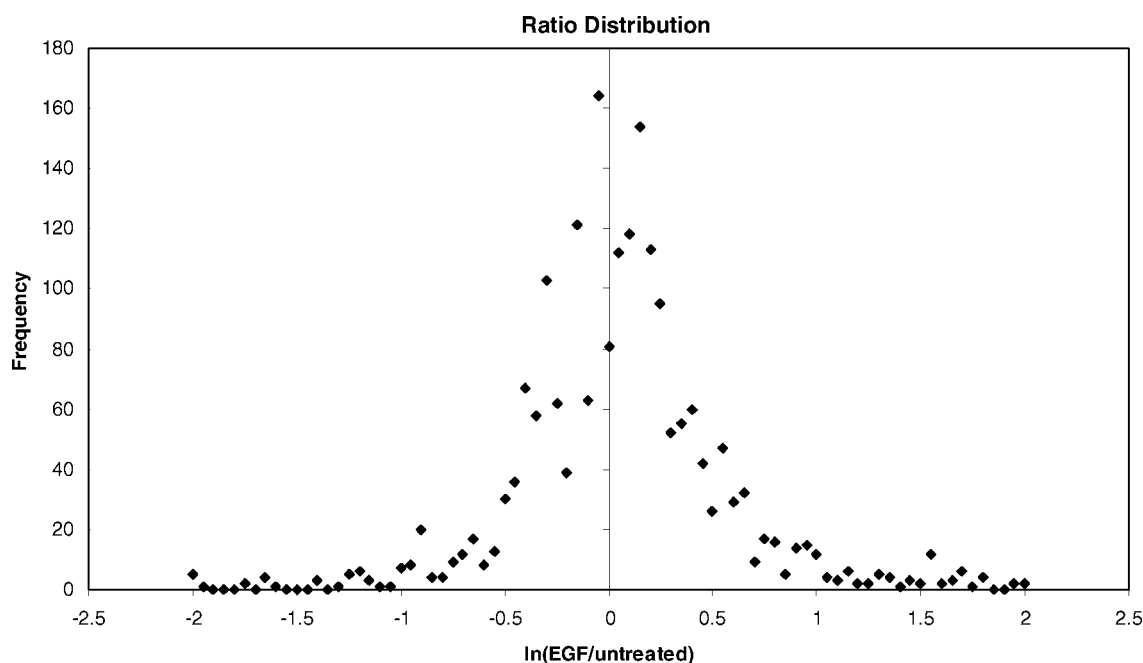
digested with trypsin and subjected to peptide-based IMAC. Different amounts of this loading material were tested for the peptide-based IMAC, and each eluate was analyzed by single-dimension LC-MS/MS. The results, displayed in Supplementary Table 2 in Supporting Information (S-Table 2), show a dramatic increase in the number of phosphopeptides and in the ratio of phosphopeptides to unmodified peptides identified. Additionally, we determined the optimal amount of the above double-IMAC sample to use for analysis by MudPIT, the results of which are displayed in Supplementary Table 3 in Supporting Information (S-Table 3). Altogether, it is clear that coupling protein-based IMAC with peptide-based IMAC is far superior to either method alone (based on above data and previous observations of peptide-based IMAC alone using a slightly different source material). This general observation has been made before,<sup>6</sup> though we would like to note that the study utilized a less complex sample, different IMAC loading methods, and produced a less dramatic effect from combining the two IMAC enrichment steps compared to our study.

After its assessment and optimization, we applied our double-IMAC method to a quantitative analysis of the EGF-pathway in HeLa cells, outlined in Scheme 1. We employed the use of stable isotope labeling in cell culture (SILAC) in order to make the analysis quantitative. One culture was grown in

media containing “light” versions of lysine and arginine, and a second culture was grown in media containing “heavy” versions of lysine and arginine (see Experimental Section for details). The “light” culture was stimulated with EGF and the “heavy” culture was left untreated. Protein lysates were separately prepared and then mixed 1:1 [w/w] based on total protein. The 1:1 (light/heavy) mixture was then subjected to our double-IMAC procedure described above and outlined in Scheme 1.

For the final step of this analysis, MudPIT (~24 h) was carried out using a hybrid LTQ Orbitrap mass spectrometer, where data-dependent MS/MS (MS2) and MS/MS/MS (MS3) spectra were collected as described in the Experimental Section. Briefly, the primary MS scan was carried out in the Orbitrap, followed by data-dependent MS2 and neutral loss-triggered MS3 (specific for the loss of phosphoric acid, from phosphoserine and phosphothreonine-containing peptides) carried out in the linear ion trap. MS2 and MS3 spectra were searched separately using SEQUEST as described in the Experimental Section. When the strict filtering of a 1% false-positive rate was used (see Experimental Section for more details), 4470 unique phosphopeptides were identified (see below for further description). Similar to what was seen in the optimization of the double-IMAC procedure, ~70% of the identifications were phosphopeptides and ~30% were unmodified peptides (Supplementary Table 4 in Supporting Information S-Table 4).

A total of 6270 phosphopeptides were identified from the combined “light” and “heavy” SEQUEST searches. However, when comparing these “light” and “heavy” phosphopeptides, an expected redundancy is seen. The distribution of the identified phosphopeptides originating from the “light”-EGF-treated sample, the “heavy”-untreated sample, and both is displayed in Figure 1A. Approximately 40% of the identified phosphopeptides were found as both “light” and “heavy” versions in the experiment. As a result, 4470 were observed.



**Figure 2.** Relative quantification analysis. All MS2- and MS3-derived phosphopeptide identifications were subjected to relative quantification analysis using the program Census. Of the identifications passing the quantification filters, a frequency distribution plot was generated. Plotted on the x-axis is the natural log of the ratio of light (EGF-derived) over heavy (untreated-derived) precursor phosphopeptide abundance. Plotted on the y-axis is the number of occurrences for a particular ln(ratio).

When only the amino acid sequence of these phosphopeptides are considered, a total of 3323 unique amino acid sequences (peptides) are present, containing 4729 unique phosphorylation sites. Interestingly, there were significantly more “light”-only phosphopeptides identified than “heavy”-only. This is not surprising, since one would expect to find significantly more phosphopeptides originating from the EGF-treated cells due to the activation of a signal transduction cascade that involves numerous phosphorylation events. We also looked at the distribution of MS2 versus MS3-derived phosphopeptide identifications. One can see (Figure 1B) that ~21% of the identified phosphopeptides were identified from both MS2 and MS3 spectra. Additionally, ~24% of the phosphopeptide identifications were derived from MS3 spectra alone. Together, this illustrates that the neutral loss-triggered MS3 portion of the MS method we used is a worthwhile inclusion for the analysis of this type of phosphopeptide-enriched sample.

Although we used a stringent false-positive cutoff rate, we wanted to further validate the quality of the phosphopeptide identifications from the analysis of the EGF pathway. Thus, we carried out a postsearch analysis of the MS2-derived phosphoserine and phosphothreonine-containing peptide identifications using the program DeBunker.<sup>25</sup> This program assesses the quality of the aforementioned identifications by looking for characteristic spectral features associated with these identifications (e.g., a neutral loss of phosphoric acid from the precursor ion selected for CID and from the b and y fragment ions within MS2 spectra). From this assessment, a probability can be assigned that relates to the likelihood that the phosphopeptide-spectral match is typical of most phosphoserine and phosphothreonine-containing peptides. The results of the DeBunker analysis of our data, displayed in Table 1, show that ~96% of these MS2-derived phosphopeptide identifications have a high likelihood of being a phosphopeptide. These results demonstrate that the majority of these phosphopeptide identifications are derived from high quality phosphoserine and phosphothreonine-specific MS2 spectra.

Identified phosphopeptides from the EGF pathway analysis experiment were subjected to relative quantification using the program CensuS. The relative abundance of all identified phosphopeptides is calculated by examining the chromatographic peaks of both the “light” and “heavy” precursor peptides corresponding to an MS2- or MS3-based identification. Displayed in Figure 2 is the distribution of the relative changes found for all of the quantified phosphopeptides. This figure plots the natural log ratio of the relative abundance for a phosphopeptide (*x*-axis) by the frequency of occurrence (*y*-axis) at a particular ratio. The distribution is normal and centered near zero, indicating that the majority of the phosphopeptides quantified have a relative abundance ratio (light/heavy) near or at one; meaning that those corresponding sites of phosphorylation are not regulated by EGF. Of the 4470 nonredundant phosphopeptides identified, 37% were quantifiable by CensuS (Supplementary Table S-5 in Supporting Information). This is similar to previous quantification efficiencies for a global phosphoproteomic analysis,<sup>11</sup> and is due to many factors including signal-to-noise and peak shape of the extracted ion chromatograms for both “light” and “heavy” versions of each phosphopeptide identified. For those quantified by CensuS, 15% were found to be up or down-regulated by 2-fold or more (Table S-5 in Supporting Information). For all quantification results and

corresponding DeBunker scores, see Tables S-6 and S-7 in Supporting Information. Further examination of the proteins and phosphorylation sites found to be regulated by EGF in this study is underway and will be presented in a future publication.

## Conclusion

Considering both the number of phosphopeptides identified and the amount of LC-MS analysis time used, we believe that our double-IMAC method is comparable to the most effective phosphoproteomic methods published to date.<sup>10,16,17,26</sup> Additionally, the enrichment procedures used in our method are simple, fast, and do not require specialized equipment or training, and can thus be performed in most biological laboratories prior to MS analysis. Altogether, this method allows a manageable timetable in carrying out global analyses of phosphorylation events involved in the signal transduction cascades of mammalian cells.

**Acknowledgment.** G.T.C. was supported by an NIH National Research Service Award fellowship. B.L. is supported by CFFT computational fellowship BALCH05X5. S.K.P. is supported by NIH Grant No. UOM/DMID-BAA-03. This work was additionally supported by NIH grants 5R01 MH067880 and P41 RR11823 awarded to J.R.Y.

**Supporting Information Available:** Tables of the evaluation of protein-based IMAC using EGF-stimulated cells (S-Table 1), the optimization of the peptide-based IMAC step of the double-IMAC sample preparation procedure (S-Table 2), optimization of the amount of peptide-based IMAC eluates to run via MudPIT (S-Table 3), summary of peptides identified in the EGF pathway analysis (S-Table 4), summary of phosphopeptides quantified in the EGF pathway analysis (S-Table 5), CensuS results (S-Table 6), CensuS results with corresponding DeBunker score (S-Table 7); and figure of Delta mass distributions of the peptide identifications from MS2 searches (S-Figure1). This information is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- (1) Hunter, T. *Cell* **2000**, *100*, 113–127.
- (2) Washburn, M. P.; Wolters, D.; Yates, J. R., III *Nat. Biotechnol.* **2001**, *19*, 242–247.
- (3) Loyet, K. M.; Stults, J. T.; Arnott, D. *Mol. Cell. Proteomics* **2005**, *4*, 235–245.
- (4) Reinders, J.; Sickmann, A. *Proteomics* **2005**, *5*, 4052–4061.
- (5) Andersson, L.; Porath, J. *Anal. Biochem.* **1986**, *154*, 250–254.
- (6) Collins, M. O.; Yu, L.; Coba, M. P.; Husi, H.; Campuzano, I.; Blackstock, W. P.; Choudhary, J. S.; Grant, S. G. *J. Biol. Chem.* **2005**, *280*, 5972–5982.
- (7) Pandey, A.; Podtelejnikov, A. V.; Blagoev, B.; Bustelo, X. R.; Mann, M.; Lodish, H. F. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 179–184.
- (8) Salomon, A. R.; Ficarro, S. B.; Brill, L. M.; Brinker, A.; Phung, Q. T.; Ericson, C.; Sauer, K.; Brock, A.; Horn, D. M.; Schultz, P. G.; Peters, E. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 443–448.
- (9) Beausoleil, S. A.; Jedrychowski, M.; Schwartz, D.; Elias, J. E.; Villen, J.; Li, J.; Cohn, M. A.; Cantley, L. C.; Gygi, S. P. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12130–12135.
- (10) Gruhler, A.; Olsen, J. V.; Mohammed, S.; Mortensen, P.; Faergeman, N. J.; Mann, M.; Jensen, O. N. *Mol. Cell. Proteomics* **2005**, *4*, 310–327.
- (11) Cantin, G. T.; Venable, J. D.; Cociorva, D.; Yates, J. R., III *J. Proteome Res.* **2006**, *5*, 127–134.
- (12) Ficarro, S. B.; McClelland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M. *Nat. Biotechnol.* **2002**, *20*, 301–305.
- (13) Posewitz, M. C.; Tempst, P. *Anal. Chem.* **1999**, *71*, 2883–2892.

- (14) Pinkse, M. W.; Uitto, P. M.; Hilhorst, M. J.; Ooms, B.; Heck, A. J. *Anal. Chem.* **2004**, *76*, 3935–3943.
- (15) Rush, J.; Moritz, A.; Lee, K. A.; Guo, A.; Goss, V. L.; Spek, E. J.; Zhang, H.; Zha, X. M.; Polakiewicz, R. D.; Comb, M. J. *Nat. Biotechnol.* **2005**, *23*, 94–101.
- (16) Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. *Cell* **2006**, *127*, 635–648.
- (17) Villen, J.; Beausoleil, S. A.; Gerber, S. A.; Gygi, S. P. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1488–1493.
- (18) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. *Mol. Cell. Proteomics* **2002**, *1*, 376–386.
- (19) Zhu, H.; Pan, S.; Gu, S.; Bradbury, E. M.; Chen, X. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 2115–2123.
- (20) McDonald, W. H.; Ohi, R.; Miyamoto, D. T.; Mitchison, T. J.; Yates, J. R., III *Int. J. Mass Spectrom.* **2002**, *219*, 245–251.
- (21) Bern, M.; Goldberg, D.; McDonald, W. H.; Yates, J. R., III *Bioinformatics* **2004**, *1*, I49–I54, 20 Suppl.
- (22) Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. *J. Proteome Res.* **2003**, *2*, 43–50.
- (23) Everley, P. A.; Bakalarski, C. E.; Elias, J. E.; Waghorne, C. G.; Beausoleil, S. A.; Gerber, S. A.; Faherty, B. K.; Zetter, B. R.; Gygi, S. P. *J. Proteome Res.* **2006**, *5*, 1224–1231.
- (24) MacCoss, M. J.; Wu, C. C.; Liu, H.; Sadygov, R.; Yates, J. R., III *Anal. Chem.* **2003**, *75*, 6912–6921.
- (25) Lu, B.; Ruse, C.; Xu, T.; Park, S. K.; Yates, J., III *Anal. Chem.* **2007**, *79*, 1301–1310.
- (26) Bodenmiller, B.; Mueller, L. N.; Mueller, M.; Domon, B.; Aebersold, R. *Nat. Methods* **2007**, *4*, 231–237.

PR0705441